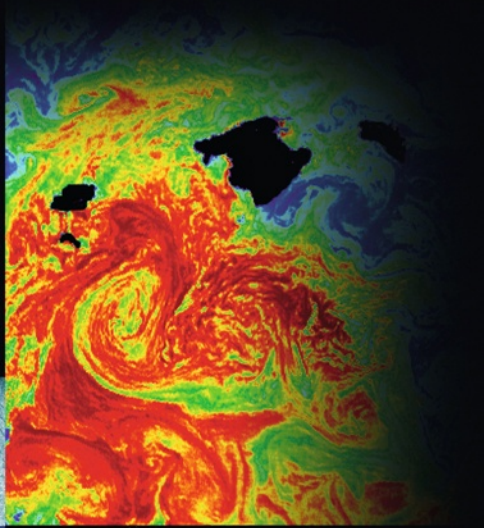
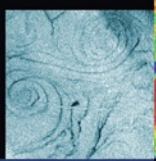
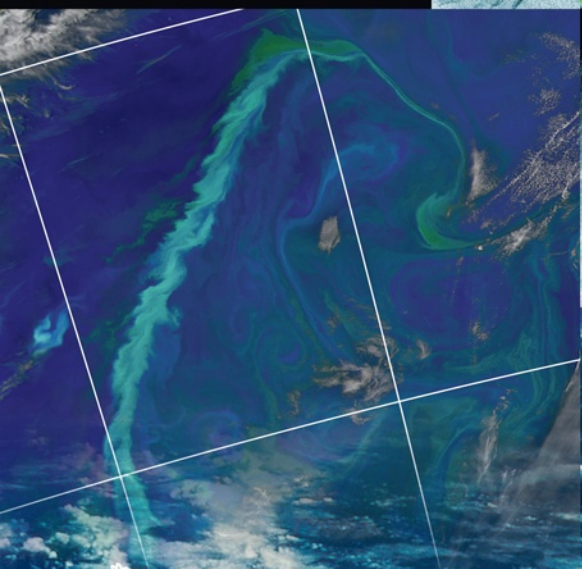
Ian S. Robinson

DISCOVERING THE OCEAN FROM SPACE

The unique applications of satellite oceanography







Springer

PRAXIS

Discovering the Ocean from Space

The Unique Applications of Satellite Oceanography

Ian S. Robinson

Discovering the Ocean from Space

The Unique Applications of Satellite Oceanography

Springer

Published in association with
Praxis Publishing
Chichester, UK

PRAXIS

Professor Ian S. Robinson
School of Ocean & Earth Science
University of Southampton
National Oceanography Centre
European Way
Southampton
UK

SPRINGER–PRAXIS BOOKS IN GEOPHYSICAL SCIENCES
SUBJECT *ADVISORY EDITOR*: Philippe Blondel, C.Geol., F.G.S., Ph.D., M.Sc., F.I.O.A., Senior Scientist, Department of Physics, University of Bath, Bath, UK

ISBN 978-3-540-24430-1 e-ISBN 978-3-540-68322-3
DOI 10.1007/978-3-540-68322-3

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010925235

© Springer-Verlag Berlin Heidelberg 2010

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: Jim Wilkie
Project management: OPS Ltd, Gt Yarmouth, Norfolk, UK

Printed on acid-free paper

Springer is part of Springer Science + Business Media (www.springer.com)

Contents

Preface xv

Abbreviations xix

List of figures xxi

List of tables xxxi

List of abbreviations and names of satellites and sensors xxxiii

List of symbols and nomenclature xli

1 Introduction 1
- 1.1 An important observational tool for planetary science 1
- 1.2 Putting remote sensing to work for oceanographers 3
- 1.3 The oceanographic scope of the book 4
- 1.4 Reference 6

2 The methods of satellite oceanography 7
- 2.1 Ocean remote-sensing techniques—a summary 7
- 2.2 The unique sampling capabilities of sensors on satellites 9
  - 2.2.1 Creating image-like data fields from point samples 9
  - 2.2.2 Satellite orbits and how they constrain remote sensing . 11
  - 2.2.3 The space-time sampling capabilities of satellite sensors . 14
- 2.3 Generic data-processing tasks 16
  - 2.3.1 Sensor calibration 17
  - 2.3.2 Atmospheric correction 19
  - 2.3.3 Positional registration 21
  - 2.3.4 Geophysical product derivation 21
  - 2.3.5 Image resampling onto map projections 23
  - 2.3.6 Composite image maps 26

2.4 Sensor types for observing the ocean 28
2.4.1 Using the electromagnetic spectrum 28
2.4.2 Ocean color radiometers 30
2.4.3 Thermal infrared radiometry for measuring sea surface temperature 35
2.4.4 Microwave radiometry 42
2.4.5 Altimetry for measuring surface slope, currents, and wave height 46
2.4.6 Oblique-viewing radars for measuring sea surface roughness 51
2.5 Platforms and sensors for satellite oceanography 54
2.6 Satellite ocean data products 54
2.7 References 66

3 Mesoscale ocean features: Eddies 69
3.1 Discovering mesoscale variability from space 69
3.2 Mesoscale ocean eddies 72
3.2.1 Eddies—ubiquitous phenomena in a turbulent ocean . . . 72
3.2.2 Lengthscales of mesoscale eddies—the Rossby radius . . 72
3.2.3 The dynamical structure of rings and eddies 76
3.3 Detecting eddies from satellites. 78
3.4 Using SSHA from altimetry to observe eddies 80
3.4.1 Revealing ocean eddies in altimeter SSHA data 80
3.4.2 Present limitations of satellite altimetry 86
3.4.3 Kinematic measurements from altimetric SSHA fields . . 88
3.4.4 The distribution of mesoscale turbulent energy 89
3.5 Observation of eddies and mesoscale turbulence in the SST field 91
3.5.1 SST signatures of eddies in infrared imagery 91
3.5.2 Microwave radiometry for viewing ocean eddies 96
3.6 Views of mesoscale turbulence from ocean color 98
3.7 Surface roughness signatures of eddies. 103
3.7.1 Hydrodynamic modulation patterns of eddies 103
3.7.2 Slick-modulated signatures of eddies 106
3.7.3 Sun glitter photography 109
3.7.4 Can imaging radar become a reliable tool for observing turbulent eddies?. 110
3.8 References 111

4 Mesoscale ocean features: Fronts. 115
4.1 Boundaries in the ocean 115
4.2 The remote-sensing signatures of ocean fronts 118
4.2.1 Sea surface temperature signatures of fronts 118
4.2.2 Can fronts be detected by altimetry? 124
4.2.3 Observing fronts in ocean color images 126

4.2.4 Frontal signatures in radar surface roughness images . . 129
4.2.5 Direct measurement of currents using Doppler analysis of SAR data 134
4.3 Tracking fronts 136
4.3.1 Mapping frontal edges 136
4.3.2 Automatic parameterization of frontal structure 140
4.4 Climatology of the major ocean fronts 142
4.5 Mesoscale frontal variability 146
4.5.1 The Gulf Stream 146
4.5.2 The Southland Front 148
4.5.3 Antarctic Circumpolar Fronts 148
4.6 Biological production associated with ocean fronts 152
4.6.1 Antarctic Circumpolar Current 152
4.6.2 Fronts in the southwest Atlantic 153
4.7 References 155

5 Ocean mesoscale features: Upwelling and other phenomena 159
5.1 Upwelling 159
5.1.1 The causes and consequences of upwelling 159
5.1.2 Aspects of upwelling detected by satellites 162
5.1.3 Upwelling regions of the world seen from space 167
5.1.4 Using satellite data in upwelling research 171
5.2 Wind-driven, offshore, dynamical features 174
5.3 Large river plumes 177
5.4 Island wakes 179
5.5 Ice edge phytoplankton blooms 181
5.6 Remote sensing in iron limitation studies 184
5.7 Making the most of satellite data for mesoscale studies: conclusions from Chapters 3–5 187
5.8 References 190

6 Planetary waves and large-scale ocean dynamics 195
6.1 Phenomena seen best from satellites 195
6.2 Detecting planetary waves from space 196
6.2.1 Producing composite anomaly datasets 197
6.2.2 Producing Hovmöller diagrams to reveal propagating features 200
6.2.3 Altimetry reveals the first compelling evidence of planetary waves 202
6.2.4 Sea surface temperature signatures 204
6.2.5 Evidence of planetary waves in ocean color 206
6.3 The characteristics of Rossby waves 206
6.3.1 A summary of planetary wave theory 206
6.3.2 How can Rossby waves be seen at the sea surface? . . . 210

6.4 Estimating planetary wave speed 212
6.4.1 Methods for analyzing Hovmöller diagrams 212
6.4.2 Radon transform 213
6.4.3 Mapping the speed of planetary waves 216
6.4.4 Meridional components of planetary wave propagation . 216
6.5 Understanding Rossby waves better 220
6.5.1 Satellite data confirm the existence of Rossby waves . . . 220
6.5.2 Revisiting Rossby wave theory 221
6.5.3 The importance of Rossby waves 223
6.6 Other large-scale propagating phenomena 224
6.6.1 Equatorial Kelvin waves 225
6.6.2 Tropical instability waves 227
6.6.3 The Madden–Julian Oscillation 231
6.6.4 The Antarctic circumpolar wave 233
6.7 References 234

7 Ocean biology from space 239
7.1 Introduction 239
7.2 Phytoplankton blooms 240
7.2.1 An unfolding new view of phytoplankton distribution . . 240
7.2.2 The global distribution of chlorophyll 246
7.2.3 Scientific exploitation of satellite ocean color data 251
7.2.4 Coccolithophores 254
7.3 Primary production 255
7.3.1 Theoretical background 255
7.3.2 Methods for estimating production from remote sensing 258
7.3.3 Estimating PAR from space 262
7.3.4 Measurements of primary production 264
7.4 Fisheries 267
7.4.1 General considerations 267
7.4.2 Fisheries management and research 268
7.4.3 Operational applications to specific fisheries 270
7.4.4 Aquaculture 271
7.5 Habitats in shallow tropical seas 272
7.6 Coral reefs—a wider role for satellite data 279
7.7 Marine biology in the future 282
7.8 References 283

8 Ocean surface waves 293
8.1 Introduction 293
8.2 Measuring ocean waves—principles 294
8.2.1 Characterizing ocean waves in terms of measurable parameters 294
8.2.2 Wave energy and spectra 295
8.2.3 Significant wave height 296

8.2.4 Measuring ocean waves from an altimeter 298
8.2.5 Observing waves with the synthetic aperture radar (SAR) 300
8.2.6 Wave spectrometry 304
8.3 Measuring ocean waves—practical systems 307
8.3.1 Altimeters for measuring SWH 307
8.3.2 SWH data products from altimeters 309
8.3.3 Synthetic aperture radars 312
8.3.4 ASAR wave-related products 313
8.4 Applications of wave data from satellites 317
8.4.1 Applications of SWH 317
8.4.2 Applications of SAR 318
8.5 Using satellite data in wave prediction models 319
8.5.1 Wave prediction models 319
8.5.2 Use of satellite data with wave models 320
8.5.3 Assimilating satellite data into models 321
8.6 Wave climate 322
8.7 Assessment and future perspectives 326
8.8 References 328

9 Wind over the sea 333
9.1 Measuring wind over the sea from satellites 333
9.1.1 Scatterometry 334
9.1.2 Wind data from SAR 336
9.1.3 Wind data from altimeters 336
9.1.4 Microwave radiometry 337
9.1.5 The alternatives to satellite measurements 340
9.2 Oceanography and wind data 341
9.2.1 Differences between analysis winds and satellite winds . . 342
9.2.2 Which type of wind data should be used to study ocean phenomena? 342
9.3 Tropical cyclones over the ocean 344
9.3.1 Detecting and predicting tropical cyclones 344
9.3.2 Use of ocean remote sensing to study hurricane–ocean interaction 347
9.4 Satellite winds for offshore wind farms 350
9.5 References 354

10 Fluxes through the air–sea interface 359
10.1 Introduction 359
10.2 Determining fluxes 361
10.2.1 General principles 361
10.2.2 Theoretical basis of flux parameterizations 362
10.3 Satellite data available for surface fluxes 363
10.3.1 Sea surface temperature 364
10.3.2 Wind 365

10.3.3 Sea surface roughness 366
10.3.4 Significant wave height and wave age 367
10.3.5 Water vapor 368
10.3.6 Air temperature at sea level 369
10.3.7 Gas concentrations in the surface sea and the ABL . . . 369
10.4 Measuring fluxes from space 370
10.4.1 Radiative flux 370
10.4.2 Gas flux 372
10.4.3 Turbulent heat flux 378
10.5 Satellite flux measurements in future? 382
10.6 References 386

11 Large ocean phenomena with human impact 391
11.1 Introduction 391
11.2 El Niño 393
11.2.1 The ENSO phenomenon. 393
11.2.2 Observing an El Niño from satellites 402
11.2.3 Observing an El Niño in sea surface temperature from satellites 404
11.2.4 Applying altimetry to the study of El Niño 407
11.2.5 Satellite-observed wind fields and ocean surface currents. 410
11.2.6 Chlorophyll 415
11.2.7 Rainfall over the ocean 418
11.2.8 Synergy 419
11.3 Monsoons 421
11.3.1 Introduction 421
11.3.2 Illustrating the Indian monsoon using satellite data 422
11.3.3 Interannual variability of the Indian monsoon 424
11.4 Sea ice distribution 426
11.4.1 Introduction 426
11.4.2 Measuring sea ice from space 427
11.4.3 How is the distribution of sea ice changing? 431
11.5 Tides, sea level, surges, and tsunamis 435
11.5.1 A surveyor's benchmark in the sky 435
11.5.2 Mean sea level 437
11.5.3 Storm surges 441
11.5.4 Tsunamis 442
11.6 Conclusion 444
11.7 References 447

12 Internal waves 453
12.1 Introduction 453
12.1.1 Ocean internal and interfacial waves 453
12.1.2 The importance of internal waves in physical and biological oceanography 456

12.2 Internal wave signatures detected with SAR 457
12.2.1 Introduction. 457
12.2.2 Internal, solitary wave packets observed by SAR 459
12.2.3 Identification of internal wave trains and their propagation direction 462
12.2.4 Hydrodynamic and film modulation 463
12.2.5 Internal wave mean propagation speed 468
12.2.6 Inversion of polarity in SAR signatures of internal waves 468
12.3 Internal waves and ocean color 471
12.3.1 Observations 471
12.3.2 Remote sensing and depth distribution of ocean chlorophyll 474
12.3.3 A model for interpreting ocean color signatures of internal tides 475
12.3.4 Internal waves and primary production 477
12.4 Impact of remote sensing on our knowledge of internal waves . . 479
12.5 References 480

13 Shelf seas, estuaries, and coasts 485
13.1 Introduction 485
13.2 Observing shelf seas from space 486
13.2.1 What is distinct about the remote sensing of shelf seas? 486
13.2.2 Variability scales in shelf seas 488
13.2.3 Shelf edge phenomena 492
13.2.4 Thermal signatures of shelf sea dynamical phenomena . . 497
13.2.5 Remote sensing of suspended sediments in shelf seas . . 508
13.2.6 Monitoring ecosystems and water quality 516
13.3 Coastal altimetry 523
13.3.1 Challenges and opportunities for altimetry in coastal seas 523
13.3.2 Potential applications of coastal and shelf altimetry . . . 524
13.3.3 Practical approaches to improving altimeter accuracy in shelf seas 527
13.4 Coastal and estuarine remote sensing 528
13.4.1 Important edges of the ocean 528
13.4.2 A mismatch of scales? 529
13.4.3 Coastal remote-sensing applications using satellite data. . 532
13.5 References 534

14 Putting ocean remote sensing to work 539
14.1 Satellites and applied oceanography 539
14.1.1 Introduction 539
14.1.2 The fundamental importance of ocean monitoring and forecasting 540
14.1.3 Motivation for scientists to engage in applied ocean remote sensing 541

14.2 Integrated ocean-forecasting systems 543
14.2.1 What is operational oceanography? 543
14.2.2 Combining satellite oceanography and ocean models for operational tasks. 547
14.2.3 Assimilating satellite data into ocean-dynamical models . 551
14.3 Ecosystem modeling 555
14.3.1 How can satellite ocean color data support operational applications? 555
14.3.2 Marine ecosystem models, scientific principles, and operational purpose 555
14.3.3 Ways in which ocean color data are used in ocean modeling 558
14.3.4 Sequential assimilation to constrain ecosystem state variables 562
14.3.5 Characterizing light penetration in numerical models . . 564
14.3.6 Alternative approaches to ocean color assimilation 565
14.4 Preparing satellite data for operational use 569
14.4.1 Providing merged data from multiple sensors/satellites . 570
14.4.2 GHRSST: A case study on preparing SST data for operational applications 575
14.5 Oil spill monitoring 583
14.5.1 Introduction 583
14.5.2 How can oil spills be monitored routinely from space? . 584
14.5.3 CleanSeaNet, a European service for oil spill detection . 586
14.6 Using satellite data for climate monitoring 588
14.6.1 Introduction 588
14.6.2 The ocean's role in the climate system 589
14.6.3 Essential climate variables 591
14.6.4 Ocean datasets used for climate 597
14.7 References 602

15 Looking forward 607
15.1 Achievements 607
15.1.1 Oceanographic discoveries from satellite data 607
15.1.2 Does ocean science need remote sensing? 609
15.2 Securing the future for ocean remote sensing 610
15.2.1 Essential satellite oceanography 610
15.2.2 Limitations of existing sensors and platforms 611
15.2.3 Future sensors, platforms, and systems for observing the ocean 613
15.3 Challenges for satellite oceanographers 617
15.4 References 619

Index 621

To Diane, Steve, and Phil

and in memory of my brother
John Malcolm

Preface

Although this is a companion volume to my previous book *Measuring the Oceans from Space*, published in 2004, it approaches the subject from a different angle by providing a broad-ranging introduction and review of the applications of satellite remote sensing across the field of oceanography. During the five years it has taken to prepare this book what has motivated me is the attraction of the subject itself. In searching the scientific literature I have been excited by the interesting, new, and important elements of satellite oceanography being discovered around the world. I hope this book will help others to discover for themselves how sensors in space are giving us a new perspective of oceanographic phenomena and are increasingly enabling ocean science to serve the needs of modern civilization.

Primarily I have written with postgraduate and senior undergraduate students in mind. Readers of my previous book will know that it aimed to present a systematic explanation of the diverse methods of satellite oceanography, the breadth of scientific knowledge, and the richness of innovative technology developed over 30 years since the first dedicated ocean satellite was flown in 1978. But satellite oceanography is too important a subject to treat just as a specialist topic for those with an interest in observational instrumentation and data-processing techniques. In recent years the availability of high-quality, processed satellite data has delivered remarkable and inspiring images of the ocean, which can almost tell their own stories.

Therefore this book, *Discovering the Ocean from Space*, is written to open the eyes of all oceanographers, students, and seasoned researchers alike, to show you how much we can learn about the ocean from space. The chapters are arranged in relation to the different ocean phenomena that are seen from satellites, rather than organized by sensor or techniques. The emphasis is on revealing the ocean rather than presenting the methodology, although there are many cross-references to *Measuring the Oceans from Space* where the underpinning principles and remote-sensing methods are explained. The one exception to this is Chapter 2 where a condensed summary of the methods and elements of satellite oceanography has been inserted so that

the volume is self-contained when used as a textbook for a course on satellite oceanography to a class of students from all branches of ocean science.

With teaching in mind, I hope that the extensive use of illustrations and color images will attract students to browse and then become interested in the ocean phenomena themselves because of the insights that come from looking at pictures. Once your interest is aroused, I hope you will use the many citations of oceanographic research papers to follow up the topics in more detail than there is room to include in a single volume.

However, another theme that runs through the book is that of applied oceanography, culminating in Chapter 14 on "Putting ocean remote sensing to work". The widespread availability of satellite data has opened up many new opportunities for making oceanographic science relevant to the challenges of modern civilization, including safety and security of those who work at sea, the better management of the health of marine ecosystems, and the monitoring of ocean climatology. I believe that this will lead to increased career opportunities in which marine science is applied to industrial, commercial, and environmental management contexts as well as research. I hope this book will help to prepare oceanography students for such careers by demonstrating the important role of satellite data in ocean monitoring and forecasting.

Another reason for writing the book has been to demonstrate how important satellite observations have become to the subject of ocean science as a whole. It is easy to take for granted the data from satellites that are used routinely in research and in the application and teaching of oceanography. In the early days of ocean remote sensing, satellite ocean data were often fortuitous by-products from a wider pattern of investment in space technology, or from space-based meteorological observing systems. However, in the 21st century the continued use of satellites to observe the ocean needs to be justified for its own sake, in terms of the benefits for operational applications, for innovative oceanographic research, and for essential climate monitoring. I hope that by assembling in one volume a variety of diverse applications of satellite data this book will be useful to those who argue the case for continued funding to maintain long-term satellite monitoring of the ocean.

I owe some of my readers an apology and an explanation. When *Measuring the Oceans from Space* was published in 2004, it promised a second volume in 2005 called "Understanding the Ocean from Space" which would describe the applications of satellite oceanography. Originally intended as Part 3 of the first volume, the growing amount of material forced us to transfer it into a second volume. However, the demands of my work as Professor of Oceanography from Space at Southampton University and head of the Laboratory for Satellite Oceanography at the National Oceanography Centre grew considerably whilst at the same time more and more new applications of ocean remote sensing appeared in the literature. The confident promise to my publishers of a second volume within a year soon became an embarrassment as I found little spare time to devote to its completion. Emails from friends and colleagues around the world let me know I could not quietly forget this promise. Finally, five years late, with the support of two young colleagues, José da Silva and Susanne Fangohr who helped me as co-authors of a chapter each, the unfailing encouragement

of my local publisher, Praxis, and upheld by the heroic patience and support of my wife, the book is finally completed. Of course a lot more has been published about the ocean applications of remote sensing in those five years and, in attempting to keep up to date, the book has grown to over 600 pages. But I am very glad that the delay has allowed me to include mention of several exciting developments from scientists around the world, which have strengthened the content of the book, and led me to rename it with its new title of *Discovering the Ocean from Space*. I hope all my readers will taste something of the excitement of discovery that I have experienced in reading the scientific papers that underpin the contents of these pages.

I ought to explain to my European readers that, in switching to the use of North American spellings of our common English language, I am following the policy of the international publisher, Springer. I am very happy to do so if that will make the work more accessible to readers in other parts of the world, where our British spelling of kilometres and colour perhaps seems rather quaint. Nonetheless, I must warn my students in England that I shall not be changing the habit of a lifetime and will continue to correct your essays if you start to talk about color and meters!

Finally there are two things I want to say to those readers who are students learning about this subject for the first time. One is to encourage you to get into the habit of testing the ideas presented in this book by following up the cited references. I have done my best to summarize accurately the results of others and relate them to a wider context, but in reducing and condensing the work of others some ideas can get overlooked or misrepresented. I have aimed to provide enough key references to enable you to start a literature search of your own when you find a topic of particular interest. The other thing is to wish you enjoyment as you read this book and excitement as you learn about the unique applications of satellite oceanography. But your journey of discovery need not be limited to this book. As explained in Chapter 2, high-quality satellite ocean data products in digital image form are freely accessible and you can easily explore them yourself on your personal computer. Who knows what stunning images of ocean phenomena are already acquired but unnoticed within space agency databases, waiting to be discovered by someone reading this book?

Acknowledgments

In writing a book like this I have drawn knowledge, information, images, and input from many sources. I have been supported in a variety of ways by many people. Yet when a book has taken six years to complete, it is not easy to remember them all, which makes writing this acknowledgments section a difficult task. Let me apologize at the start if I inadvertently overlook someone whom I should be explicitly thanking.

The scientific content of this book is largely a distillation of the work of others, selected and presented through my own perspective on the subject. I have attempted to acknowledge by citations the key authors of the ideas in the book. Even though the references listed in each chapter are not by any means exhaustive they are intended to provide a good starting point for a deeper study of the topic. Where a figure or table is copied from, or based on, the work of another author I have tried always to make that clear. In the few cases where a straight copy of a figure has been used, the material was either free for use or permission has been gained. In most cases I have drawn the figures myself, sometimes using parts of other figures but as far as possible starting from scratch. In cases where figures are based on digital datasets downloaded from agency databases the Internet source is stated. For image processing and enhancement I have relied largely on the Bilko image-processing software used for training courses in marine remote sensing. The figures have been refined to publication quality using Adobe Illustrator, and I am grateful to Neil Shuttlewood, the typesetter, whose eagle eye does not let me get away with a lowering of his high standards for graphical quality.

I owe a debt of gratitude to several people who have helped me with the text. My friends and colleagues, Susanne Fangohr and José da Silva, agreed early on to co-author a chapter each and then had to wait patiently while I finished the rest of the book. Their willingness to share the load prevented me from being overwhelmed by the scale of the task. Another good friend, Craig Donlon, in the midst of his busy job at ESA, took time to provide detailed criticism of Chapters 14 and 15 where much of the subject matter is a distillation of emerging ideas about operational oceanography

and there are fewer published references and less of an established scientific consensus. To Philippe Blondel, the series editor at Praxis, fell the task of reviewing the complete text and I have very much appreciated his constructive support and critical judgment throughout. I am sure that the inputs of these colleagues have improved the book, although I must take full responsibility for any errors that may have survived their critical review. Furthermore, I have to acknowledge the combined efforts of Philippe Blondel and Clive Horwood, publisher of Praxis, who recognized when the demands of my main job left almost no spare time for working on the book, but were able to supply just the right balance of patience, encouragement, and pressure to keep the momentum going.

Unlike our colleagues in the humanities, it is not considered a priority task for science professors in British universities to write books which distill the essence of a subject into a book. The quality of original scientific papers, competitive success in winning research funds, and the effective teaching of students are the primary targets against which our performance is evaluated. The writing of books is more of an optional extra and must be squeezed into our workload when there is time. I am therefore particularly grateful that the School of Ocean and Earth Science at Southampton University granted me a sabbatical semester in 2008 during which I was enabled to spend several months in the peace and beauty of Sweden working to finish the book. In order for me to gain this freedom, several of my colleagues in the Laboratory for Satellite Oceanography at Southampton were willing to cover my ocean remote-sensing teaching duties, including Peter Challenor, Paolo Cipollini, David Cromwell, Susanne Fangohr, Richenda Houseago-Stokes, Graham Quartly, Colette Robertson, Helen Snaith, and Meric Srokosz. My regular teaching colleagues, Neil Wells, Harry Bryden, and Bob Marsh, have also helped out to allow me more time for writing.

I am particularly grateful to Werenfrid Wimmer and David Poulter, research staff who have supported me in a variety of ways. The same is true for those research students whom I have supervised since 2004, including Nico Caltabiano, Stephanie Henson, Chris Jeffery, Mounir Lekouara, Violeta Sanjuan-Calzado, Anna Sutcliffe, and Gianluca Vulpe. As their knowledge of specific topics of ocean remote sensing has grown they have both broadened my grasp of the wider field and illuminated for me the details of particular corners.

I am grateful also to friends and colleagues from around the world, whom I meet at conferences or who send me papers, for helping me keep up with developments in ocean remote sensing. Neither should I forget those of my former students who are now themselves teaching in Brazil, Mexico, Portugal, the U.S.A., and elsewhere around the world, who have kept asking when the new book would be ready. Here is your answer.

Finally I must acknowledge once more that I could not have finished this book without the patient support, love, and care of my wife Diane.

Figures

2.1 Information flow in ocean remote sensing 8
2.2 Sketch showing how the IFOV defines the measurement footprint during the sample integration time 10
2.3 Swath-filling geometry of a rectangular, line-scanning sensor 11
2.4 The two types of orbit used for Earth-observing satellites 12
2.5 Ground track of a typical near-polar, low-Earth orbit 13
2.6 A single day's coverage over Europe 14
2.7 Diagram representing the space-time sampling characteristics of four types of sensor 16
2.8 Outline of data-processing tasks to convert raw satellite data into ocean products 17
2.9 Examples of map projection types 25
2.10 The electromagnetic spectrum, showing the regions exploited by typical remote-sensing instruments 28
2.11 Different remote-sensing methods and classes of sensors used in satellite oceanography 29
2.12 A typical spectrum of Earth-leaving radiance in the visible and near-IR part of the spectrum 31
2.13 Factors which affect light reaching an ocean color sensor 31
2.14 Infrared emission spectra of black bodies 37
2.15 Schematic to illustrate the principle of using band-differential response to the atmosphere as the basis of atmospheric correction algorithms 38
2.16 The conical scanning arrangement for the ATSR 39
2.17 Characteristic temperature profiles at the sea surface 40
2.18 Physical dependences that determine microwave radiation measured above the atmosphere when viewing the open sea 44
2.19 The relationship between different distance quantities used in altimetry 47
2.20 Sketch of typical measurements of σ_0 as a function of incidence angle and sea state 52
3.1 Early infrared images of the ocean 70

3.2 Sea surface temperature measured by the AVHRR sensor. 71
3.3 Extracts from images of data products derived from MODIS on the Terra satellite, January 27, 2005 73
3.4 Spinup to geostrophic equilibrium from the onset of a pressure gradient force 73
3.5 Part of a geostrophic flow field with steady flow along isobars 74
3.6 A simplified two-layer ocean and the difference between barotropic and baroclinic gravity wave speeds 75
3.7 The basic structure of a simple eddy in the northern hemisphere 77
3.8 Along-track altimeter records from TOPEX/Poseidon 81
3.9 Example of the two-dimensionally smoothed SSHA field from a single orbit cycle of Jason-1. 82
3.10 Sea surface height anomaly over the Arabian Sea on August 4, 1993, and map of the formal error estimate 84
3.11 Sea surface height anomaly over the Mediterranean Sea on May 10, 2006 . . 85
3.12 Sea surface altimetry data products for the Southern Ocean off South Africa on August 25, 1993 87
3.13 Root mean square of sea level anomaly obtained from 11 years of sea surface height. 90
3.14 Eddy kinetic energy in the southwest Atlantic Ocean on November 5, 2005 . 91
3.15 Mean eddy kinetic energy in the Mediterranean 92
3.16 Sea surface temperature field derived by MODIS on Aqua, April 18, 2005, showing the meanders of the Gulf Stream 93
3.17 Map of level 2 SST data from ATSR nighttime image on May 9, 1992 over the Balearic Islands. 94
3.18 SST maps of the Southern Ocean poleward of South Africa, as measured by the AMSR-E microwave radiometer. 97
3.19 Map of chlorophyll concentration derived from MODIS on Aqua, April 18, 2005. 98
3.20 A break in the clouds over the Barents Sea on August 1, 2007 reveals a large, coccolithophore bloom 99
3.21 Map of chlorophyll concentration derived from the SeaWiFS overpass of the Gulf of Aden 100
3.22 SeaWiFS chlorophyll-*a* composite image, and sea level anomaly in 1999 . . . 102
3.23 Comparison between a 1 km resolution AVHRR IR image and a 100 m resolution ERS-1 SAR image on October 3, 1992. 104
3.24 ERS-1 SAR image over the Kuroshio Current in the northwest Pacific Ocean, acquired on December 23, 1994 105
3.25 Ocean mesoscale eddies in Pacific Ocean east of Japan 107
3.26 ERS-1 SAR image of the Tyrrhenian Sea north of Sicily, acquired on September 19, 1993 108
3.27 Example of spiral eddies in the Mediterranean Sea off Egypt on October 7, 1984 110
4.1 Section through an ocean front (northern hemisphere) 116
4.2 Secondary dynamical processes that may occur at ocean fronts 117
4.3 SST image of the front where the warm Agulhas Current detaches from the east African coast 119
4.4 Part of a brightness temperature image from the ATSR 120
4.5 Application of high-pass digital filters to the image in Figure 4.4, to enhance visualization of ocean fronts. 121
4.6 Application of a spatial variance filter to the image shown in Figure 4.4 . . . 123

4.7 Relationship between the filamentary along-front velocity structure in a multiple-core front, ADT, MDT retrieved from altimetry, a best-fit geoid, and the SSHA from altimetry . . . 125
4.8 Chlorophyll concentration revealing the region where the warm Agulhas Current, with low chlorophyll, detaches from the east African coast . . . 127
4.9 The Falklands (Malvinas) current visualized by its ocean color signature . . . 129
4.10 ERS SAR image of a region east of Taiwan showing convergent fronts in the ocean but also a potentially misleading atmospheric front . . . 130
4.11 ERS SAR image over the Lombok Straits showing a number of local surface convergent fronts . . . 132
4.12 ERS-1 SAR image acquired on January 7, 1995 over the East China Sea . . . 133
4.13 Component of apparent surface current, U_D, detected by Doppler centroid analysis of SAR data . . . 134
4.14 Normalized radar cross-section and Doppler velocity analyzed from a wide-swath image obtained by Envisat on February 6, 2003 . . . 136
4.15 Normalized radar cross-section and Doppler velocity analyzed from wide-swath images obtained by Envisat ASAR on four successive overpasses 3 days apart between September 13 and 22, 2007 . . . 137
4.16 Detection of fronts on AVHRR SST image by the multi-image algorithm . . . 139
4.17 The extraction window used for the front-following algorithm . . . 140
4.18 The hyperbolic functional form used to represent the temperature profile across an isolated front . . . 141
4.19 Long-term, seasonal-averaged frequency of thermal fronts occurring within each 9.28 km resolution pixel of the Pathfinder SST dataset across the Pacific Ocean . . . 144
4.20 Long-term, annual composite frontal probability map in the East China Sea for the 1985–1996 period derived from the AVHRR Pathfinder SST dataset . . . 145
4.21 Probability maps for the East China Sea showing seasonal breakdown of data presented in Figure 4.20 . . . 147
4.22 Mean gradient of observed sea surface temperature from ATSR . . . 149
4.23 Time–latitude plot of meridional gradient of sea surface height between 1994 and 1997 . . . 150
4.24 A typical SSH gradient field south of Australia and New Zealand overlaid with mean best-fit SSH contours optimized for the whole period of observations . . . 151
4.25 Mean summer chlorophyll concentrations south of Australia and New Zealand 153
4.26 Six ranges of chlorophyll-*a* magnitudes, and information from the three mean fields . . . 154
5.1 Coastal upwelling . . . 160
5.2 Equatorial upwelling . . . 161
5.3 SST in the Benguela upwelling region of the southwest Atlantic Ocean . . . 163
5.4 Chlorophyll-*a* concentration in the Benguela upwelling region of the southwest Atlantic Ocean . . . 164
5.5 Wind vectors from QuikScat over the Benguela upwelling region from the evening overpass of QuikScat on March 2, 2005 . . . 165
5.6 Cross-section through an eastern ocean margin showing the typical thermal structure, the equatorward boundary current, and associated SSH when there is no upwelling and when there are upwelling-favorable winds . . . 166
5.7 The major upwelling zones around the world . . . 167

5.8 Evidence of equatorial upwelling in the 6-year cumulative average map of chlorophyll concentration as measured by the MODIS sensor on Aqua 168
5.9 Benguela upwelling monthly average for February 2004 of SST and chlorophyll 169
5.10 The Canary upwelling along the coast of northwest Africa 170
5.11 Upwelling along the coasts of Peru and Chile 172
5.12 Upwelling along the Oregon and California coasts 173
5.13 Weekly average wind speed and direction from QuikScat over Central America for week ending February 18, 2006 176
5.14 Upwelling during a "Norte" event in the Gulf of Tehuantepec off the Pacific coast of Mexico 177
5.15 Chlorophyll map of west Equatorial Atlantic Ocean derived from Aqua MODIS on September 30, 2006 179
5.16 Schematic cross-section through the sea showing wind shadow, shear lines, and upwelling driven by Ekman transport downwind of an isolated oceanic island 180
5.17 Three maps of chlorophyll-*a* concentration over the Galapagos region, derived from SeaWiFS ocean color data 182
5.18 Aqua MODIS view of the Ross Ice Shelf on February 24, 2008 184
5.19 SeaWiFS-derived chlorophyll image of the bloom resulting from the SOIREE iron enrichment experiment 185
5.20 Contours of the ADT field overplotted onto the satellite-derived chlorophyll-*a* field surrounding the Crozet Plateau, for the week of October 23–30, 2004. . 187
5.21 Map showing how the timing of bloom initiation varies with position over the Crozet Plateau 188
5.22 Contours of the ADT field overplotted onto the satellite-derived chlorophyll-*a* field surrounding the Crozet Plateau 189
6.1 How values of an ocean variable, sampled at full resolution on a level 2 grid in sensor co-ordinates, are allocated to the corresponding level 3 geographical grid 197
6.2 Creating a climatology and anomalies based on weekly composites 199
6.3 NCEP monthly SST climatology. 201
6.4 The "data cube" produced by vertically stacking successive two-dimensional maps of a satellite data time series 202
6.5 Map of SSHA in the Indian Ocean, and evidence of planetary waves 203
6.6 Hovmöller plot at 25°S of the SST anomaly derived from the ATSR. 205
6.7 Hovmöller plots of chlorophyll and SSHA 206
6.8 Torque experienced by water columns moving south or north at a tropical north latitude. 207
6.9 The first baroclinic mode of planetary waves in the northern hemisphere . . . 208
6.10 How planetary wave speed depends on the slope of wave signatures in time–longitude (Hovmöller) plots 212
6.11 The principle of the Radon transform 214
6.12 Example of the full two-dimensional transform for a Hovmöller field such as Figure 6.11 214
6.13 Example of the Radon transform, and the corresponding variance–direction plot 215
6.14 Global maps of planetary wave speed measured using Radon transforms of Hovmöller plots 217
6.15 Zonal mean speed of planetary waves detected by their signature in different data types. 218

6.16 Polar plot of energy from the three-dimensional Radon transform of TOPEX/Poseidon SSH anomalies . . . 219
6.17 Ratio between planetary wave speed measured from multimission altimetry data and first-mode, theoretical Rossby wave speed . . . 221
6.18 Currents, vertical displacements, and Coriolis forces in a baroclinic, equatorial Kelvin wave . . . 226
6.19 SSHA from TOPEX/Poseidon plotted in longitude, time at the Equator . . . 227
6.20 60-day sequence at 5-day intervals of AMSR-E 3-day composite SST images 228
6.21 Time–longitude plots of temperature in the equatorial Atlantic . . . 230
6.22 MJO index over the period of study; longitude–time plot; the Niño3 index. . 232
6.23 Approximate positions of the ACC and sea ice limits around Antarctica, and the relative phase distribution of SST and surface pressure anomalies in the ACW phenomenon . . . 234
7.1 Image created from an extract of a level 2 chlorophyll image product from SeaWiFS, showing the spring bloom off Nova Scotia . . . 241
7.2 Global level 3 daily image showing chlorophyll concentration retrieved from SeaWiFS, for April 21, 2001. . . . 242
7.3 Part of the global, daily, level 3 chlorophyll image from SeaWiFS shown in Figure 7.2, at its full resolution. . . . 243
7.4 Global, level 3, 8-day composite image showing chlorophyll concentration retrieved from SeaWiFS. . . . 244
7.5 Global, level 3, monthly composite image showing chlorophyll concentration retrieved from SeaWiFS. . . . 244
7.6 Northeast Atlantic extracts from alternate, 8-day, level 3 chlorophyll composites from SeaWiFS showing the northward progression of the spring bloom 245
7.7 Global composite image of all SeaWiFS chlorophyll data acquired from the mission launch in September 1997 until the end of 2007 . . . 246
7.8 Full-scale extract from Figure 7.7 to the northeast of Australia, showing the production associated with islands, atolls, and reefs . . . 248
7.9 Atlantic Ocean extracts from four monthly climatologies of chlorophyll derived from SeaWiFS data between 1997 and 2007 . . . 248
7.10 North Atlantic extract showing the monthly chlorophyll anomaly for April 2001. . . . 250
7.11 Coccolithophores . . . 254
7.12 Typical variation of production rate, P, with irradiance, E . . . 257
7.13 Idealized structure of the vertical distribution of biomass in the upper ocean, used in models of primary production. . . . 260
7.14 Examples of PAR distributions derived from SeaWiFS data . . . 263
7.15 Annual primary production within the World Ocean . . . 265
7.16 Net primary production in April 2007. . . . 266
7.17 Timing of the maximum phytoplankton biomass in the northwest Atlantic from February to July, and relationship between larval haddock survival index and local anomalies in bloom timing. . . . 270
7.18 Schematic showing how the apparent spectral reflectance of the sea bed appears to vary when viewed through different depths of seawater. . . . 276
7.19 Broad-scale habitat map of the Caicos Bank derived by supervised classification from a SPOT XS dataset, and false-color composite image of SPOT multispectral image data . . . 278
7.20 Example of a NOAA Coral Reef Watch map of bleaching hotspots . . . 281

7.21	Map of DHW index produced by NOAA Coral Reef Watch	281
8.1	Interaction of an altimeter pulse with a rough sea surface	298
8.2	Range and azimuth directions for a SAR viewing ocean waves	302
8.3	ERS-1 SAR image showing long surface waves off Prawle Point in the English Channel	302
8.4	Radar spectrometer geometry showing how range resolution achieved by time-sampling the echo of a nearly normal radar beam achieves a coarser resolution in the ground-track direction	305
8.5	How a conical scanning wave spectrometer radar achieves views in all directions	306
8.6	Spatial distribution of the ground track of an altimeter on a satellite	309
8.7	Significant wave height data products produced by NASA/JPL from the Poseidon altimeter on Jason-1	311
8.8	An Envisat ASAR, wave mode, level 1 product showing the amplitude image of a single-look, complex radar backscatter cross-section for a wave-mode imagette	314
8.9	Wave heights and mean wave propagation direction retrieved from an Envisat ASAR, single-look, complex image over the Gulf of St. Malo in the English Channel	315
8.10	(a) Envisat ASAR image of the Pacific Ocean south of Santa Barbara, showing the northern group of Channel Islands, on January 20, 2006, overlaid with color denoting significant wave height and wave. (b) Closeup of a part of (a), revealing the signatures of the swell on the SAR image.	316
8.11	Wave heights and mean wave propagation direction, retrieved from an Envisat ASAR, single-look, complex image over the Gulf of St. Malo in the English Channel	319
8.12	Seasonal climatology of mean significant wave height over the northeast Atlantic	324
8.13	Sensitivity of wintertime significant wave height	326
9.1	Daily coverage of ocean surface winds measured by ascending (morning) passes of QuikScat on August 9, 2008	335
9.2	Daily coverage of ocean surface winds measured by descending (evening) passes of QuikScat on August 9, 2008	335
9.3	Wind speed data retrieved from the Jason-1 altimeter on August 6–7, 2008	337
9.4	Wind speeds retrieved from ascending overpasses of the AMSR-E microwave radiometer on August 9, 2008	338
9.5	Wind vector output from Windsat produced in near-real time by NOAA	339
9.6	NCEP reanalysis of global winds on August 9, 2008	341
9.7	Hurricane Ivan over the Gulf of Mexico, as revealed by the NOAA AVHRR visible waveband radiometer on September 14, 2004.	345
9.8	Surface wind field from the ERS-l scatterometer in Tropical Cyclone Elsie	346
9.9	The location and intensity of Hurricane Katrina at intervals of 6 hours	349
9.10	Distributions of wind power density derived from QuikSCAT	353
9.11	Wind speed map derived from an ERS-2 SAR scene acquired on February 25, 2003	354
10.1	Two-layer model of air–sea interaction showing a layer dominated by turbulent mixing and one governed by molecular diffusion on either side of the air–sea interface	361
10.2	Three levels of mean square sea surface slope and air–sea flux at given wind speeds	366

10.3 Variation with water temperature of the solubility of CO_2 and the product $s \cdot (Sc)^{-0.5}$ . . . 373
10.4 Parameterizations of the gas transfer velocity for different wind speeds 374
10.5 Global values of the correction factor R . . . 375
10.6 Mean annual net sea-to-air flux for CO_2 . . . 377
10.7 Global distribution of monthly mean values of the Dalton number 379
10.8 Variation of C_T with wind speed for different air temperatures 380
10.9 Fifteen-year mean and standard deviation of latent and sensible heat flux derived from satellite data . . . 382
10.10 Monthly mean climatology of global latent heat fluxes calculated on the basis of satellite data for the period 1987–2005 . . . 383
10.11 Monthly mean climatology of global sensible heat fluxes calculated on the basis of satellite data for the period 1987–2005 . . . 384
11.1 Equatorial section from west to east across the Pacific Ocean, schematically outlining air–sea interaction patterns . . . 396
11.2 Typical maps of monthly mean, near-surface temperature in the tropical Pacific Ocean for different phases of the El Niño cycle . . . 397
11.3 Typical longitudinal sections of monthly mean temperature distribution with depth along the Equator in the Pacific Ocean for different phases of the El Niño cycle . . . 398
11.4 Maps of the near-surface temperature anomaly corresponding to monthly mean, near-surface temperature for different phases of the El Niño cycle . . . 399
11.5 Time series of ENSO indicators, 1950 to present . . . 400
11.6 Monthly composite SST distributions over the equatorial Pacific Ocean . . . 404
11.7 Sequence of monthly SST anomaly maps of the equatorial Pacific, every 2 months during 1997–1998 . . . 405
11.8 Monthly mean sea level anomaly maps of the equatorial Pacific for every second month during 1997 and 1998 . . . 408
11.9 Plot of the SSHA measured by TOPEX/Poseidon along the Equator over the width of the Pacific Ocean . . . 409
11.10 Time series of the sea level anomaly averaged over the El Niño-3/4 region, and corresponding part of the ONI . . . 411
11.11 Hovmöller plot of monthly mean, zonal wind speed . . . 412
11.12 OSCAR surface current data products . . . 413
11.13 Hovmöller plot of monthly mean, zonal surface currents at the Equator over the Pacific Ocean . . . 414
11.14 Maps of chlorophyll monthly mean concentrations in the eastern equatorial Pacific Ocean . . . 416
11.15 Maps of the chlorophyll concentration anomaly in the eastern equatorial Pacific Ocean . . . 417
11.16 Rainfall patterns over the tropical Pacific Ocean associated with an El Niño 419
11.17 Sensitivity of satellite-derived rainfall over the sea . . . 420
11.18 Monthly mean wind vectors retrieved from QuikScat over the Arabian Sea . 423
11.19 Monthly composite SST images from Pathfinder version 5 processing of AVHRR infrared data over the North Indian Ocean . . . 424
11.20 Sea surface height anomaly maps over the Indian Ocean . . . 425
11.21 Satellite-derived maps of chlorophyll concentration at different stages of the Indian Monsoon . . . 426
11.22 Daily map of sea ice concentration around the Antarctic on August 22, 2009 429

11.23 Monthly sea ice extent in the Antarctic Ocean 430
11.24 Monthly sea ice extent in the Arctic Ocean 431
11.25 Time series of OSI-SAF sea ice concentration maps for the Arctic Ocean. . . 432
11.26 Annual time series and trendline between 1979 and 2008 of sea ice extent averaged over a month 433
11.27 September, monthly sea ice extent in the Arctic Ocean 434
11.28 Global mean sea level from the multimission SSALTO-DUACS data altimetry dataset 438
11.29 Local trends of global mean sea level from the multimission SSALTO-DUACS data altimetry dataset for the period October 1992 to January 2008. 440
11.30 Sensitivity of wintertime sea level to the North Atlantic Oscillation 440
11.31 Tsunami wave heights, and 20 Hz sea level anomaly. 444
12.1 Slick bands associated with internal waves off Cape Cod 454
12.2 Lines of constant phase produced by a small cylindrical paddle oscillating with constant frequency 456
12.3 ERS-1 SAR image dated August 21, 1994 of the region of Cape Cod showing two trains of internal solitary waves emanating from Race Point Channel . . 458
12.4 Processes associated with the passage of a linear oceanic internal wave 460
12.5 An internal solitary wave packet consisting of solitons of depression with decreasing amplitude 461
12.6 A TerraSAR-X image dated June 23, 2008 showing a typical example of double-sign signatures, and a TerraSAR-X image dated July 4, 2008 of the same region showing internal wave signatures 465
12.7 SAR image showing signature transition from double- to single-negative sign 466
12.8 Predicted backscatter contrasts across an IW packet with decreasing amplitudes 467
12.9 Wide-swath ENVISAT ASAR images showing several successive trains generated by tidal flow at the Spanish and French continental shelves 469
12.10 ERS-2 SAR image dated July 23, 1998 acquired over the Gulf of Cadiz. . . . 470
12.11 Internal waves, surface waves, and SAR image intensity variation when depression solitary waves move into shallower water 470
12.12 The Bay of Biscay, showing depth contours and the coasts of northern Spain and western France, and time series of the observed thermal structure from an XBT survey 472
12.13 Chlorophyll concentration from SeaWiFS on September 4, 1999 and coincident internal waves from the ERS-2 SAR on September 3, 1999. 473
12.14 Typical chlorophyll profiles plotted as a function of geometrical depth. 476
12.15 Schematic plot of chlorophyll profile and observation of the DCM by the satellite sensor. 477
12.16 Chlorophyll concentration along section X–Y in Figure 12.13 observed by SeaWiFS and as modeled by da Silva *et al.* (2002) 478
13.1 World map showing regions where the continental shelf extends more than about 25 km from the coast 487
13.2 Different shelf edge processes that create remote-sensing signatures 492
13.3 SST composite image from AVHRR NOAA-18 over northwest European shelf seas for the week ending April 8, 2006 showing cooler water over the shelf . 493
13.4 Enhanced color composite consisting of normalized water-leaving radiance generated from MODIS data for June 2, 2006 over northwest European coastal waters 495
13.5 SST weekly composite image from AVHRR NOAA-18 over northwest

European shelf seas for the week ending May 6, 2006 showing a plume of warmer water 496
13.6 SST monthly composite image from AVHRR NOAA-18 over U.K. western approaches for the month of June 2004. 497
13.7 Monthly averaged chlorophyll-*a* concentration for June 2004 derived from SeaWiFS 498
13.8 Typical, weekly composite SST distributions in shelf seas around the U.K. at four times of the year 499
13.9 Bathymetry of northwest European shelf seas 501
13.10 Schematic cross-section of isotherms through a shelf sea tidal-mixing front . 503
13.11 Seven-day median composite SST distributions derived from the AVHRR showing the formation of tidal mixing/stratification fronts in U.K. shelf seas during summer of 2007 504
13.12 Map of the modeled contour of the parameter $\log(h/u^3)$ 506
13.13 Sections showing chlorophyll and temperature sections through the Ushant front in July 1975 507
13.14 SeaWiFS chlorophyll image acquired on July 11, 1999, showing enhanced chlorophyll-*a* concentration along the line of the Celtic Sea tidal front and the western Irish Sea front. 508
13.15 Level 2, MODIS, normalized water-leaving radiance at 551 nm from a mainly cloud-free overpass of U.K. shelf seas on February 11, 2008 510
13.16 MODIS SST image from the same overpass as Figure 13.15 512
13.17 Part of Figure 13.15 enlarged as a gray-tone image to reveal the fine-resolution streaks aligned with the current or the bathymetry 512
13.18 Monthly composite of normalized water-leaving radiance at 551 nm from MODIS data for August 2007 513
13.19 TSM derived from a MERIS scene over the North Sea, acquired on March 27, 2007, and the corresponding signal depth 514
13.20 SAR image showing detailed shallow-water bathymetry 515
13.21 A bloom of harmful *Karenia mikimotoi* to the east of the Orkneys. 521
13.22 A near real–color composite from SeaWiFS over the Baltic Sea on July 24, 2003, showing a surface manifestation of a bloom of *Nodularia spumigena* 522
13.23 Proposed geographical domain in which a specialized coastal altimetry data product is required 526
13.24 Typical length and timescales of coastal and estuarine processes compared with the spatial- and temporal-sampling capabilities of typical classes of satellite image data 529
14.1 The GMES Marine Core Service, showing its scope and its role for assimilating satellite and *in situ* observations from several suppliers and feeding integrated ocean information to end users. 546
14.2 The narrow, direct approach of deriving a particular ocean product from a specific sensor of a particular space agency 548
14.3 A model-based approach in which POMs from satellite sensors are fed into a model system 550
14.4 Sequential assimilation scheme for physical variables in an ocean general circulation model. 553
14.5 Dominant inter-compartmental nitrogen flows in a four-compartment, NPZD, ecosystem model 557

14.6 The conventional view of how satellite ocean color data can interface with an ecosystem model 565
14.7 Sources of error and potential loss of information in the conventional assimilation of satellite-derived chlorophyll data. 566
14.8 How ecosystem information can be fed back from the model to inform ocean color data-processing choices before derived products are assimilated into the model. 567
14.9 An alternative approach to assimilating ocean color data into an ecosystem model. 568
14.10 Typical example of daily coverage of SST from six different SST data products 577
14.11 The content of a GHRSST L2P file 580
14.12 On November 13, 2002, following a heavy storm off the Atlantic coast of Spain, the oil tanker *Prestige* split in two and sank 585
14.13 Approximate revisit interval for SAR acquisitions showing the dependence on latitude of SAR coverage for oil spill monitoring 587

Tables

2.1	Levels of satellite data products from different stages of processing	18
2.2	Definition of common radar bands used for ocean remote sensing	30
2.3	Satellites carrying important ocean-viewing sensors	55
2.4	Details of major satellite ocean color sensors	57
2.5	Recent and current series of high-quality satellite infrared radiometers	58
2.6	Recent and current series of satellite microwave radiometers	59
2.7	Recent and current series of satellite altimeters	60
2.8	Recent and current satellite synthetic aperture radars	61
2.9	Recent and current satellite scatterometers measuring wind speed and direction	61
2.10	Access to useful sources of satellite-derived ocean data products from the Web	62
2.11	Access to useful image data-viewing and manipulation tools	64
7.1	Relative contribution of different ocean zones to total primary production	265
7.2	High-resolution visible and near-infrared sensors	274
8.1	Altimeters providing measurements of wave height since 1978	308
8.2	Details of wave data products from Jason-1	310
8.3	Details of wave data products from Envisat RA-2	310
10.1	Parameterizations of gas transfer velocity	374
11.1	Changes in meteorological conditions around the world apparently affected by the occurrence of the "Warm Episode"	394
13.1	Length and timescale of shelf sea processes and phenomena observed from satellites	489
13.2	Sampling capabilities of sensors used to observe shelf seas	491
14.1	The main classes of SST sensor systems on satellites	577
14.2	Essential climate variables identified by GCOS	593
14.3	GCOS climate-monitoring principles	594
14.4	Specific GCOS monitoring principles applicable to satellite systems	596
14.5	ECVs for the oceanic domain, largely dependent on satellite observations, showing corresponding data products and FCDRs required	597
14.6	GCOS requirements for datasets and products used for ECVs	598
15.1	Ocean variables required to be monitored operationally from satellites	612

Abbreviations and names of satellites and sensors

Satellites or satellite series are indicated by † and sensors by *.

2D-FFT	Two Dimensional Fast Fourier Transform
AATSR*	Advanced Along Track Scanning Radiometer
ABL	Atmospheric Boundary Layer
ACC	Antarctic Circumpolar Current
ACC (Bdy)	Southern boundary of the ACC
ACW	Antarctic Circumpolar Wave
ADCP	Acoustic Doppler Current Profiler (*in situ* ocean instrument)
ADEOS†	ADvanced Earth Observing System (Japanese polar-orbiting platform)
ADT	Absolute Dynamic Topography
AIRS*	Atmospheric Infrared Sounder
AIS	Automatic Identification System
ALOS	Advanced Land Observing Satellite
AMI*	Advanced Microwave Instrument (dual SAR/scatterometer on ERS-1 and 2)
AMSR	Advanced Microwave Scanning Radiometer
AMSR-E*	AMSR version flown on Aqua
AMSU-A*	Advanced Microwave Sounding Unit A (for atmospheric sounding)
ANN	Artificial Neural Net
Aqua†	NASA's EOS polar-orbiting platform with an afternoon overpass
Argo	A system of globally distributed floats which profile the ocean's density structure
ASAR*	Advanced SAR on Envisat
ASCAT*	Advanced SCATterometer flown on MetOp

ASST	Averaged SST
ATBD	Algorithm Theoretical Basis Document
ATSR*	Along Track Scanning Radiometer
AVHRR*	Advanced Very High Resolution Radiometer
B	Bacteria
CASI	Compact Airborne Spectral Imager
CDOM	Colored Dissolved Organic Matter
CEA	*Commissariat à l'Energie Atomique*
CEI	Chlorophyll Extension Index
CEOS	Committee on Earth Observing Satellites
CERSAT	Center for Satellite Exploitation and Research at IFREMER, France
CF	Coriolis Force
Chl	Chlorophyll
CHRIS*	Compact High Resolution Imaging Spectrometer
CLS	*Collecte Localisation Satellites* (French research company)
CMIS*	Conically scanned Microwave Imager and Sounder
CMOD	Empirical model of C-band microwave backscatter *vs.* wind speed and direction
CNES	*Centre National d'Etudes Spatiales* (French Space Agency)
COARE	Coupled Ocean–Atmosphere Response Experiment
Coriolis†	Satellite of U.S. Naval and Air Force Research Laboratories
CROZEX	CROZet natural iron bloom EXperiment
CRW	Coral Reef Watch
CSET	Cross Shore Ekman transport
CZCS*	Coastal Zone Color Scanner
DCM	Deep Chlorophyll Maximum
DHW	Degree Heating Week
DIC	Dissolved Inorganic Carbon
DMC	Disaster Monitoring Constellation
DMS	Dimethyl Sulfide
DMSP†	Defense Meteorological Satellite Program (U.S.)
DON	Dissolved Organic Nitrogen
DUACS	Data Unification And Combination System
EAP	East Atlantic Pattern
ECMWF	European Center for Medium-range Weather Forecasts
ECV	Essential Climate Variable
EEZ	Exclusive Economic Zone
EIGEN	European Improved Gravity model of the Earth by New techniques
EIGEN-GRACE03S	A GRACE-based static gravity field model derived within the EIGEN initiative
EKE	Eddy Kinetic Energy
EKW	Equatorial Kelvin Wave

EM	ElectroMagnetic
EMSA	European Maritime Safety Agency
EnKF	Ensemble Kalman Filter
ENSO	El Niño–Southern Oscillation
Envisat†	Major polar platform for ESA's Earth Observing System
EO	Earth Observation (typically refers only to satellite observations)
EOF	Empirical Orthogonal Function
EOS	Earth Observing System
ERA-40	ECMWF Re-Analysis of meteorological variables for 1957–2001
ERD	Environmental Research Division
ERS-1, 2†	ESA Remote Sensing satellite series
ERSEM	European Regional Seas Ecosystem Model
ERSST	Extended Reconstructed Sea Surface Temperature
ESA	European Space Agency
EUMETSAT	EUropean Organization for the Exploitation of METeorological SATellites
EuroGOOS	European Global Ocean Observing System
fAPAR	fraction of Absorbed Photosynthetically Active Radiation
FCDR	Fundamental Climate Data Record
FOV	Field Of View
GANDER	Global Altimeter Network Designed to Evaluate Risk
GCM	General Circulation Model (of the ocean or atmosphere)
GCOM	Global Change Observation Mission
GCOS	Global Climate Observing System
GCR	Global Climate Record
GDR	Geophysical Data Record (from an altimeter)
GDS	GHRSST Data Specification
GEO	Group on Earth Observations
GEOSS	Global Earth Observing System of Systems
GFO	Geosat Follow On
GHRSST	Group for High Resolution Sea Surface Temperature
GHRSST-PP	GODAE High Resolution SST Pilot Project
GLI*	GLobal Imager (Japanese visible, near-IR, and thermal IR sensor)
GLONASS†	Global Navigation Satellite System
GMES	Global Monitoring for Environment and Security
GNSS-R	Global Navigation Satellite System Reflectometry
GOCE*	Gravity and Ocean Circulation Explorer
GODAE	Global Ocean Data Assimilation Experiment
GOOS	Global Ocean Observing System
GOSIC	Global Observing Systems Information Center
GPCP	Global Precipitation Climatology Project
GPS	Global Positioning System

GRACE*	Gravity Recovery And Climate Experiment
GSFC	Goddard Space Flight Center
GW	Great Whirl (eddy in the Arabian Sea)
HAB	Harmful Algal Bloom
HDF	Hierarchical Data Format
HNLC	High Nutrient Low Chlorophyll
HOAPS	Hamburg Ocean–Atmosphere Parameters and Fluxes from Satellite data
HR-DDS	High Resolution Diagnostic DataSet
HRV	High Resolution Visible
IFOV	Instantaneous Field Of View
IFREMER	*Institut français de recherche pour l'exploitation de la mer*, translated as French Research Institute for Exploitation of the Sea
IGDR	Interim Geophysical Data Record (from an altimeter)
IM	Image Mode
IOCCG	International Ocean Color Co-ordinating Group
IOP	Inherent Optical Property
IPCC	Intergovernmental Panel on Climate Change
IR	InfraRed
ISCCP	International Satellite Cloud Climatology Project
ISW	Internal Solitary Wave
IT	Internal Tide
ITCZ	InterTopical Convergence Zone
IW	Internal wave
JAMSTEC	Japan Agency for Marine–Earth Science and TEChnology
Jason-1, 2	Satellites for altimetry continuing the series started by T/P
JCOMM	Joint Committee for Oceanography and Marine Meteorology
JGOFS	Joint Global Ocean Flux Study (international collaborative research project)
K-dV	Korteweg–de Vries
L2P	Level 2 Preprocessed
LAI	Leaf Area Index
LIDAR	LIght Detection And Ranging
LTSRF	Long Term Stewardship and Reanalysis Facility
MCC	Maximum Cross Correlation
MCS	Marine Core Service
MCSST	MultiChannel Sea Surface Temperature
MDT	Mean Dynamic Topography
MERIS*	MEdium Resolution Imaging Spectrometer
MERSEA	Marine Environment and Security for the European Area
MetOp†	European polar-orbiting operational meteorological satellite series (ESA/Eumetsat)
MFS	Mediterranean Forecasting System

MISST	Multi-sensor Improved Sea Surface Temperature for GODAE
MJO	Madden–Julian Oscillation
MOI	Mediterranean Oscillation Index
MODIS*	MODerate-resolution Imaging Spectrometer
MPI	A method of SAR wave spectrum inversion developed at the *Max Planck Institut für Meteorologie* at Hamburg
MSHED	Multi Sensor Histogram Edge Detection
MSL	Mean Sea Level
MSS	Mineral Suspended Sediment
MTF	Modulation Transfer Function
MTOFS	*Measuring the Oceans from Space*, the companion volume (Robinson, 2004)
MW	MicroWave
MWR	MicroWave Radiometer
NAO	North Atlantic Oscillation
NASA	National Aeronautics and Space Administration (U.S.)
NBC	North Brazil Current
NCAR	National Center for Atmospheric Research
NCEP	National Center for Environmental Prediction (U.S.)
NECC	North Equatorial Counter-Current
NEODAAS	NERC Earth Observation Data Acquisition and Analysis Service
NetCDF	Network Common Data Format
NOAA	National Oceanographic and Atmospheric Administration
NOP	Numerical Ocean Prediction
NPOESS†	National Polar-orbiting Operational Environmental Satellite System (U.S.)
NPP†	NPOESS Preparatory Project
NSIDC	National Snow and Ice Data Center
NWP	Numerical Weather Prediction
OA	Objective Analysis
OCTS*	Ocean Color and Temperature Sensor (Japanese, on ADEOS-1)
OFS	Ocean Forecasting System
OI	Optimal Interpolation
OLCI	Ocean and Land Color Instrument
ONI	Ocean Niño Index
OOS	Ocean Observing System
OSCAR	Ocean Surface Current Analysis–Real time
OSDR	Operational Sensor Data Record
OSI-SAF	Oceans and Sea Ice Satellite Applications Facility
OSTIA	Operational sea Surface Temperature and sea Ice Analysis
P	Phytoplankton
PALSAR	Phased Array L-band Synthetic Aperture Radar

PAR	Photosynthetically Available Radiation
PDF	Probability Distribution Function
PF	Pressure Force
PF	Polar Front (in Sections 4.5.3 and 4.6.1)
PFZ	Potential Fishing Zone
PMEL	Pacific Marine Environmental Laboratory
POM	Primary Ocean Measurement
PON	Particulate Organic Nitrogen
Poseidon*	Radar altimeter (CNES)
PSB	Patagonian Shelf Break
PSR	Photosynthetically Stored Radiation
QuikScat*	Satellite with a dedicated scatterometer mission (U.S.)
r.m.s.	root mean square
ROFI	Region Of Freshwater Influence
ROWS*	Radar Ocean Wave Spectrometer
sACCf	southern ACC front
SAF	Sub Antarctic Front
SAR*	Synthetic Aperture Radar
SCIAMACHY*	SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY
SeaWiFS*	Sea-viewing Wide Field-of-view Sensor
SeaWinds*	Ku-band scatterometer flown on QuikScat (NASA)
SERIES	Subarctic Ecosystem Response to Iron Enrichment Study
SEVIRI†	Spinning Enhanced Visible and Infrared Imager
SG	Southern Gyre (eddy in the Arabian Sea)
SIOP	Specific Inherent Optical Property
SIRAL	Synthetic Aperture Interferometric Radar Altimeter on the ESA CryoSat mission
SIZ	Seasonal Ice Zone
SLA	Sea Level Anomaly
SLAR	Side-looking airborne radar
SLC	Single Look Complex
SLSTR	Sea and Land Surface Temperature Radiometer
SMMR	Scanning Multichannel Microwave Radiometer
SMOS†	Soil Moisture and Ocean Salinity (satellite mission)
SOFeX	Southern Ocean iron (Fe) Experiment
SOI	Southern Oscillation Index
SOIREE	Southern Ocean Iron Enrichment Experiment
SPM	Suspended Particulate Material
SPOT†	*Satellite probatoire d'observation de la Terre* (CNES)
SPRA	Semi Parametric Retrieval Algorithm
SRAL	Synthetic Aperture Interferometric Radar Altimeter for ESA Sentinel-3 satellites
SSALTO	*Segment Sol multimissions d'ALTimétrie, d'Orbitographie et de localisation precise*

SSHA	Sea Surface Height Anomaly
SSI	Surface Solar Irradiance
SSM/I*	Special Sensor Microwave Imager
SSM/T-2*	Special Sensor Microwave Temperature Sounder
SSS	Sea Surface Salinity
SST	Sea Surface Temperature
SSTL	Surrey Satellite Technology Ltd.
STAR	Center for Satellite Applications and Research
SWH	Significant Wave Height
SWIMSAT†	Proposed real-aperture radar system to measure directional spectra of ocean waves from space
T/P	TOPEX/Poseidon
TC	Tropical Cyclone
TCDR	Thematic Climate Data Record
TCHP	Tropical Cyclone Heat Potential
Terra	NASA's EOS polar-orbiting platform with a morning overpass
TIW	Tropical Instability Wave
TM*	Thematic Mapper
TMI	TRMM Microwave Imager
TOA	Top Of Atmosphere
TOGA	Tropical Ocean–Global Atmosphere
TOPEX	TOPographic EXperiment: radar altimeter (NASA)
T/P	TOPEX/Poseidon altimetry mission (NASA/CNES)
TRMM†	Tropical Rainfall Measuring Mission
TSM	Total Suspended Matter
TSS	Total Suspended Sediment
TUI	Temperature-based Upwelling Index
TZCF	Transition Zone Chlorophyll Front
UI	Upwelling Index
UNFCCC	United Nations Framework Convention on Climate Change
VGPM	Vertically Generalized Production Model
VHRR*	Very High Resolution Radiometer
VIIRS*	Visible and Infrared Imager Radiometer Suite
WCRP	World Climate Research Program
Windsat*	Multifrequency polarimetric microwave radiometer flown on Coriolis
WM	Wave Mode
WWW	World Weather Watch
Z	Zooplankton

Symbols and nomenclature

Although the bulk of this book is not written from a theoretical standpoint, there are a few places where it is convenient to express the scientific principles in mathematical terms. This list provides a reference defining the symbols used. Some are common throughout the book, but most are limited to one or two chapters. There is some duplication where a symbol has different meanings in different chapters or sections. For this reason, the symbols are listed mainly by chapter. To avoid ambiguity, care must be taken to relate the definition to the chapter or section.

Symbol	*Property represented*	*Units (if applicable)*	*Section*
	Symbols common throughout the book		
f	Coriolis parameter	s^{-1}	
f_0	Central value of f in a latitude range	s^{-1}	
g	Acceleration due to gravity	m s^{-2}	
u	Eastward (or meridional) component of sea surface velocity	m s^{-1}	
v	Northward (or zonal) component of sea surface velocity	m s^{-1}	
x	Eastward component of distance	km	
y	Northward component of distance	km	
z	Vertical distance co-ordinate (positive upwards or downwards depending on context)		
β	Gradient of Coriolis parameter with latitude	s^{-1} deg^{-1}	
φ	Latitude	deg	
ρ	Seawater density	kg m^{-3}	
σ_0	Normalized radar backscatter cross-section		
Ω	Earth rotation rate	s^{-1}	3, 6

(*continued*)

Symbol	*Property represented*	*Units (if applicable)*	*Section*
	Symbols in Chapter 2 (some ambiguity between sections)		
B_f	Microwave spectral radiance per unit frequency bandwidth	W m^{-2} str^{-1} s	2.4.4
G	Constant of gravitation		2.2
H_{sat}	Height of satellite above reference ellipsoid	m	2.4.5
$H_{1/3}$	Significant wave height	m	2.4.5
h	Height of satellite above the ground	m	2.2
h	Height of sea surface above reference ellipsoid	m	2.4.5
h_{atm}	Displacement of sea surface caused by atmospheric pressure	m	2.4.5
h_{dyn}	Dynamic height of sea surface above the geoid after tidal and atmospheric pressure corrections	m	2.4.5
h_{geoid}	Height of geoid above reference ellipsoid	m	2.4.5
h_{tide}	Tidal displacement of sea surface	m	2.4.5
h_{SSHA}	Local sea surface height anomaly	m	2.4.5
k	Boltzmann's constant	1.38×10^{-23} J K^{-1}	2.4.4
$L(\lambda, T)$	Spectral radiance per unit wavelength	W m^{-2} m^{-1} str^{-1}	2.4.3
M	Mass of the Earth	kg	2.2
r	Distance above the Earth center of a satellite orbit	km	2.2
R	Earth radius (~6,378 km)	km	2.2
R_{alt}	Altimetry-measured distance between satellite and sea surface	m	2.4.5
R_λ	Radiance reflectance at optical waveband centered at wavelength λ		2.4.2
S_n	Signal for channel n of an IR radiometer		2.4.3
T	Orbit period of an Earth-orbiting satellite	s	2.2
T	Temperature	K	2.4.3
T_{bn}	Radiance expressed as equivalent blackbody brightness temperature for channel n of an IR radiometer	K	2.4.3
T_s	Sea surface temperature	K	2.4
$\Delta_{i,j}(T_b)$	Difference between T_b in band j and band i	K	2.4.3
λ	Wavelength	m	2.4
	Symbols in Chapter 3		
c	Cross-frontal offset in a front detection algorithm	m	3.4.2
h	Depth of water in calculation of longwave speed	m	3.2
h	Sea surface height anomaly in calculating ocean kinematic properties	m	3.4
h_1	Thickness of upper layer of a two-layer ocean	m	3.2
L	Characteristic lengthscale of a fluid phenomenon	m	3
L_{Rb}	Baroclinic Rossby radius of deformation	m	3
R_e	Reynolds number	dimensionless	3

Symbol	*Property represented*	*Units (if applicable)*	*Section*
	Symbols in Chapter 3 (*cont.*)		
S_n	Normal component of strain in ocean surface current field	s^{-1}	3.4
S_S	Shear component of strain in ocean surface current field	s^{-1}	3.4
V	Characteristic flow velocity in a fluid phenomenon	$m\ s^{-1}$	3
W	Okubo–Weiss parameter	s^{-2}	3.4
ν	Kinematic viscosity	$m^2\ s^{-1}$	3
ρ_0	Mean density over the water column	$kg\ m^{-3}$	3
ω	Vorticity of ocean surface current about vertical axis	s^{-1}	3.4
	Symbols in Chapter 4		
a	Lengthscale characterizing the width of a front		4.3.2
T_b	Half-magnitude of the frontal temperature difference	K	4.3.2
T_p	Modeled frontal temperature profile	K	4.3.2
T_0	Temperature at frontal line	K	4.3.2
U_D	Component of apparent surface current in the radar range direction detected by Doppler centroid analysis	m^{-1}	4.2.5
x'	Pixel co-ordinate within an image extract, orthogonal to y'		4.3.2
y'	Pixel co-ordinate within an image extract, orthogonal to x'		4.3.2
θ	Orientation of a front relative to the y' co-ordinate		4.3.2
	Symbols in Chapter 6		
c_x	Wave phase speed in the x direction	$m\ s^{-1}$	6.3
c_{nx}	Eastward phase speed of the nth mode Rossby wave	$m\ s^{-1}$	6.3
k	Zonal wavenumber of a Rossby wave	m^{-1}	6.3
l	Meridional wavenumber of a Rossby wave	m^{-1}	6.3
V	Representative dynamical variable within a Rossby wave		6.3
x'	Dummy variable within the Radon transform		6.4
y'	Dummy variable within the Radon transform		6.4
θ	Orientation relative to the time co-ordinate in a time–longitude plot, defining a variable of the Radon transform		6.4
λ_n	Rossby radius of deformation for the nth mode	m	6.3
ω	Frequency of a Rossby wave	s^{-1}	6.3
ω_n	Frequency of the nth mode Rossby wave	s^{-1}	6.3

(*continued*)

Symbol	*Property represented*	*Units (if applicable)*	*Section*
	Symbols in Chapter 7		
a_c^*	Chlorophyll-specific light absorption coefficient	m^2 $(gChl)^{-1}$	7.3
C	Chlorophyll concentration	gC m^{-3}	7.3
E_{PAR}	Photosynthetically available radiation	Ein m^{-2} s^{-1}	7.3
J_C	Chemical energy equivalent of a unit mass of carbon fixed	kJ $(gC)^{-1}$	7.3
P	Primary production rate, at which carbon is fixed by photosynthesis per unit time per unit volume of water	gC m^{-3} s^{-1}	7.3
P_{tot}	Primary production rate integrated through the water column	gC m^{-2} s^{-1}	7.3
φ_μ	Quantum yield	molC Ein^{-1}	7.3
ψ^*	Cross-section for photosynthesis	m^2 $(gChl)^{-1}$	7.3
	Symbols in Chapter 8		
H_s	Significant wave height	m	8
h	Depth of water in calculating gravity wave speeds	m	8
k	generic wavenumber for surface waves	m^{-1}	8
S	Ocean wave energy spectrum		8
T	Period of an ocean surface wave	s	8
T_m	Mean period of a sea surface gravity wave field	s	8
T_p	The peak period (at the maximum of surface wave spectrum)	s	8
T_z	The zero-crossing period of a surface gravity wave field	s	8
u^*	Friction velocity	m s^{-1}	8
V_{gr}	Group velocity of sea surface gravity waves	m s^{-1}	8
V_{ph}	Phase speed of sea surface gravity waves	m s^{-1}	8
ζ	Vertical displacement of sea surface by surface waves	m	8
σ_ζ	Standard deviation of sea surface wave elevation	m	8
	Symbols in Chapter 10		
C_D	Drag coefficient over the sea		
C_E	Dalton number (transfer coefficient for latent heat flux)		
C_p	Phase speed of the dominant frequency in the ocean wave spectrum		
C_T	Stanton number (transfer coefficient for sensible heat)		
c_p	Specific heat of air at constant pressure		
D	Diffusivity of specified gas in seawater	m^2 s^{-1}	
F	Generic flux of some property from ocean to atmosphere		10.2.2
F_{gas}	Flux of a specified gas from ocean to atmosphere	mole(gas) m^{-2} s^{-1}	

Symbol	*Property represented*	*Units (if applicable)*	*Section*
	Symbols to Chapter 10 (*cont.*)		
H_s	Significant wave height	m	
K_x	Generic transfer coefficient for sea–air flux		10.2.2
k	Transfer velocity of specified gas across the sea–air interface	$m\ s^{-1}$	
k_b	Bubble-mediated gas transfer velocity		10.4
k_d	Direct gas exchange transfer velocity		10.4
L	Latent heat of vaporization of water at T_a	$J\ kg^{-1}$	
L_p	Wavelength of the dominant frequency in the ocean wave spectrum		
pX_a	Partial pressure of specified gas on the air side of the interface		
pX_w	Partial pressure of specified gas on water side of the interface		
Q	Net heat exchange at the sea surface (positive out of the sea)	$W\ m^{-2}$	
Q	Algorithm-derived estimate of surface specific humidity		10.3.5
Q_b	Net longwave radiation	$W\ m^{-2}$	
Q_E	Latent heat flux	$W\ m^{-2}$	
Q_H	Sensible heat flux	$W\ m^{-2}$	
Q_S	Net shortwave radiation	$W\ m^{-2}$	
q_s	Specific humidity of air at the sea surface		
q_z	Specific humidity of air at the reference height z		
R	Generic interface impedance for sea–air flux		10.2.2
R	Correction factor to allow for non-linear wind dependence when monthly averaged winds are used for flux integrations		10.4
$R_L^{\downarrow}$	Longwave back radiation from atmosphere to ocean	$W\ m^{-2}$	10.4.1
Sc	Schmidt number (μ/D)		
s	Solubility of the specified gas in water at the temperature and salinity of the sea surface skin		
T_a	Air temperature (at reference height z above the sea surface)	K	
T_s	Air temperature at the sea surface	K	
u^*	Friction velocity	$m\ s^{-1}$	
u_s	Horizontal wind speed at the sea surface	$m\ s^{-1}$	
u_z	Horizontal wind speed at the reference height, z, above the surface	$m\ s^{-1}$	10.2.2
W	Atmospheric total precipitable water derived by microwave radiometry		10.3.6
W_b	Total column water vapor derived by microwave radiometry		10.3.6
X	Generic ocean property associated with a sea–air flux		

(*continued*)

Symbol	*Property represented*	*Units (if applicable)*	*Section*
	Symbols to Chapter 10 (*cont.*)		
z_0	Roughness height of the sea surface		
α	Charnock constant		
μ	Kinematic viscosity of water	$m^2\ s^{-1}$	
ρ	Surface air density (at temperature T_a and pressure p_z at the reference height z above the sea surface)	$kg\ m^{-3}$	
τ	Wind stress	$N\ m^{-2}$	
	Symbols in Chapter 12		
c	Slope relative to horizontal of internal wave rays		12.1.1
c	Coefficient in K-dV equation defined in Equation (12.5)		12.2.2
$c(z)$	Depth distribution of Chlorophyll concentration	$gChl\ m^{-3}$	12.3.2
c_g	Group speed of the Bragg-scattering sea surface waves	$m\ s^{-1}$	
c_p	Phase speed of the Bragg-scattering sea surface waves	$m\ s^{-1}$	
C	Phase speed of internal soliton (see Equation 12.9)	$m\ s^{-1}$	
H_1, H_2	Thickness of upper, lower layers of a two-layer ocean	nm	
$g(z)$	Depth weighting of remotely sensed optical signal		12.3.2
h	Depth of total water column in two-layer ocean	m	
K	Optical diffuse attenuation coefficient	m^{-1}	12.3.2
N	Buoyancy, or Brunt–Väisälä, frequency	s^{-1}	
r	Ratio of H_1 and H_2		
U_x	Radar range component of the surface velocity field associated with the internal wave train	$m\ s^{-1}$	
Z_{90}	Light penetration depth	m	12.3.2
α	Coefficient in K-dV equation defined in Equation (12.6)		
γ	Coefficient in K-dV equation defined in Equation (12.7)		
η	Interface displacement in internal wave equation	m	
ρ_1, ρ_2	Density of upper, lower layers of a two-layer ocean	$kg\ m^{-3}$	
σ	Frequency of internal waves	s^{-1}	
τ	Relaxation time of the modulated Bragg waves	s	
λ	Notional wavelength of an internal soliton (see Equation 12.10)	m	
	Symbols in Chapter 13		
h	Depth of the water column	m	
U	Characteristic amplitude of tidal current	$m\ s^{-1}$	
u	Mean depth-averaged amplitude of the dominant tidal harmonic current	$m\ s^{-1}$	

1

Introduction

1.1 AN IMPORTANT OBSERVATIONAL TOOL FOR PLANETARY SCIENCE

The ocean of planet Earth still holds many secrets. This book aims to show its readers how the use of remote-sensing devices on Earth-orbiting satellites has revealed hitherto unseen aspects of the sea. It points to new ways of understanding the ocean and new insights in ocean science, which have developed only since Earth observation (EO) technology granted us a unique vantage point in space from which to measure aspects of the ocean. It demonstrates the applications of "satellite oceanography", showing it to be an exciting tool which in future should unlock more of the ocean's mysteries. It also describes how the particular sampling capabilities of sensors above the Earth can be put to work in the more operational tasks of monitoring, forecasting, and managing the marine environment.

After a century in which explorer-scientists reached every part of every continent over land and ice we generally accept that there remain no significant geographical discoveries to be made in our world. After four decades of increasingly sophisticated technology—which have enabled us to descend into the ocean deeps, fly through the highest parts of the atmosphere, and probe the planet's interior with geophysical tools—there is also a tendency to assume that the science of the Earth and its environment is broadly understood, apart from clarification of some details. As a consequence popular opinion now looks beyond the Earth for a "final frontier" to explore. Indeed, such is the readiness to believe that the behavior of our own planet is known and predictable, that political leaders in technologically advanced nations talk of adopting the exploration of our neighboring planet Mars as a project to inspire the pioneering spirit of their people and to stimulate new technological endeavor. Yet it is profoundly mistaken to overlook the outstanding scientific challenges which still remain to understand the science of the Earth as a system. Moreover, it is foolish to ignore mankind's urgent requirement to be able to

monitor and predict changes of our own global environment, for on this will depend the future stability of human civilization.

In questions of how the Earth operates physically, chemically, and biologically as an integrated system, the role of the ocean is not fully grasped. Within the hydrosphere it is recognized that the ocean tends to have a stabilizing effect on physical climate, due to its much longer time constant for change than that of the atmosphere. Yet the large-scale and long-term processes in the ocean which determine its role in climate change are not properly known or understood, and neither is the relationship between processes occurring at different length and time scales. Because the ocean is a fluid it is constantly changing across a wide spectrum of scales. These span from centimeters and seconds for small surface waves to thousands of kilometers and several decades for the exchange of water in ocean basins between the surface layer and the abyss. Interactions between biological, chemical, and physical processes in the ocean can occur at all scales in between these extremes. It is the unique capacity of satellite remote-sensing systems to sample "snapshots" of the detailed spatial distribution of ocean variables over hundreds to thousands of kilometers, repeating those measurements regularly for decades, which gives them a key role in measuring and then understanding ocean variability. This book will show a variety of ways in which satellite data have begun to open up new opportunities for scientific study of the ocean, and point to the long-term scientific role which the methods of satellite oceanography should occupy in the future.

Moreover, beyond a significant contribution to improving scientific understanding, there is a further important role for satellite ocean remote sensing in a number of operational settings where that knowledge can be applied. The regular repeated sampling that is a characteristic of remote sensing allows routine monitoring of ocean parameters, yielding observations that can benefit various different users of the sea, especially when the data are supplied in a timely way. These include information needed by mariners about waves, wind, and currents, or data for environmental quality managers concerning natural phenomena such as the occurrence of algal blooms or anthropogenic events like oil spills. Even when sampling of certain parameters by satellites is not frequent enough to support effectively continuous monitoring, the long-term accumulation of data can be used for constructing probabilities of extreme events, which is another valuable contribution in support of marine-related industries. Examples of these and similar applications of the methods of satellite oceanography to operational and commercial tasks are also presented in this book, with the emphasis on cases where the satellite data make a unique contribution that is not matched by conventional measurements.

While satellite data have been used in ocean science for over 25 years, the most significant advances in measurement methods have come in the last 10 years, leading to a sufficient variety of applications to allow a book like this to be selective in presenting the most successful, interesting, or innovative examples within each of the topics chosen to span the breadth of oceanography. Yet perhaps just as important as the examples presented here is the inspiration that they may give to some of those who read this book, to first dream, then develop, and finally demonstrate new applications of satellite ocean data within their own field of interest. There are many

important questions and tasks facing ocean and Earth system scientists at the start of the 21st century. My hope is that some of those reading this book will decide that understanding the ocean is just as challenging and worthy of their creative endeavor as probing the Solar System. Besides, it might turn out to be rather more crucial for the security and comfort of future generations of mankind that we are able to monitor effectively the environmental health of our planet through the pulse of the ocean.

A parallel objective for the book is to assemble a collection of examples which demonstrate the effectiveness of satellite data when applied to oceanography and together make a compelling case for maintaining the satellite programs on which they are based. Until recently much of the cost of space hardware and development has been financed with the aim of promoting technological innovation, but those circumstances are changing. The time is approaching when the users of satellite data will need to pay (probably indirectly rather than directly) to ensure the continuity of satellite programs. Since the whole oceanographic community will be called on to justify the costs involved, it is important that they are made aware of how important satellite data have become to a number of branches of marine science. Hopefully this book will contribute significantly to that awareness.

1.2 PUTTING REMOTE SENSING TO WORK FOR OCEANOGRAPHERS

The underlying aim of this book is to educate the reader in the ways of using satellite data to discover new perspectives of the ocean which extend their understanding of oceanography and ocean processes. It is not a direct exposition of the basic principles and methods of ocean remote sensing, since those are already addressed in some detail by Robinson (2004) in the companion volume, *Measuring the Oceans from Space* (hereafter *MTOFS*). Rather, by demonstrating and explaining how satellite ocean data are currently used in a variety of situations for novel scientific research and important operational tasks, the application of the methods described in *MTOFS* will be illustrated. Therefore this volume completes the story started in *MTOFS* so that the two volumes together offer a comprehensive view across the body of knowledge sometimes called "satellite oceanography". If the first volume presented the tools of ocean remote sensing, this volume shows how they can be put to work to produce results of real benefit both to marine science and to operational oceanography.

However, while the two books are intended to complement each other with little overlap, they have both been written to be complete within themselves. A reader who is unsure of what satellite data have to offer can learn from the oceanographic examples presented here in *Discovering the Ocean from Space* whether remote sensing promises results in their own particular field of research or operational interest. It aims to be an accessible way into the subject for the general oceanographic reader, before they embark on studying the basic methodology of sensors and measurement techniques in *MTOFS*.

Emphasis in this book is therefore placed on what can be learned about the ocean rather than how remote-sensing techniques work. Most of the chapters which follow are devoted to a particular type of phenomenon or field of oceanographic science, and how this is observed from satellites. Within each chapter the various remote-sensing techniques capable of observing the phenomenon are described, but the reader is referred to *MTOFS* for any detailed discussion of sensor and data-processing methodology. Exceptions to this are where a special methodology needs to be explained because it is essential for a particular application. This is the case where a synergetic approach uses data from more than one type of sensor. Because *MTOFS* is structured around chapters based on particular techniques it has little opportunity to discuss multisensor analysis and processing methods, which are therefore given further consideration here.

Most chapters also contain a simple introduction to background oceanographic theory, so that the treatment of each topic is self-contained. However, the aim here is to provide just sufficient explanation to clarify the contribution which satellite data make to an improved understanding of the phenomenon, and then point the reader to further references if they want to study the oceanography of the phenomenon in more depth. The focal points for each chapter are the particular examples of observations of ocean processes or phenomena, of which there may be several separate case studies. The discussion of each of these is intended to clarify first of all the nature of the information that can be derived from remote sensing, and then to explore the extent to which satellite data provide unique measurements or insights. In some cases it becomes clear that the richest harvest of new knowledge comes from combining satellite and *in situ* measurements. Sometimes the satellite contribution is only marginal, while in others it will be seen that the advent of Earth observation techniques has revolutionized the way oceanographers go about their work.

Where appropriate each chapter presents examples not only of cases where new scientific insights are revealed but also of operational applications of the technique. An underlying theme is to consider the influence which the availability of satellite data has on the quality with which an operational activity can be performed, as well as the converse—what would be the consequence if the satellite data were no longer available? This is consistent with one of the aims, mentioned in Section 1.1, to provide a broad critique of the impact of remote-sensing data in oceanography. The discussion also looks forward to the potential for improving the techniques and the possible gains counted in terms of either a fuller understanding of how the ocean works, or a better ability to monitor and forecast it.

1.3 THE OCEANOGRAPHIC SCOPE OF THE BOOK

As explained above, the point of view of the whole book is directed at the ocean phenomena and processes which are particularly well revealed by satellite data, without expounding in any depth the methodologies of remote sensing, which can be found in *MTOFS*. Nonetheless, Chapter 2 does give an introduction to the principal sensors used for ocean remote sensing, identifies the unique sampling

characteristics of instruments on satellites and notes the stages of data processing required to convert raw measurements into oceanographically useful data. The emphasis is on those factors which give satellite ocean data their distinctive characteristics, both strengths and weaknesses. This brief summary of ocean remote-sensing methodology ensures that the book provides a self-contained introduction to satellite oceanography. It can therefore serve as a textbook for a course in ocean remote sensing where the emphasis is on applications rather than techniques.

The next three chapters commence the survey of oceanography viewed from space by exploring the ways in which our knowledge of mesoscale ocean processes has benefited from remote sensing. Chapter 3 looks at mesoscale eddies, Chapter 4 at ocean fronts, and Chapter 5 at upwelling and other related ocean features. Each of these phenomena are readily observed by several different satellite oceanography techniques. Chapter 6 then considers a particular class of large-scale ocean features whose length and timescales make them particularly well suited to satellite detection. These are the large-scale, wave-like phenomena which propagate in a regular way across ocean basins, such as Rossby waves and tropical instability waves. For many years these features were very difficult to detect at all, so that their unambiguous signatures in global image datasets represents a singular success for satellite oceanography. Chapter 7 moves beyond physical oceanography to review what can be learned about ocean biology from space. It touches on a variety of topics, including primary production, fisheries, and coral reefs.

The next three chapters explore what satellites can tell us about physical processes occurring at the air–sea interface, which is of course that part of the ocean viewed directly by most remote Earth-observing sensors on satellites. Chapter 8 considers the phenomenon of ocean surface gravity waves, differentiating between wave properties detected by particular sensor classes and pointing out how remotely sensed wave data differ from conventional measurements. Chapter 9 examines ways in which the measurement of the wind over the sea and its detailed spatial variability are used for marine applications. Chapter 10, written with Susanne Fangohr, looks at how the fluxes of heat and gases between the ocean and the atmosphere can be estimated from satellite data, an approach combining data from several sensors that is still being developed but shows considerable promise.

Chapter 11 develops the theme of air–sea interaction processes at the larger scale and shows how remote sensing can give us a unique view of several important topics. These include El Niño, Indian Ocean monsoons, the distribution of sea ice, low-frequency variability of sea surface height and secular changes in sea level. Satellite data provide a unique global perspective on these phenomena, which contribute to what we experience as short-timescale climate variability.

Two more chapters consider ocean phenomena at somewhat smaller scales. Chapter 12, written with José Carlos da Silva, is devoted to the phenomenon of internal waves, a subject that has advanced considerably by the use of remote-sensing methods. Chapter 13 looks more generally at phenomena found in shelf seas, touching on topics such as seasonal stratification and associated tidal mixing fronts, the occurrence of algal blooms, and the monitoring of water quality parameters. It also touches on how remote sensing can be used for monitoring coasts and

estuaries. Here the scales of interest tend to be smaller than is optimum for using satellite data, but used in conjunction with more conventional data sampling they still have an important role to play.

The remaining substantive chapter (Chapter 14) directly addresses the way in which satellite data are incorporated into operational ocean monitoring and forecasting. It includes a short section on oil spill detection, monitoring, and management, although most of the chapter picks up a rather different theme, describing how satellite data can be integrated with measurements from *in situ* sensors by the use of numerical, model-based, ocean-observing systems. This approach seeks to maximize the complementarity of the various observational and modeling tools now available to oceanographers, with the aim of establishing global, regional, and local ocean-forecasting systems comparable with the way in which meteorological observations are nowadays integrated into numerical weather prediction models. Although such ocean-forecasting schemes are only now in their infancy, they promise to make a fundamental difference to how we monitor the state variables of the ocean, not only for scientific analysis but also to provide essential operational information for users of the sea. Chapter 14 also introduces the new ways in which satellite measurements of the ocean are being used to produce essential climate variables, reliably calibrated records spanning many years, from which evidence of trends and long-period variability will allow the role of the ocean in global climate change to be characterized. If ongoing programs of ocean-monitoring sensors on satellites are to be extended indefinitely into the future, then they will need to be justified by these operational and climate-monitoring roles.

The short concluding chapter draws together the themes running through the book, with a discussion of what the near future may hold for the development of this interesting and important subject.

1.4 REFERENCE

Robinson, I. S. (2004), *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

2

The methods of satellite oceanography

2.1 OCEAN REMOTE-SENSING TECHNIQUES—A SUMMARY

This book is primarily about the oceanographic applications of remote sensing, and is written to complement the detailed descriptions and discussion of the techniques of satellite oceanography published in the companion volume *MTOFS* (Robinson, 2004). However, it would be unfortunate if a reader new to the whole field were left in complete ignorance of how the data presented in the rest of this book have been acquired. This chapter is therefore provided to give a very basic introduction to the methods of ocean remote sensing. It aims to summarize the essentials of the subject and to provide the minimum knowledge that a university graduate in oceanography ought to have, without going into all the detail that someone working in the fields of satellite oceanography research or applications would need. Thus although it only skims the surface, the reader who needs further information to make an explanation clear, or who is stimulated to find out more about a particular technique, can be confident of being able to do so by consulting *MTOFS* (Robinson, 2004) where a much fuller set of relevant references will also be found. On the other hand, the reader who is already familiar with Robinson (2004) can safely skip this chapter.

Figure 2.1 illustrates schematically what is involved in measuring properties of the ocean using a sensor that is typically hundreds or thousands of kilometers from the sea surface. An electromagnetic signal of a particular kind leaves the sea carrying information about one of the primary observable quantities which are the *color*, the *radiant temperature*, the *roughness*, and the *height* of the sea. This signal must pass through the atmosphere where it may be changed, and where noise may be added to it, before it is received by the sensor which detects particular properties of the radiation and converts each measurement into a digital signal to be coded and sent to the ground. The sensor geometry restricts each individual observation to a particular instantaneous field of view (IFOV). In order to convert the numbers

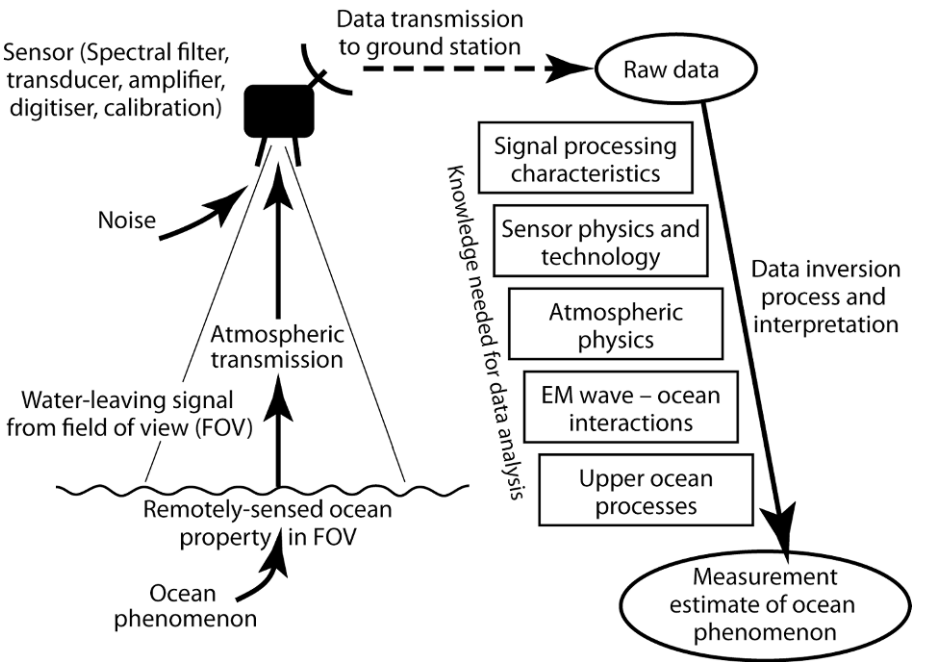


Figure 2.1. Schematic of information flow in ocean remote sensing.

received at the ground station into scientific measurements of useful precision and quantifiable accuracy, the remote-sensing process represented in the left-hand side of Figure 2.1 must be inverted digitally using the knowledge and information identified on the right-hand side.

It is their acquisition from a unique vantage point in space which gives satellite data their special character and so in Section 2.2 our brief summary of methods takes a look at the way image datasets are acquired, identifies the different satellite orbits available for remote sensing, and considers how both of these factors affect the space-time sampling capacity of the datasets. Then Section 2.3 gives an overview of the generic data-processing tasks that are implied in the right-hand side of Figure 2.1. These consist of operations that must be performed on the raw data received from the satellite in order to turn them into estimates or measurements of ocean parameters, with quantifiable accuracy, suitable for use in the scientific analysis or operational applications described in the rest of this book.

Section 2.4 introduces the diverse techniques of satellite oceanography, which use distinct parts of the electromagnetic spectrum and different aspects of radiation to measure particular properties of the ocean. Although several pages each are devoted to the different methods, ocean color, thermal infrared temperature detection, passive-microwave radiometry, altimetry, and oblique-viewing radars, these are very abbreviated descriptions of what are today extensive subjects each worthy of a book for themselves. A summary of the more important satellites and sensors used

for ocean remote sensing is provided by Section 2.5, and Section 2.6 concludes the chapter with guidance to readers on how to browse, access, and manipulate some of the wide variety of satellite ocean data available through the Internet.

2.2 THE UNIQUE SAMPLING CAPABILITIES OF SENSORS ON SATELLITES

The use of Earth-orbiting satellites as platforms for ocean-viewing sensors offers a number of unique advantages such as the opportunity to achieve wide synoptic coverage at fine spatial detail, and repeated regular sampling to produce time series several years long. It is these capabilities that distinguish satellite remote sensing from all other oceanographic observing techniques. The capacity for synoptic imaging depends primarily on the spatial sampling characteristics of the sensor, which are ultimately limited by detector sensitivity and the data flow capacity of the telecommunications system between the satellite and ground stations. Another set of limitations follow from the unavoidable constraints imposed by the physical laws of satellite orbital dynamics. Ultimately the sampling characteristics of different satellite oceanography methods depends on the sensor–platform combination. This section provides an outline of the important issues, but a more detailed discussion of sampling by remote sensors can be found in chapters 3 and 4 of *MTOFS* (Robinson, 2004).

2.2.1 Creating image-like data fields from point samples

What makes satellite data so useful and interesting for many users is their unique capability for dense two-dimensional spatial sampling which enables images to be formed corresponding to the surface distribution of the measured variable. But unlike the "snapshot" pictures we obtain from cameras, remotely sensed image data fields consist of millions of individual scientific measurements built up over a short length of time from a regular sampling pattern over the ground. Typically just one sensor is used, which ensures consistency of sensitivity for all the samples making up the image dataset. Those remote-sensing instruments that use an array of detectors must ensure uniform intercalibration of all elements.

The sea or ground area observed by a single detector is limited to its instantaneous field of view (IFOV) which is defined by a given directional spread relative to the pointing direction. Two-dimensional sampling to cover the sea surface is achieved by utilizing any relative motion between the platform and the ground and by pointing the sensor in a systematic sampling pattern. The instantaneously acquired measurement of an ocean property would be a single value representing the average property over the region defined by the intersection of the IFOV with the ground. However, because every sensor requires a finite time to record a measurement, during which the pointing of the sensor moves a finite distance over the ground, the effective "footprint" of each measurement must be somewhat larger

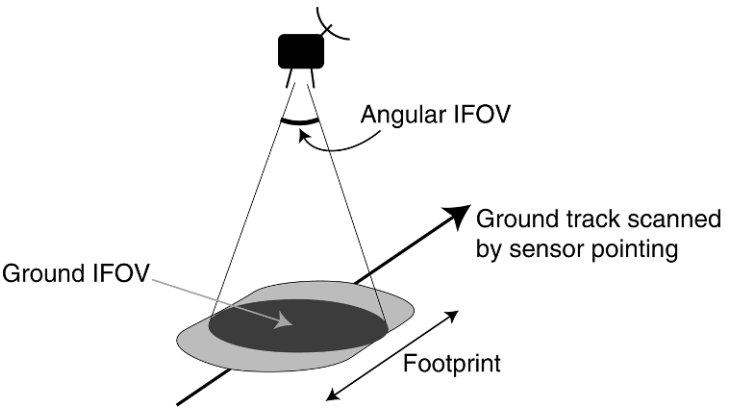


Figure 2.2. Sketch showing how the instantaneous field of view (IFOV) defines the measurement footprint during the sample integration time.

than the ground IFOV (as illustrated in Figure 2.2). The footprint determines the spatial resolution of the sensor.

Some sensors simply make downward-looking observations at periodic intervals while the satellite moves over the ground, to give an average value of radiation or other property from an area of the Earth surface, typically centered at the nadir point (the point on the Earth immediately below the satellite). A nadir-viewing altimeter is an example of this type of sensor. The only way to produce image-like datasets from such sensors is to wait until the satellite track has covered the ground with sufficient density to enable the variable to be smoothly mapped from the available point measurements, but this may take many days to achieve.

To obtain a truly near-instantaneous image requires a sensor that explicitly scans sideways across the satellite track direction. Figure 2.3 illustrates a typical arrangement where scan lines are perpendicular to the satellite track. Normally it is arranged for the sensor to scan a complete line and return to start the next in the time it takes for the satellite subpoint on the ground to travel a distance equal to the footprint size in the along-track direction. Thus the scan line spacing matches the sensor spatial resolution and adjacent scan lines are contiguous as shown. Along the scan lines, the sensor is arranged to take one sample in the time it takes for the pointing to swing through an angle equivalent to the IFOV, thus matching the sample spacing to the sensor resolution in the scan direction also. In this way a wide swath of ground is imaged, normally centered on the satellite subtrack.

The detailed scanning mechanism varies from sensor to sensor. Note that in general the footprint size increases and changes shape slightly towards the extremity of wide swath scan lines. In some cases the scan geometry may not be rectangular but curved, as when a conically scanning mirror is used instead of a rectangular scanning mirror. In the case of a geostationary satellite (see Section 2.2.2) the whole platform rotates about a north–south axis to achieve scanning parallel to lines of latitude. At the same time the sensor's field of view is rotated north–south to point at different latitudes for each satellite rotation, thus achieving a coverage of the whole face of the globe as visible from that location in space. For microwave devices, the scanning

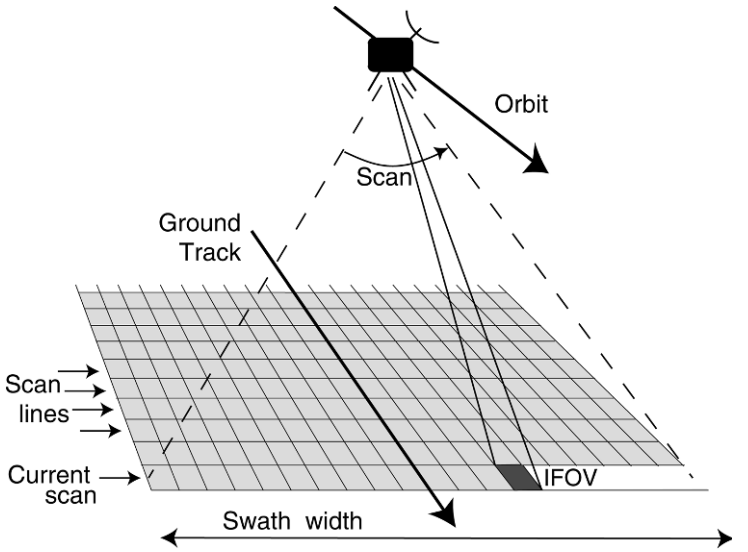


Figure 2.3. Swath-filling geometry of a rectangular, line-scanning sensor.

may be achieved by electronic (beam steering) or radar signal-processing methods without the need for mechanically moving reflectors. In this case it may be possible to reduce the geometric distortion inherent in mechanical scanning. A fuller discussion of scanning and imaging methods is given in section 4.1 of *MTOFS* (Robinson, 2004).

It is worth emphasizing the fact that a remote sensor integrates the incoming radiation over the IFOV and so the estimates it makes of ocean properties correspond to averages over the measurement footprint. In a well-designed scanning system in which the sea surface is covered by contiguous but not overlapping footprints, the resulting set of measurements are directly comparable with the way a two-dimensional model describes the sea, representing ocean variables as averages within each cell of a rectangular grid. For many applications of satellite data this provides a distinct advantage compared with conventional *in situ* instruments that make single-point measurements in the sea. To compare such *in situ* data with models requires measurements that are representative of the whole cell, difficult to achieve if subcell-scale variability is large unless many different measurements can be spatially averaged within the cell. Remote-sensing observation avoids this point-sampling problem, although it is encountered in a different way during the procedure of validating satellite data by comparing them with *in situ* measurements.

2.2.2 Satellite orbits and how they constrain remote sensing

Earth-orbiting satellites are constrained by forces due to gravitation and inertia. Based on Newtonian dynamics, the period, T, for a satellite to travel once round

a circular orbit at distance r above the center of the Earth is:

$$T = 2\pi\sqrt{\frac{r^3}{GM}}, \tag{2.1}$$

where G is the constant of gravitation; M is the mass of the Earth; and $GM = 3.98603 \times 10^{14}\,\mathrm{m^3\,s^{-2}}$. In terms of the satellite height h above the ground and the Earth radius R (about 6,378 km) $r = R + h$ and so

$$T = 2\pi\sqrt{\frac{(R+h)^3}{GM}}. \tag{2.2}$$

There are just two basic types of orbit useful for ocean remote sensing, geostationary and near-polar, as illustrated in Figure 2.4. The *geostationary orbit*, at a height of about 35,785 km, has a period of one sidereal day (~23.93 h) which is the time taken for the Earth to rotate through 360°. Placed over the Equator, the satellite flies west–east at the same rate as the Earth's rotation, so it always remains fixed at the same place in the sky relative to objects on the ground, allowing it to view the ground at any sampling frequency. Being fixed it can see only that part of the world within its horizon and cannot usefully view much beyond about 7,000 km in any direction measured from the satellite nadir point on the Equator, at the longitude of the satellite.

In a *near-polar orbit* the satellite flies at a much lower altitude, typically between about 700 km and 1,350 km, for which the orbital period is about 100 min (Equation 2.2). It thus completes between 14 and 15 orbits a day, during which the Earth rotates once, so the satellite marks out a ground track crossing about 14 times northeast–southwest (descending tracks) and the same number of southeast–northwest ascending tracks. The tracks are distributed evenly around the globe, with successive orbits following a track about 24° of longitude to the west of the previous orbit as shown in Figure 2.5. If the satellite in orbit returned to its starting point exactly after one day then it would go on repeating the same 14 or 15 orbit

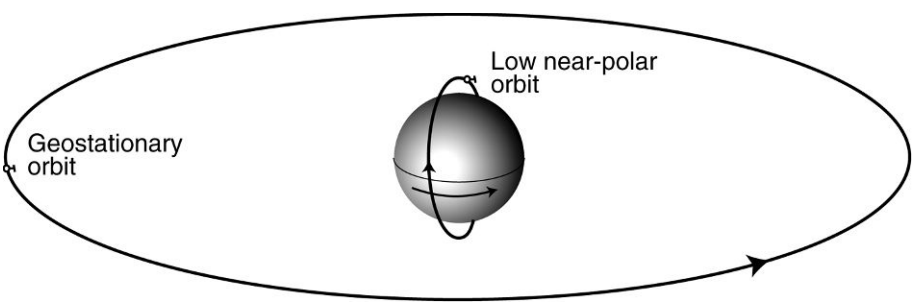


Figure 2.4. The two types of orbit used for Earth-observing satellites, drawn approximately to scale. The geostationary orbit is about 36,000 km above the Earth. The near-polar orbit is typically between 700 km and 1,000 km above the Earth surface.

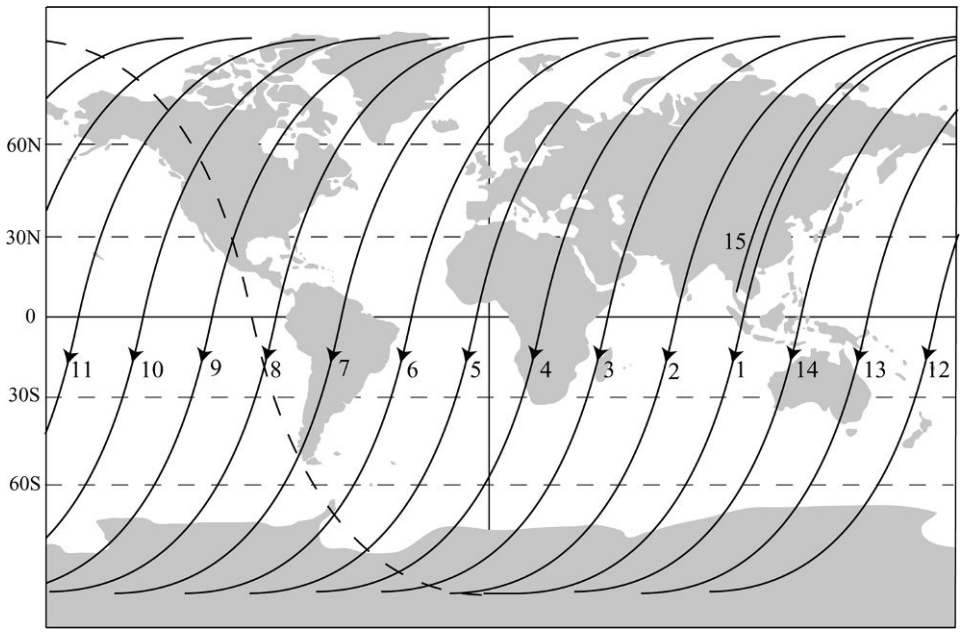


Figure 2.5. Ground track of a typical near-polar, low-Earth orbit, showing all the descending passes for one day and one ascending pass (dashed).

tracks and never visit the spaces between them. Instead the orbit is normally planned so that it takes a longer time, typically between 3 days and 35 days known as the orbit repeat period, before it starts to cover its earlier track exactly. The longer the orbit repeat period the greater the number of different orbit tracks over the Earth surface that are completed within the cycle, and so the smaller the spaces between the tracks.

Most low, near-polar orbits used for Earth observation satellites are arranged to be *Sun-synchronous*. By choosing an inclination that is slightly greater than 90° (i.e., their path does not quite reach the poles) the orbit plane can be constrained to precess at a rate of once per year relative to the stars. This locks the overpasses to the position of the Sun and means that every orbit always crosses the Equator at the same local solar time. For most ocean-observing sensors this is very convenient, since it ensures that the longitudinal position of the Sun does not change from one sample to the next, even though the solar latitude inevitably changes with the annual cycle. However, for altimetry a Sun-synchronous orbit is to be avoided since it aliases the solar semidiurnal tidal constituent, because the solar tidal phase will be exactly the same every time the satellite revisits the same location on the sea surface.

More information about orbits can be found in section 3.2 of *MTOFS*. Section 11.6.3 of *MTOFS* explains more about tidal aliasing in altimetry.

2.2.3 The space-time sampling capabilities of satellite sensors

While a sensor on a geostationary platform is unable to see beyond its restricted horizon it does stay in the same place all the time and is therefore potentially capable of high-frequency sampling. However, it takes time to scan in spatial detail the full Earth disk field of view and its great height above the ground also makes it difficult to achieve fine spatial resolution. Geostationary sensors typically offer a revisit interval of less than 30 min and spatial resolution of 3 km to 5 km. This gives them the highest frequency time sampling of all satellite sensors, but relatively poor spatial resolution.

In contrast, a scanning sensor on a polar platform can potentially cover the whole Earth in a single day, as long as the swath is at least about 2,700 km wide, which is the distance at the Equator between the ground track of successive polar orbits (see Figure 2.6a). In this case each point on the Earth surface will be viewed at least once from a descending track and once from an ascending track. For a Sun-synchronous orbit, if the ascending track is in local daytime the descending track will be at night, or *vice versa*. An even wider scan permits more samples per day as swaths from successive orbits overlap at the Equator, while at higher latitudes overlapping occurs for much narrower swaths. Nonetheless, except in polar regions, the regular sampling interval for a single polar orbiter is never less than several hours.

For much narrower swaths as illustrated in Figure 2.6b, normally associated with fine-resolution imaging sensors, the time between successive views of the same location depends on the orbit repeat period. If this is just a few days then the sensor revisit interval will be the same as the repeat period. However, for too short a repeat period the spacing between the complete set of ground tracks will still be wider than

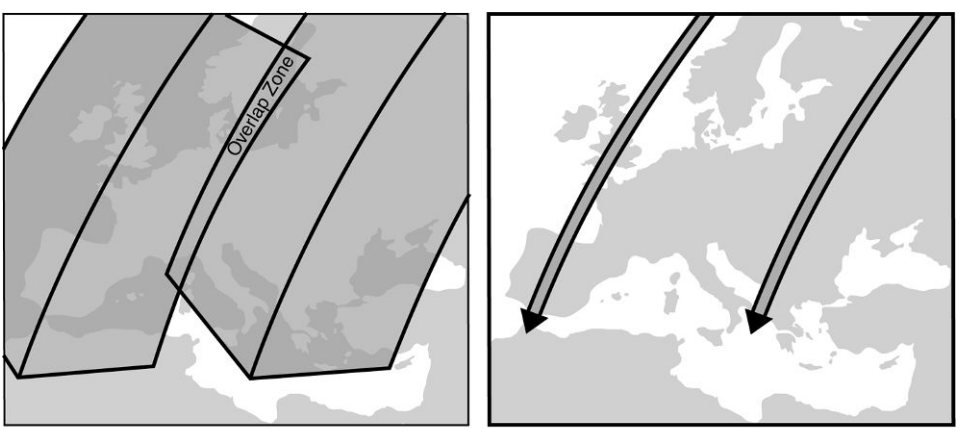


Figure 2.6. A single day's coverage over Europe by (a) a wide (>2,000 km) swath sensor and (b) a narrow (<200 km) swath sensor. In each case the two tracks represent the typical spacing between successive orbits of a polar orbiter at an altitude of about 1,000 km.

the narrow swath and so the sensor will miss many parts of the Earth surface altogether. Global coverage by a sensor whose swath is about 200 km would take about 15 days to accomplish, and so an orbit repeat cycle of at least 15 days is required for this to be achievable. Sensors with intermediate width swaths (e.g., 500–1,000 km) will still take several days to obtain total coverage of the equatorial latitudes, but the revisit interval is likely to be shorter than the orbit repeat period, the more so at higher latitudes.

For nonscanning instruments such as altimeters that sample only along the ground track, the longer the orbit repeat period is, the greater the spatial coverage and the finer the sampling grid that can be achieved. However, it takes an extended period of time to build up the resulting map of data. Whereas for scanning sensors the spatial-sampling grid depends on the scanner design, and is usually matched to the sensor spatial resolution, for a nonscanning sensor it is the orbit pattern that dictates the spatial-sampling grid.

For scanning and nonscanning sensors alike, there is evidently a well-defined tradeoff between spatial and temporal-sampling capability, which is discussed in more detail by chapter 4 of Robinson (2004). It is important to appreciate these fundamental constraints when designing an ocean-observing system for operational purposes. For example, the only way to ensure that even a wide-swath sensor can sample every 6 hours from a near-polar orbit is to fly sensors on two satellites. Ideally a combination of spatial and temporal resolution should be selected in order that important phenomena can be adequately sampled. For example, if mesoscale eddies are to be monitored then the spacing between orbit tracks should not be wider than their variability lengthscale, nor should the repeat cycle be longer than the characteristic lifetime of an eddy. Otherwise some eddies may be missed altogether.

Every individual remote-sensing instrument has its particular space-time sampling capabilities, depending on both the sensor itself and the platform on which it is placed. These can be expressed in a graphical form, as in Figure 2.7 which represents the sampling characteristics as a defined rectangular area in a two-dimensional space-time field, for the generic sensor types discussed in the foregoing paragraphs. The vertical axis represents a logarithmic lengthscale and the horizontal axis a logarithmic timescale. The lower boundary of the region allocated to a particular sensor represents the smallest spatial scale that can be detected by the sensor (i.e., its spatial resolution). Similarly the left-hand boundary represents the shortest time interval over which variations in the ocean can be detected (i.e., the temporal-sampling resolution). The bottom left-hand corner therefore represents the best spatiotemporal resolution that is possible using that sensor. Note that a clear field of view with no obstructions is assumed. In the case of a sensor type that can view only under clear-sky conditions, cloud cover would degrade the temporal resolution and have the effect of shifting the left-hand boundary farther to the right.

The top boundary for each sensor's sampling capability box corresponds to the largest extent of spatial coverage that can be obtained for a near-instantaneous view. This is the size of individual images collected at a particular time, and is generally defined by the swath width, although it could be larger in the along-track direction.

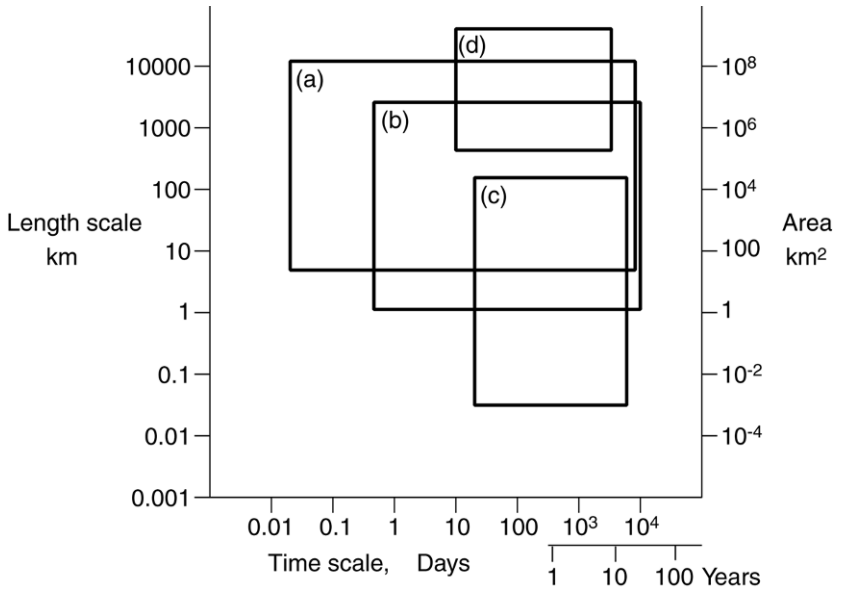


Figure 2.7. Diagram representing the space-time sampling characteristics of four types of sensor. (a) Scanning radiometer on a geostationary platform. (b) Wide-swath, medium-resolution (1 km) scanning sensor on a polar-orbiting platform. (c) Narrow-swath, fine-resolution (20 m) scanning sensor on a polar-orbiting platform. (d) Nadir-sampling nonscanning sensor (e.g., altimeter) on a polar-orbiting platform.

Of course this does not represent the full coverage of the sensor which, for sensors on polar platforms, is global. The right-hand boundary represents the timespan of available data, and depends on the lifetime over which useful data have been collected. In the examples illustrated a continuous series of at least 10 years is assumed.

The region enclosed within the boundaries represents the space-time sampling space for that sensor. The height of the box indicates the range of lengthscales that can be resolved, and the width the range of timescales. Diagrams like this are useful not only for comparing different sensors, but also for relating the sensor sampling capability to the requirements of a particular application; specifically in matching it to the space-time variability scales of the ocean phenomenon to be observed, as discussed further in section 4.5 of *MTOFS*.

2.3 GENERIC DATA-PROCESSING TASKS

It is important that those who make use of satellite-derived data products should be aware of the calibrations, corrections, analyses, and resampling that may have been applied to the products before they received them, since these processes have impacts that are relevant for their oceanographic interpretation and application. This section therefore provides a short overview of these tasks, which correspond to the informa-

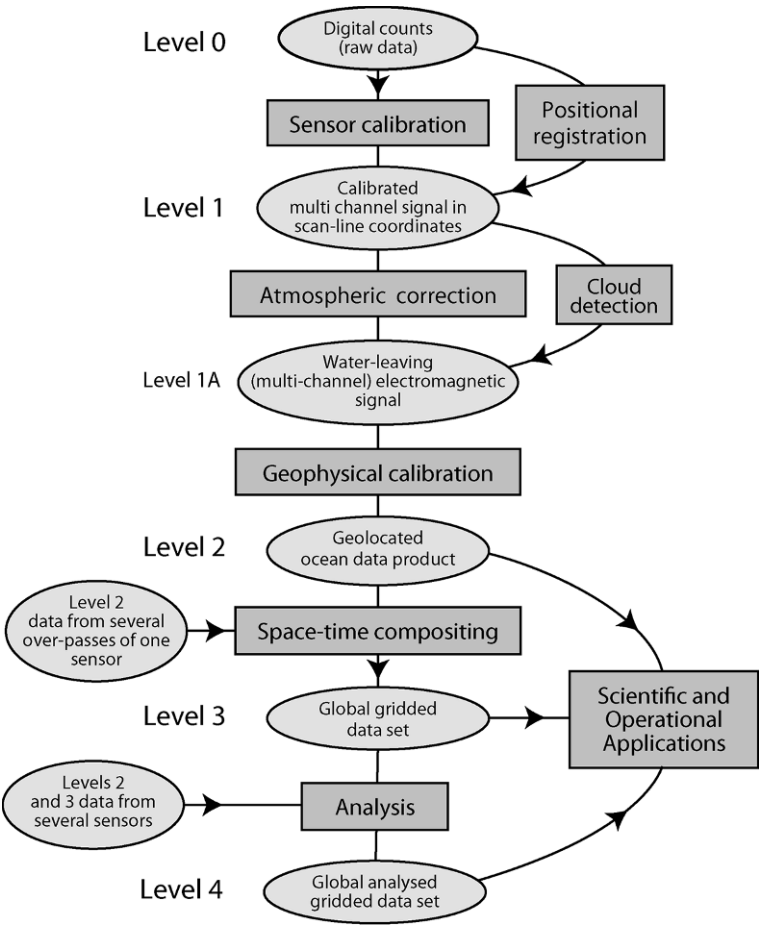


Figure 2.8. Outline of data-processing tasks to convert raw satellite data into ocean products.

tion "unpacking" that is necessary (see Figure 2.1) if the raw data acquired from a satellite are to be turned into useful quantitative information about an ocean variable, property, or parameter. Figure 2.8 summarizes the sequence of tasks and indicates the different "levels" of data products that correspond to each stage of data processing. The data product levels are defined in Table 2.1. Some tasks specific to particular satellite oceanography methods will also be described in Section 2.4. A more complete discussion of generic processing tasks can be found in *MTOFS* (sections 3.4.2 and 5.2).

2.3.1 Sensor calibration

The sensor calibration stage of data processing should convert the raw data received from the sensor into an estimate of the electromagnetic property which the sensor is intended to detect, such as the radiance entering the sensor in a given waveband. It

Table 2.1. Levels of satellite data products from different stages of processing.

Level	*Description of product*
0	Raw data received from the satellite, in standard binary form.
1	Image data in sensor co-ordinates. Individual calibrated channels, of measurements made at the satellite.
1.5 (or 1a)	In special cases, level 1 data with atmospheric correction applied.
2	Atmospherically corrected and calibrated image dataset of water-leaving radiance or derived oceanic variable. Geolocated, but normally presented in image co-ordinates.
3	Composite images of derived ocean variable resampled onto standard map base. Averaged in space and time from several overpasses of level 2 data. Derived from a single sensor. May contain gaps.
4	Image representing an ocean variable averaged within each cell of a space-time grid, for which gaps at level 3 have been filled by data analysis including interpolation. The analysis may merge several level 2 and/or level 3 datasets from a number of sensors and may also use *in situ* observations or model output.

must account for all factors affecting information flow from the environmental signal entering the sensor to the receipt of the digital data signal at the ground. The first requirement is for a sensor calibration model that will determine the radiance, phase, or other property, of radiation entering the satellite sensor, given the value transmitted by the satellite and received at the ground station. This model is needed to invert two processes: the conversion of received radiation into an electrical response (typically a voltage amplitude or frequency) and the conversion of that response into a digital value. The accuracy of the calibration model must be defined in order to characterize the errors introduced in the transducer and the digitizer. It can be assumed that the transmission of the digital signal is error-free, if the signal is received at all.

For good sensor calibration to be achieved, all the sensor components must be carefully calibrated and characterized before launch. However, for many sensor types, this by itself is not enough to ensure confident in-flight calibration. Radiometers, for example, normally require the use of calibration radiant sources to allow calibration to be regularly updated. Since the sources as well as the sensors may degrade, elaborate checking systems use other calibration targets, such as the Moon for visible wavelength light. The slow degradation of a sensor calibration over its lifetime may only be fully defined a considerable time after the data were first acquired. Consequently the historic archive of data from a particular sensor may be reprocessed and the derived products reissued, up to several years after they were first acquired. Users of satellite data need to be aware of this possibility.

If the data are made publicly available after sensor calibration has been applied then they are referred to as a level 1 dataset. Not all agencies make data of this type available, since it does not serve a very practical purpose for data users, although it is useful for scientific analysis into the underlying remote-sensing processes.

2.3.2 Atmospheric correction

A serious disadvantage for oceanographers of remote sensing from Earth observation satellites is that we must look through another medium—the atmosphere—to see the ocean. The atmosphere is opaque to electromagnetic radiation at many wavelengths, and there are only certain wavelength windows through which radiation may be fully or partially transmitted. Atmospheric gas molecules themselves may absorb or scatter radiation, and in addition water vapor, aerosols, and suspended particles of dust will do the same. If water droplets are present in the form of clouds, they may completely change the transmission properties of the atmosphere.

A major task for the analysis of satellite data is to take into account the effect of the atmosphere. At its worst this requires the detection of phenomena which render the data unusable (such as cloud for visible and infrared sensors or heavy rain for microwave radiometers). Otherwise it involves estimation of the effect that the atmosphere has had on measured electromagnetic radiation, between leaving the sea and reaching the sensor.

Cloud detection

Sensors using the visible and infrared parts of the spectrum cannot view the ocean through cloud. When handling data from such sensors, the next processing task after applying sensor calibration is to detect which pixels are obscured by cloud and which are clear. The purpose of cloud detection operations is to flag pixels as cloudy or as cloud-free. In some cases the flagging can be more sophisticated, indicating probabilities of cloud corruption, or indicating which of several different cloud detection tests has been positive. A comparable process is also adopted for analyzing microwave radiometer data. Although clouds are transparent to microwaves and do not present a problem, cells of very heavy precipitation in the atmosphere are opaque to microwaves and need to be detected.

It must be appreciated that clouds can never be "removed" from image datasets to leave a cloud-free field. To do so would imply that we knew what the ocean conditions were like beneath the cloud, but that is precisely what the cloud prevents us from seeing! To be able to present cloud-free datasets derived from visible and infrared sensors requires either (a) that data from several overpasses have been combined in a composite, or (b) that gaps in the data where clouds obscured the sea have been filled by some type of analysis such as optimal interpolation based on cloud-free pixels, or using the most recent previous cloud-free data, or substituting climatological values. Sometimes a combination of (a) and (b) is used. Whether such gap-filling is acceptable depends on the application. The user should always adopt a

critical approach to such data to ensure that erroneous conclusions are not drawn from aspects of the data which are themselves purely artifacts, and not real data at all.

One of the most challenging aspects of visible and infrared remote sensing is detecting those clouds which only partly cover a pixel and may be very hard to distinguish from clear-sky conditions. If incorrectly accepted as clear, they will yield anomalous values of the ocean variable being measured, and can corrupt a dataset. If there is a danger of such false values introducing a bias (e.g., to a climate dataset), then it is best to err on the side of caution. Although it may be disappointing to throw away some good data because they are suspected of being cloud-contaminated, this may be worth it in order to maintain confidence in the quality of the overall dataset. On the other hand, for operational applications long-term biases may not matter so much and a more relaxed approach to cloud contamination can be adopted.

Atmospheric correction strategies

Once any pixels obscured by cloud have been detected and flagged, attention must be turned to the remaining "clear-sky" pixels. In general these still contain contributions from the atmosphere, which corrupt the water-leaving signal that we must estimate from the top-of-atmosphere signal, which is what the satellite sensor actually measures. The atmosphere scatters light, especially in the visible and near-infrared wavebands. Some of the radiant energy from the sea to the sensor, which represents the true "signal", is deflected away from the sensor by atmospheric scattering. In addition solar radiation is scattered into the field of view of the sensor to add "noise" to what is observed. In the thermal infrared and the microwave parts of the spectrum, used to measure the temperature of surfaces emitting the radiation, absorption and emission by the atmosphere is the major problem.

The detailed efforts made to eliminate atmospheric effects are specific to individual sensor types. In general it must be noted that for most contexts it is not possible to base an atmospheric correction on *a priori* knowledge of the atmosphere itself. Significant atmospheric interference with the remote-sensing signal comes from atmospheric constituents like water vapor or aerosols whose distribution is not known to sufficient accuracy. Thus we must rely on the remote-sensing measurement itself to contain enough information about atmospheric variability to allow a correction to be made. This is where the use of multiple spectral channels provides an answer. By sampling in several different wavebands, chosen because of their different responses to both the ocean-leaving signal and the effect of the atmosphere, strategies have been developed which allow atmospheric variability to be accommodated under most conditions.

Datasets that have had atmospheric correction applied used to be referred to sometimes as level 1.5 or 1A data products. However, in the case of recent Earth-observing missions, water-leaving radiance datasets in individual visible wavebands, atmospherically corrected but without any further geophysical algorithm being applied, are referred to as level 2 products.

2.3.3 Positional registration

By positional registration we mean the identification in geographical co-ordinates of the place to which a remotely sensed measurement refers (i.e., the location on the ground or sea surface to which the sensor was pointing when it recorded the measurement). Sometimes this is referred to as image navigation, or geolocation. The adjustment of an image to bring it into conformation with a map base is sometimes called geometric correction or rectification of the image.

Fundamentally, the problem is one of knowing where the satellite was when a measurement was made, and in which direction the sensor was pointing. The accuracy with which this information is required depends on the type of sensor and the application. For single measurements with a downward-pointing sensor providing an area-integrated measurement over a $50\,\text{km} \times 50\,\text{km}$ square, precision in locating the nadir point need not be better than a few kilometers. In such cases it is sufficient to know the expected satellite orbit and nominal pointing direction.

For other nonimaging sensors, a more precise positional registration is required, involving ground station tracking of the satellite and accurate onboard monitoring and transmission of the satellite vehicle's pitch, roll, and yaw. For some remote-sensing systems this information from diverse sources is brought together and published with the sensor measurement. The precise track of the satellite is known as its ephemeris. Given the tremendous improvement in satellite navigation in recent years using the U.S. global positioning system (GPS) or its Russian counterpart (GLONASS), it is now possible in principle to derive positional registration based on navigational models for all image data. A greater uncertainty than the satellite location is knowing the direction in which the sensor is pointing.

It is worth noting here that, although positional registration is often essential information for both atmospheric correction and geophysical calibration of the data, it is never advisable to use geolocation information to resample a dataset until after these other operations have been performed. Correction and calibration algorithms often use multispectral information in which the ratios or differences between different data channels is crucial. Resampling an image onto a new geographical base requires resampling of the pixels, during which band ratio information may be distorted. If resampling is unavoidable, the use of "nearest neighbor" substitution should be chosen rather than a more elaborate process which averages values from different pixels and exacerbates the problem.

2.3.4 Geophysical product derivation

When the raw data have been calibrated and atmospheric correction has been applied, the next stage in the processing chain is to manipulate the data in order to derive estimates of a particular ocean variable. Because the resulting product is fundamentally different from the measurement originally made by the sensor, it should be considered as a measure or estimate of an ocean variable rather than a measure of electromagnetic radiation. This process is often termed *geophysical calibration*. In practical terms it is normally performed simply by applying a set of

arithmetic or algebraic manipulation rules, the algorithm, to the calibrated and atmospherically corrected electromagnetic data observed by the sensor.

Developing calibration algorithms

The geophysical calibration of remote sensors generally requires a combination of empirical and theoretical methods. It is customary to start the construction of a calibration model by basing it on a theoretical analysis of the physical processes by which an oceanographic parameter influences the electromagnetic radiation observed by the sensor. In the case of emitted thermal infrared radiation, for example, the theoretical model of radiation emission from black bodies goes a long way towards providing a complete calibration algorithm. Visible wavelength remote sensing of suspended particles in water is less amenable to theoretical analysis. While quite a lot is known about the optical processes in the ocean and atmosphere they are so complex as to make it impossible to produce from first principles a calibration of radiance data in terms of suspended sediment load. In the case of active radar sensors the physical processes which control radar backscatter from rough surfaces cannot be modeled by readily invertible analytical functions and so a more empirical approach is taken.

While it is desirable, if possible, to build calibration algorithms on a sound, physically based, theoretical model, eventually nearly all geophysical algorithms tend to require a set of calibration data in which observations from the satellite are matched to *in situ* measurements of the ocean variable which is to be derived, obtained at the same time and location as the satellite overpass. If the geophysical algorithm is to be widely applicable, it needs to be based on *in situ* matchup data that are representative of a wide range of locations, and conditions. Otherwise the algorithm will be biased towards conditions represented by the matchup data, where it performs well, while its performance may be much poorer in other conditions.

Once a geophysical calibration has been performed on satellite data from a single overpass, it is typically distributed as a level 2 dataset in its original "native" grid consisting of the original pixels arranged in rows and columns matching the sensor scan lines and the along-track direction of the satellite.

Ocean product validation

Whether the geophysical algorithm is based on purely theoretical foundations or on empirical matchup datasets, its performance cannot be critically assessed unless there is also in place a robust system of validation. This should generate a set of error statistics indicating the bias and standard deviation of the data products produced from the satellite sensor when compared against a set of reliable independent measurements of the same ocean property.

Validation, at its simplest, consists of comparing the value of an ocean variable determined from the satellite data with an *in situ* measurement of the same variable, coincident in space and time. In order to be meaningful this must be performed a large number of times. There is a need to test the validity of the algorithm using data spanning the whole range of variable values for which it is specified for use. This

means that data from a single cruise in a limited region, however many coincident points are obtained, will not be adequate to validate a global algorithm. Hence the advantage of an array of drifting buoys, or the use of long-transect ships of opportunity. There is also a need to test the algorithm under a variety of those environmental conditions which have an influence on it or the earlier parts of the data-processing chain. For example, if the geophysical products are sensitive to the atmospheric correction, validation must be performed in a variety of representative atmospheric conditions before full confidence in its performance can be justified. Validation must also be continued over the lifetime of the sensor.

Finally it should be obvious that the validation dataset must not be the same as those used to develop an empirical calibration, since in that case it would not be truly independent. Those who use data provided by others in developing new scientific or operational applications of satellite remote sensing must ensure for themselves that the data products have been validated, at least to the degree of reliability required by the application.

2.3.5 Image resampling onto map projections

Another stage in the chain of formal processing required to produce useful maps of ocean variables derived from satellite data is to present them in a geographically intelligible form. Of all the processes considered so far, this is perhaps the least significant because it does not essentially affect the accuracy or applicability of the data. At the image navigation stage mentioned in Section 2.3.3 a model should have been established for the data which enables the location of each pixel in the original satellite image to be defined in terms of latitude and longitude. This information by itself is adequate for many applications of the data, including matching to *in situ* observations.

However, one of the great strengths of remotely sensed data is the capacity to provide detailed spatial views, and to communicate information by allowing a user to view an image of the ocean. Images are painted on a screen or drawn by a printer in rectangular form corresponding to the rows and columns of the matrix in which the parameter values are notionally stored in digital form by the computer. The orientation of the rows and columns is defined by the viewing geometry of the sensor, and does not in general bear any relation to normal geographical conventions for drawing maps. Moreover if the image is drawn in "satellite co-ordinates" there may well be distortion of shapes compared with their form on the ground. For these reasons resampling onto a new grid conforming to a standard map projection is often performed.

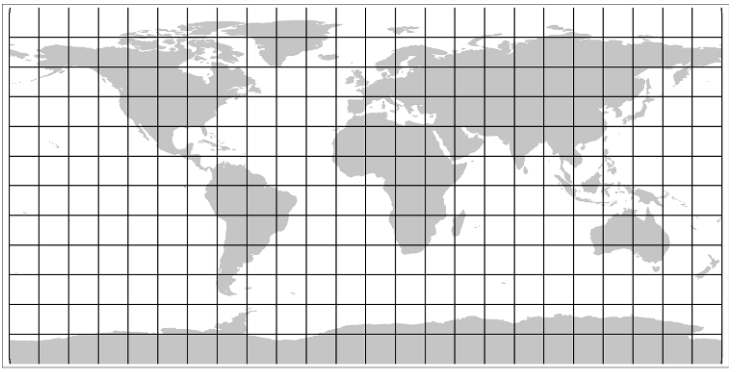
The choice of map projection can itself be an issue of some importance for producers and users of satellite data. The purpose of a map projection is to present data, whose positions are defined on a spherical surface, as if they were distributed on a flat plane such as a computer screen or printed page. For a local region spanning no more than a few hundred kilometers, geometric distortion is minor and there are few problems. For larger dimensions of thousands of kilometers, approaching the radius of the Earth, the distortions are much greater. Resampling

data onto the map projection may (a) require data from several sampled points to be located in just one new pixel on the map projection, while (b) other pixels on the new map may contain no satellite data samples at all. In this case careful thought must be given to the rules used for transforming satellite data, whose geolocated position in latitude and longitude places it on a spherical surface, to the map on a plane surface. It is reasonable to suppose that we would preserve the most complete information from the satellite data by applying weighted averages in case (a) and using interpolation from neighboring pixels in case (b). However, as noted in Section 2.3.3, if the differences or ratios of the values in different data channels (e.g., different wavelengths of a multispectral dataset) are to be obtained and exploited from the resampled version of the data, it is essential that the much simpler "nearest neighbor substitution" is used to create the map because only that approach will preserve the original channel ratios as measured by the sensor.

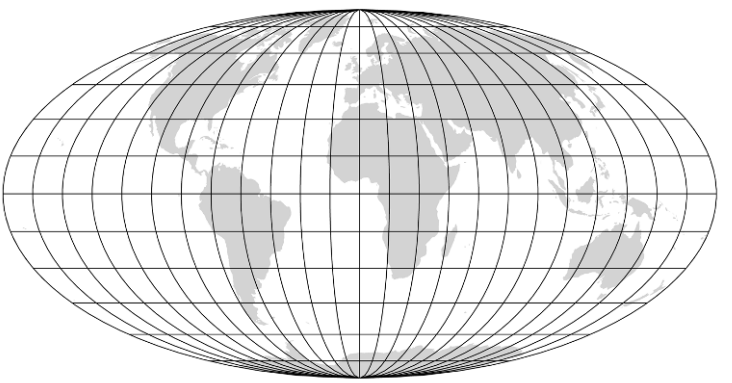
A further difficulty of interpretation arises with the presentation of global datasets on a single flat map. Put at its simplest the problem is this: a spherical surface has no edges but a flat map is always bounded by an edge. For ocean data at low to mid-latitudes the solution is to use cylindrical projection (see Figure 2.9a) in which lines of latitude are horizontal and lines of longitude are always vertical. The main issue then is deciding at which longitude to cut what would otherwise be a continuous strip of data. In the illustration, longitude 180° has been chosen but this splits the Pacific Ocean into two parts. When interpreting such maps it is important not to forget that in reality the data are continuous across the left and right boundaries. However, at high latitudes such maps are inappropriate because east–west distances are greatly exaggerated. A Mollweide type of projection (Figure 2.9b), in which lines of longitude converge to the poles, preserves relative surface areas at different latitudes but at the expense of gross shape distortion. However, like cylindrical projection, it completely obscures any sense of the neighbor relations between high-latitude locations at different longitudes. For this a polar projection (Figure 2.9c) is needed, in order to do justice to the capacity of wide-swath sensors in near-polar orbits to map the polar regions coherently.

For global maps of many ocean properties, the polar regions are of little interest and so cylindrical maps are satisfactory. However, where properties over polar regions are also important (e.g., where sea ice and sea surface temperature are being mapped together), it is necessary to present the data on three separate maps: a global cylindrical map projection and two polar projections, one for each pole.

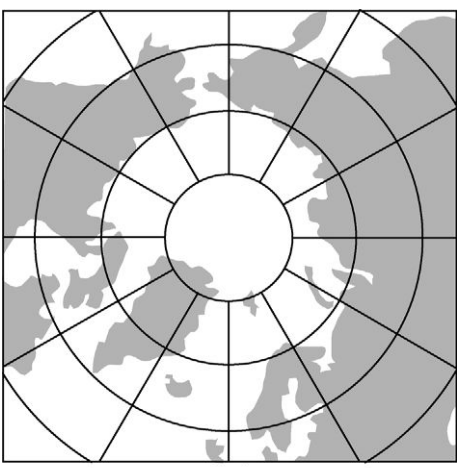
In future it is likely that the problems of producing datasets in specific map projections, leading to distortion and problems of interpretation, will be overcome by present trends in information technology. There is increasing availability of software that allows any set of data for which the pixel locations are defined in latitude and longitude to be presented in a wide variety of projections. Interactive software offers a three-dimensional view of the Earth and allows users to navigate their vantage point and viewing direction so as to explore the data fully. As this becomes widely available, there will be less need to prepare data in special map projections. In the long term, scientific ocean datasets retrieved from satellites can



(a)



(b)



(c)

Figure 2.9. Examples of map projection types. (a) Cylindrical projection, satisfactory at low to mid-latitudes, amplifies areas at high latitudes. (b) Mollweide projection, preserves equal area at the expense of shape distortion. (c) Polar projection, avoids the loss of a coherent polar view inherent in (a) and (b).

be archived with a spatial density that matches the sampling capacity of the sensors from which they are derived, while viewing software will automatically perform the resampling necessary to display the data to the user at their chosen vantage point. Such an approach will also be able to offer the user the capacity to animate time sequences of mapped data in order to observe the evolution of dynamical processes.

2.3.6 Composite image maps

An important process which may be performed to produce a satellite ocean data product is that of combining multiple sets of data from the same sensor but different overpass times and different locations into a single map of data. The result is a composite dataset, generally referred to as a level 3 data product which differs in a number of ways from a level 2 dataset derived from a single overpass. Its typical characteristics are

- It is almost always on a regular geographical grid, resampled onto pixels which are normally oriented with north upwards and incremented in intervals measured in latitude and longitude.
- The pixel size is normally larger than the native pixel in the level 2 image.
- A level 3 composite image is often global in coverage, or extends over a large region such as a single ocean. Its extent is almost always larger than can be encompassed by a single swath of the sensor that supplies its data.
- It contains data from several different overpasses, ascending and descending.
- It is based on data from overpasses within a given timespan.
- It should ideally contain statistical measures of the variability of the level 2 data sources from which it was constructed, such as the number and standard deviation of independent level 2 samples whose average becomes the composite value.

The advantage of a composite image is that it extends the geographical scope of an image dataset beyond the confines of a synoptic image derived from a single overpass, and it is able to fill up over time the gaps that are left either by the inadequate coverage pattern of swaths on a single day, or by the localized dropout of data because of cloud cover or other problems.

However, the construction of a composite image requires the application of rules which can impact the character of the resulting dataset. The main question to address is what to do with multiple entries into a single cell of the composite image. Given the downscaling that is normally involved, several pixels in one of the contributing level 2 images will contribute to a single cell of the composite. Normally the values from each of these should be averaged. Over the integration time for the composite several further level 2 datasets may contribute. These also will be added to the averaging process. However, in some cases only the latest value may be selected, or only those values which, in processing to apply the geophysical calibration, generated the highest confidence value. There is evidently the possibility that subtle changes of the rules for constructing the composite could significantly change the result.

Other design choices for a level 3 composite relate to the size of the pixels and the integration time. In some situations different composites are created from ascending and descending overpasses, which would be appropriate when, for a sensor on a Sun-synchronous satellite, these correspond to different local times of day such as day and night. The balance of advantage between different options depends on the use to be made of the composite data. It should always be remembered that a composite does not contain all the information present in the original level 2 datasets, although careful construction of a composite should enhance the applicability of the data. Composites are particularly useful for observing large-scale ocean phenomena that may be missed by the limited geographical coverage of individual level 2 datasets, and this will be examined in more detail in Section 6.2.1. The user should note carefully what rules have been applied in its production, and consider what implications these may have for interpreting and applying the data. In principle a level 3 composite should contain only data from contributing level 2 images, with the likelihood that some gaps will remain in the end product because of persistent cloud or other sampling limitations. Nonetheless users of composite datasets should check whether they do, in fact, contain inputs from more than the supposed contributing data. For example, persistent data gaps may have been filled by interpolation, or by substitution of the data from the previous integration period, or even from a climatological-averaged dataset.

Strictly, as soon as data from other sources are introduced, or spatial gaps are filled by interpolating adjacent pixels or even simply by smoothing the data field, the outcome should be referred to as a level 4 dataset. Whereas a level 3 product should be based only on what the satellite has measured, a level 4 product is created with the intention of providing the user with a complete data field that contains no gaps. It therefore represents the result of an analysis to estimate what the spatial distribution of a particular ocean variable was at a particular period of time, and so may be referred to as an "analysed field". Although it may be primarily derived from satellite data, we should expect it to contain more than just measurements from a single satellite. It may contain information from different types of satellite sensors, and be blended with *in situ* observations.

The processes used to fill gaps and smooth them will have constrained them to conform to certain expectations of what they should be like. Therefore, although for most applications it is far more helpful for users to be given a seamless and smooth distribution of an ocean variable rather than a sparse patchwork of what was actually observed, its interpretation must always bear in mind the assumptions that went into its construction. For example, if the production of a level 4 analysis of an ocean variable implicitly smooths observed data spatially or temporally with a low-pass filter, it would be inappropriate to use it for investigating high-frequency variability in that ocean property. The most useful level 4 products should be accompanied by confidence values for each pixel that will alert the user to issues where misinterpretation might occur. For example, it is to be expected that the confidence value will be lower for those pixels where no observations have been available to influence the analysis. This will be discussed further in Section 14.4.2 which explains

how analyzed sea surface temperature data products are prepared for operational applications.

2.4 SENSOR TYPES FOR OBSERVING THE OCEAN

2.4.1 Using the electromagnetic spectrum

All sensors employed on ocean-observing satellites use electromagnetic (EM) radiation to view the sea. The capability of particular sensors to measure certain properties of the ocean and how well they can view through the atmosphere or penetrate clouds depends critically on which part of the EM spectrum they use (for a fuller discussion see *MTOFS*, chapter 2). Figure 2.10 shows regions of the electromagnetic spectrum that are of relevance to remote sensing and in particular the bands occupied by the four broad classes of satellite sensors used for viewing the ocean. The types of sensor found in these broad classes, the primary measurement they make, and the parameters that can be derived from them are summarized in Figure 2.11.

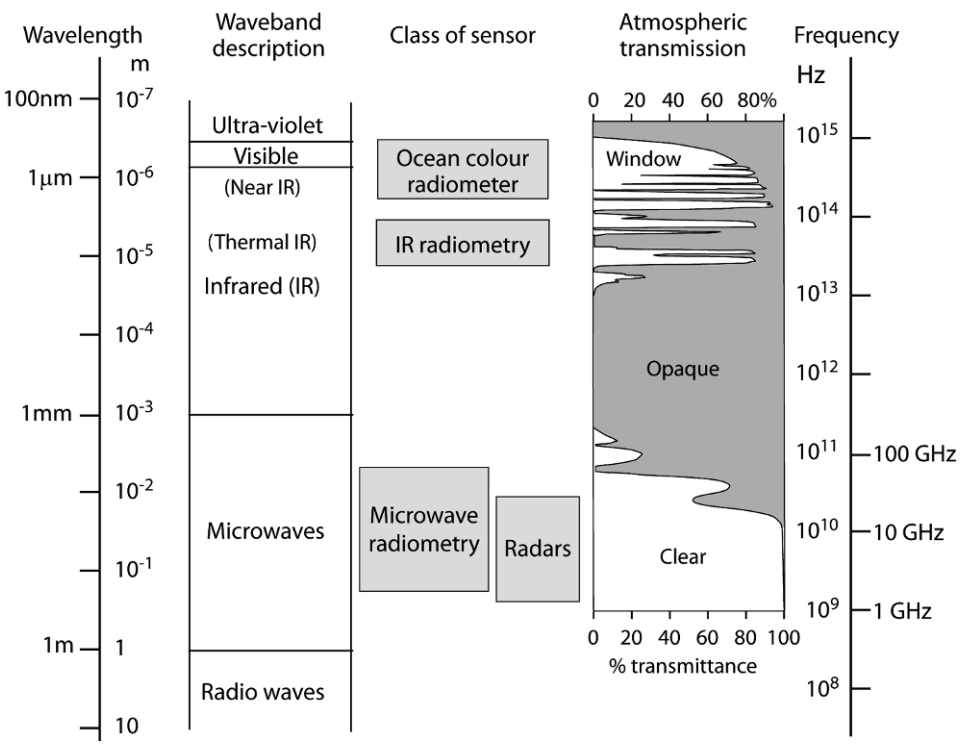


Figure 2.10. The electromagnetic spectrum, showing the regions exploited by typical remote-sensing instruments.

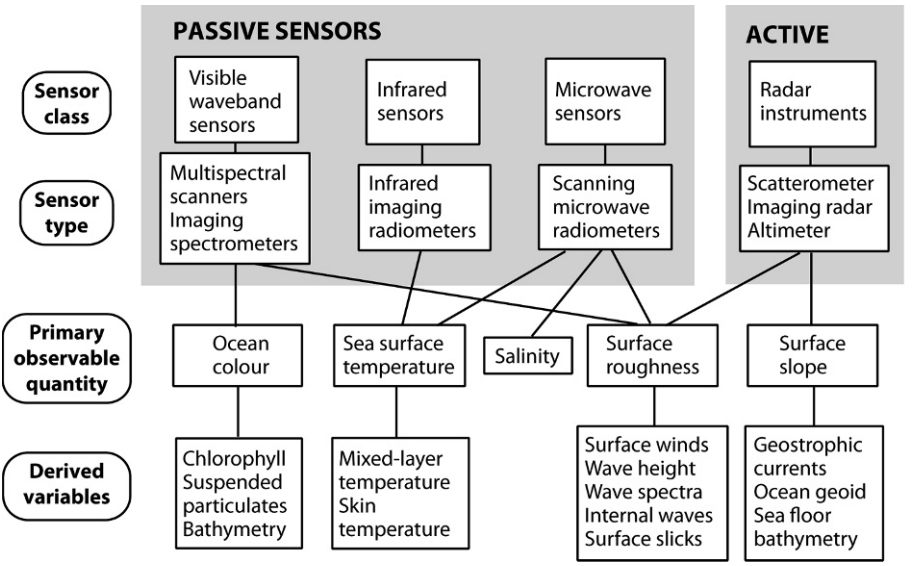


Figure 2.11. Schematic illustrating the different remote-sensing methods and classes of sensors used in satellite oceanography, along with their applications (from Robinson, 2004).

Figure 2.10 also shows how transmittance of the atmosphere varies with EM wavelength, which accounts for why sensors are only found in certain wavebands. For much of the EM spectrum the atmosphere is opaque and therefore unusable for remote sensing of the ocean. However, there are a number of "window" regions where most of the radiation gets through although it may be attenuated to some extent. One of these windows extends from the visible part of the spectrum (between 400 nm and 700 nm, used by the human eye) into the near-infrared (IR). This is used by radiometers that observe sunlight reflected from the ground and the ocean. In the context of this book these are termed ocean color sensors (Section 2.4.2 outlines how they work). Between wavelengths of about 3.5 μm and 13 μm are found a number of narrow windows exploited by IR radiometers. This is called the thermal IR part of the spectrum because much of the radiation has been emitted by surfaces according to their temperature. In ocean remote sensing it is used for measuring sea surface temperature (SST).

At much longer wavelengths, greater than a few millimeters, the atmosphere becomes almost completely transparent. This is referred to as the microwave part of the spectrum, normally differentiated spectrally in terms of frequency rather than wavelength. Because it is widely exploited for many technological aspects of modern civilization, including radio and TV broadcasts, telecommunications, mobile telephony, and so on, certain parts have to be specially reserved for remote sensing and are allocated by international regulation. They are found as discrete narrow-frequency bands within the broad regions indicated in Figure 2.10 for microwave radiometry and radars. Microwave radiometers are passive sensors, simply measur-

ing the naturally ambient radiation that is emitted by the ocean, atmosphere, and land surfaces. Radars are active microwave devices which emit pulses and measure echoes from the sea surface, in order to gain information about some aspect of the surface. Radar frequency bands are sometimes referred to by the letters indicated in Table 2.2.

Table 2.2. Definition of common radar bands used for ocean remote sensing.

Band	*Frequency* (GHz)	*Wavelength*
L	0.390–1.55	19.35–76.9 cm
S	1.55–4.20	7.14–19.35 cm
C	4.20–5.75	5.22–7.14 cm
X	5.75–10.9	2.75–5.22 cm
K_u	10.9–22.0	1.36–2.75 cm
K_a	22.0–36.0	8.33–13.6 mm

There is a variety of different types of radar, which can be distinguished by the direction in which they point, the length and modulation of the emitted microwave pulse, and the way the echo from the sea surface is analyzed. Radars can be classed as either viewing straight down at the nadir point below the platform, or viewing obliquely to encounter the surface at an incidence angle between 15° and 60°. The nadir sensors measure surface height or slope and are called altimeters. Those viewing obliquely measure a property called σ_0, the normalized radar backscatter cross-section, which is related to surface roughness at lengthscales comparable with the radar wavelength.

In the rest of Section 2.4 the measuring capability of each sensor class is summarized in a separate subsection. The principles of operation of each sensor type is outlined, while the most important elements of the data-processing and interpretation tasks are highlighted. It is also important to recognize that in general the direct measurements made by a satellite sensor, even after applying atmospheric corrections when necessary, are not themselves oceanographic parameters. Apart from the example of sea surface temperature, the primary quantity observed by most sensors requires a further stage of interpretation to generate derived parameters that are useful for oceanographers, as summarized by the bottom row of Figure 2.11

2.4.2 Ocean color radiometers

The basic principle of ocean color remote sensing is straightforward. The light measured by an ocean color sensor pointing towards the sea comes originally from the Sun. Some photons of light emitted by the Sun, with energies that place them in the visible part of the spectrum, enter the sea where they are either absorbed or scattered depending on what is contained in the seawater. Those of the scattered photons that emerge again give the sea its apparent color. This is quantified by a satellite ocean color sensor which measures the amounts of different wavelengths of light reaching it. A *multispectral radiometer* typically samples a limited number of narrow wavebands, chosen to capture the main structure of the spectral shape of incoming light (see Figure 2.12). An *imaging spectrometer* samples in full detail across the spectrum, but such instruments have so far been used mainly onboard

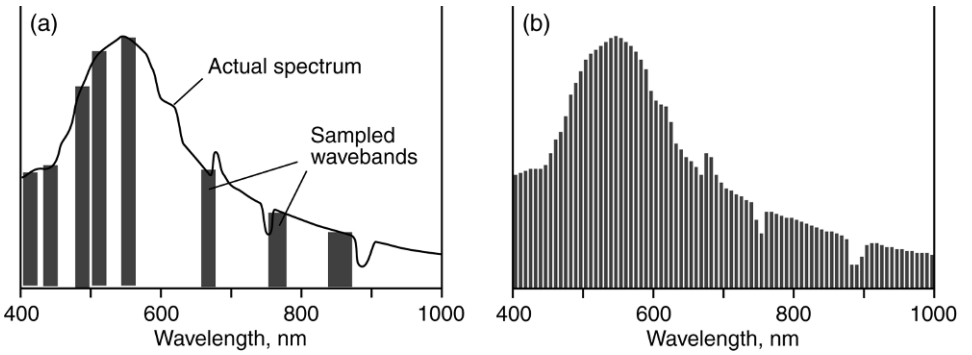


Figure 2.12. A typical spectrum of Earth-leaving radiance in the visible and near-infrared part of the spectrum, observed from a satellite above the atmosphere, as spectrally sampled by (a) a multispectral radiometer and (b) an imaging spectrometer with a spectral resolution of about 5 nm.

aircraft. From the relative magnitude of the water-leaving radiance detected by the different spectral channels of a radiometer, methods have been developed to estimate the concentration of those water constituents which give the sea its color.

The term *ocean color* is used loosely in remote sensing to refer to both the magnitude and the spectral composition of the light leaving the water. In practice it is the spectral radiance at the top of the atmosphere that is measured from a satellite. As shown in Figure 2.13, this consists of light reflected by the atmosphere, the sea surface, and (in very shallow water) the sea bed, as well as backscattered by seawater constituents. The retrieval of useful oceanographic quantities from top-of-atmosphere measurements is a challenging task, requiring careful separation of

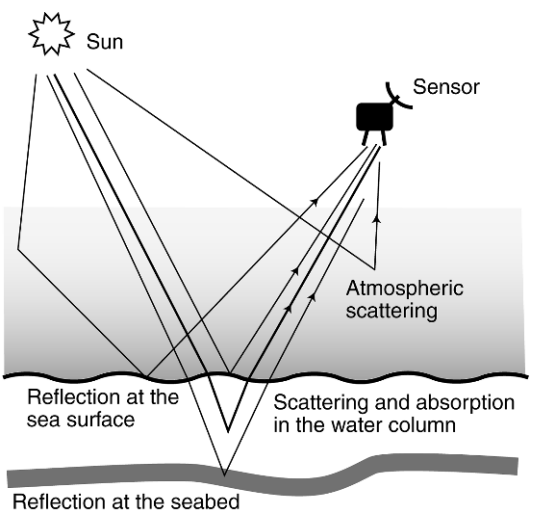


Figure 2.13. Factors which affect the light reaching an ocean color sensor.

atmospheric scattering and surface reflection from true water-leaving radiance. The rest of this section briefly outlines how this is tackled, but for a more comprehensive introduction to the physical principles underlying ocean color remote sensing, and the detailed methodology which has matured over more than 20 years, the reader should consult *MTOFS* (chapter 6 and the references therein to the scientific literature).

Atmospheric correction

The greater part of the measurement of visible wavelengths of light made looking down from a satellite orbit comes from light scattered by the atmosphere into the field of view of the sensor, which may contribute more than 90% of total measured radiance. A proportion of the light leaving the sea surface is also scattered out of the field of view and so the atmospheric correction procedure must account for both these factors in order to estimate water-leaving radiance for each of the spectral bands recorded by the sensor.

That part of atmospheric correction due to scattering by air gas molecules themselves can be calculated directly for each pixel in the field of view of an imaging sensor, from knowledge of the relative positions of the Sun, the satellite, and the pixel. Scattering is also caused by larger particles of aerosols in the atmosphere. These may be water vapor or dust particles, but unlike atmospheric gases their concentration and distribution in the atmosphere is unknown and impossible to predict. Fortunately a technique has been developed which allows the variable contribution of scattering from aerosols to be estimated from measurements made by the sensor.

The key is to use radiance measured in two spectral bands from the near-infra-red part of the spectrum. Because the sea readily absorbs almost all solar near-infrared radiation incident upon it, any light in these wavebands measured at the top of the atmosphere must have been scattered by the atmosphere or reflected at the surface. This can be used to estimate how much aerosol scattering has occurred in visible channels where the water-leaving radiance is not zero, and so the correction is accomplished. Thus near-infrared channels are essential for an ocean color sensor, although care must be taken to select the correct wavebands that do not overlap gas absorption lines in the spectrum.

Reflection from surfaces

Part of the measured signal is reflected directly from the sea surface and this has no value for quantifying the water content. If the satellite detects direct specular reflection from the Sun it will dominate all other parts of the signal. Sun glitter prevents further analysis of the signal and must be avoided if at all possible, by careful selection of the orbit and sensor geometry in relation to the position of the Sun, and by tilting the sensor away from the Sun, bearing in mind that the extent of the Sun glitter region on the sea surface varies with the roughness of the sea.

Reflection from the sea surface of sky light (i.e., sunlight already scattered by the atmosphere) cannot be avoided. Instead corrections can be made for it within atmo-

spheric correction. This means that knowledge of sea state (based on wind speed) is also needed for atmospheric correction.

The other surface which may contribute reflections to the measured signal is the sea bed. Since sunlight is fairly rapidly attenuated with depth in seawater, by both absorption and scattering in the water column, reflection from the sea bed is only a problem when the sea is both shallow and clear. In most circumstances it is not an issue, but it can create problems when interpreting ocean color data in tropical, shallow seas. Where the water is very clear then bottom reflections in shallow seas can allow the color signal to be used to detect either the depth of the sea, the character of the sea bed (sand, coral, vegetation, etc.), or both but this is not a major application for global ocean color sensors.

Interpreting ocean color

The color of the sea is not of itself an ocean variable of particular interest for most marine scientists. However, the factors which influence ocean color, such as the presence of phytoplankton, the concentration of pigments associated with primary production or dissolved organic material, and the concentration of suspended particulates are all of considerable oceanographic importance. Measurements of these properties can be derived from the color.

When atmospheric correction has been successfully applied to satellite ocean color data, the result is an estimate of the water-leaving radiance in each spectral channel in the visible waveband, normalized to reduce dependence on the Sun's elevation and the viewing incidence angle. Effectively, normalized water-leaving radiance should represent what a sensor would measure if looking straight down from an orbit that carries it just above the sea surface at the bottom of the atmosphere. This is what our eyes would detect as the color and brightness of the sea, ignoring any light reflected from the surface. The primary challenge of ocean color remote sensing is to derive quantitative estimates of the type and concentration of those materials in the water that affect its apparent color.

Photons from the Sun, with EM energy corresponding to visible wavelength, enter the sea and eventually interact with molecules of seawater or its contents. The outcome will be either that the photon is scattered, in which case it may change its direction with a chance of leaving the sea and contributing to what the sensor measures, or it will be absorbed. The probability of scattering or absorption depends on the wavelength of the light and the material which it encounters. The molecules within seawater tend to preferentially scatter shorter wavelengths of light (the blue part of the spectrum) and preferentially absorb longer wavelengths (the red end). This is why pure seawater with little other content appears blue.

The pigment chlorophyll-*a* which is found in phytoplankton has a strong and fairly broad absorption peak centered at 440 nm in the blue, but not in the green. Therefore, as chlorophyll concentration increases, more blue light is absorbed while green light continues to be scattered, and so from above the seawater looks greener. This is the basis for many of the quantitative estimates of seawater content derived

from satellite ocean color data. The typical form of an algorithm to estimate the concentration of chlorophyll (C) or phytoplankton biomass is:

$$C = A(R_{550}/R_{490})^{B}, \tag{2.3}$$

where A and B are empirically derived coefficients; and R_λ is reflectance (radiance coming out of the sea towards the sensor, normalized by incoming irradiance) over a spectral waveband of the sensor centered at wavelength λ. This is described as the green/blue ratio. In the open sea it is possible to estimate C to an accuracy of about 30% by this means. Most algorithms presently in use are somewhat more complex than (2.3) but are still closely related to it. If the sample data from which the coefficients A and B, etc. are derived is representative of many different open-sea situations then such algorithms can be applied widely in many locations.

Other substances which interact with the light and so change the apparent color of the sea are suspended particulate material (SPM) that has a fairly neutral effect on color except in the case of highly colored suspended sediments, and colored dissolved organic matter (CDOM, sometimes called "yellow substance") which absorbs strongly towards the blue end of the spectrum. Both of these affect the light along with the chlorophyll "greening" effect when there is a phytoplankton population. However, because the chlorophyll, CDOM, and SPM all co-vary within a phytoplankton population the green–blue ratio effect dominates the color and each of these materials can be quantified by an algorithm such as (2.3), as long as phytoplankton are the only major factor other than the seawater itself that affects the color. Such conditions are described as being Case 1 waters, and it is here that ocean color algorithms work fairly well to retrieve estimates of C from satellite data.

However, if there is SPM or CDOM present from a source other than the local phytoplankton population (e.g., from river runoff or resuspended bottom sediments), then we can no longer expect any simple relationship between the concentrations of these and C. In this situation green–blue ratio algorithms do not perform very well, if at all, and it becomes much harder to retrieve useful quantities from ocean color data using universal algorithms. These are described as Case 2 conditions. Unfortunately it is not easy to distinguish between Case 1 and Case 2 waters from satellite data alone. This can result in very degraded accuracy with errors of 100% if standard chlorophyll algorithms are applied in Case 2 waters. It is prudent to classify all shallow-sea areas as Case 2, particularly where there are riverine and coastal discharges or strong tidal currents stirring up bottom sediments, unless *in situ* observations confirm that Case 1 conditions apply.

Another useful measurement that can be derived from ocean color is the optical diffuse attenuation coefficient, K, usually defined at a particular wavelength such as 490 nm (i.e., K_{490}). This is also inversely correlated with the blue–green ratio because the less the attenuation coefficient, the deeper the light penetrates before it is scattered back out, the more of the longer wavelengths are absorbed, and the bluer the water appears. The algorithms for K are similar in form to (2.3) and are somewhat less sensitive to whether the conditions are Case 1 or Case 2.

Ocean color sensors and products

Table 2.4 in Section 2.5 lists the important ocean color sensors that have been flown. Although visible wavelength radiometers were among the very first Earth-observing sensors used in the 1970s, a gap of 18 years to the next launch has resulted in the development of ocean color sensors being less mature than that of other methods of satellite oceanography. The Sea-viewing Wide Field-of-view Sensor (SeaWiFS) launched in 1997 provided the first reliable, long-term, fully operational delivery of ocean color data products. Since then a number of sensors with finer spectral resolution have been flown. All the sensors in Table 2.4 fly in low (~800 km) Sun-synchronous polar orbits, providing a resolution at nadir of about 1.1 km and almost complete Earth coverage in 2 days. MERIS is also capable of a high-resolution mode with 300 m square pixels. Other color sensors have also been flown by individual countries offering less comprehensive coverage and poorer data availability than those listed.

All of the sensors listed in Table 2.4 are supported by a calibration and validation program and their data are worked up by the responsible agency into derived oceanographic products at level 2 and in some cases level 3. In all cases some measure of C is produced and an estimate of K is derived globally. These are generally reliable products and C approaches the target accuracy of 30% in open-sea Case 1 waters. However, great care must be taken when using the products over coastal water and shelf sea water (possibly Case 2) where the potentially large errors could give misleading information. Some agencies offer a number of other products estimating the concentration of certain material in the seawater such as SPM and CDOM but these are yet to be proven.

In addition, some allied products are offered by some agencies providing ocean color data. These may be partially derived from the ocean color sensor but have input from other sources: satellites, *in situ* measurements, or models. In that respect they could be classed as level 4 products. Examples of these are surface solar irradiance (SSI) and photosynthetically available radiation (PAR).

2.4.3 Thermal infrared radiometry for measuring sea surface temperature

Whereas visible waveband radiometers rely on reflected sunlight and can operate only during the local daytime, in the thermal infrared and microwave parts of the spectrum most observed radiation will have been thermally emitted by the sea surface. In this way infrared and microwave radiometers can be used directly to measure the radiation temperature of surfaces. Given knowledge about the emissivity of the sea surface this can be used to estimate the physical temperature of the water. For infrared measurements there is a close relationship between emitted infrared radiation and sea surface temperature (SST). The challenges facing the satellite oceanographer are how to remove the effect of the atmosphere from measured infrared radiance and still derive SST estimates to an accuracy within a few tenths of a Kelvin, and how to relate temperature measured by a radiometer from

above with what a thermometer placed in contact with the sea would measure. These are the issues introduced in this section.

Physical principles

An infrared sensor records the radiance detected at the top of the atmosphere in specific wavebands, λ_n. The individual measurements in each channel, n, can be expressed as an equivalent black-body brightness temperature, T_{bn}. That is the temperature at which a black body (a surface with 100% emissivity) would emit the measured radiance. At a particular wavelength, black-body emission is defined by the Planck equation:

$$L(\lambda, T) = \frac{C_1}{\pi\lambda^5[\exp(C_2/\lambda T) - 1]}, \tag{2.4}$$

where L is spectral radiance, per unit bandwidth centered at λ, leaving the unit surface area of the black body, per unit solid angle ($\mathrm{W\,m^{-2}\,m^{-1}\,str^{-1}}$); λ is the wavelength (m); T is the temperature (K) of the black body; $C_1 = 3.74 \times 10^{-16}\,\mathrm{W\,m^2}$; and $C_2 = 1.44 \times 10^{-2}\,\mathrm{m\,K}$. This must be integrated with respect to wavelength over the measured waveband and convolved with the spectral sensitivity of the sensor in order to represent radiance intercepted by a particular spectral channel. The spectral shape of Equation (2.4) and how it varies with temperature are shown in Figure 2.14. Black-body radiation is an ideal theoretical concept. The actual radiation emitted from the sea surface is a fraction, ε, called the emissivity, which is close to one for seawater and IR radiation.

Figure 2.14 also indicates the "atmospheric window" regions of the thermal infrared spectrum in which radiation passes through the atmosphere with only a small amount of attenuation, in the absence of cloud. These are found between about 3.5 μm and 4.1 μm (referred to as the 3.7 μm window), and between 10.0 μm and 12.5 μm. The latter is often used for two separate wavebands, 10.3 μm to 11.3 μm and 11.5 μm to 12.5 μm, generally referred to as the "split window" channels. All the main sensors used to measure SST by infrared radiometry use these three channels, including the two main IR sensor series referred to in this book. These are the Advanced Very High Resolution Radiometer (AVHRR) developed by the National Oceanographic and Atmospheric Administration (NOAA) in the U.S.A. and the Along-track Scanning Radiometer (ATSR) a British-designed sensor developed and flown by the European Space Agency (ESA).

To obtain T_{bn} from the digital signal S_n recorded by the sensor for waveband n requires direct calibration of the sensor using two onboard black-body targets of known temperatures which straddle the range of ocean surface temperatures being observed. This is the method adopted by the ATSR class of sensor, whereas the AVHRR uses the simpler but less accurate alternative of a single onboard black body with a view of cold space serving as an alternative to the second black body. Calibration targets are viewed once for every scan across the swath.

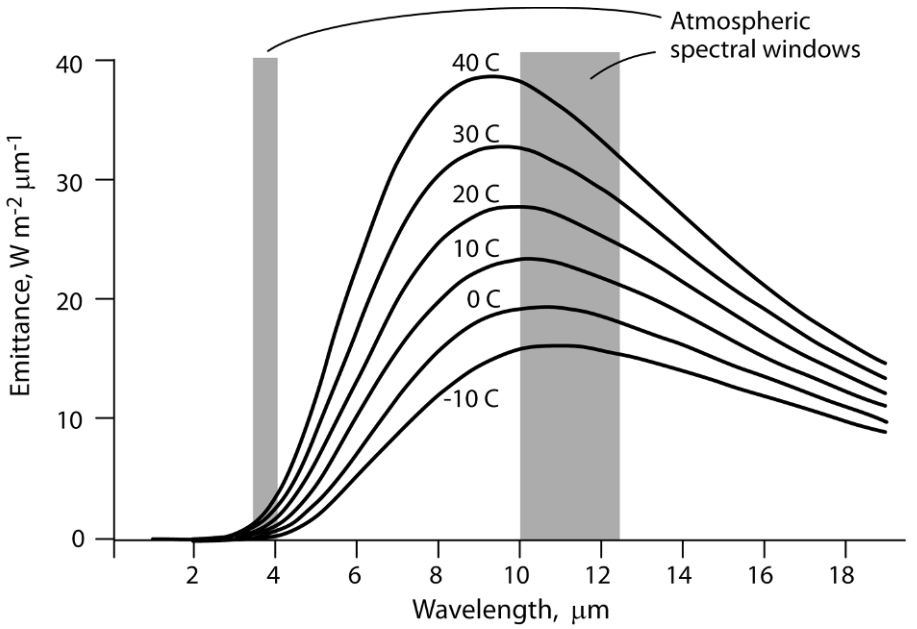


Figure 2.14. Infrared emission spectra of black bodies at temperatures between −10°C and 40°C. The gray bands show the location of atmospheric windows. At other wavelengths the atmosphere is opaque.

Atmospheric correction

Ideally we wish to measure the radiance leaving the water surface, which is determined by the skin temperature of the sea, T_s, and by the emissivity of seawater. In the thermal infrared this is greater than 0.98, but a small contribution to satellite-detected radiance comes from the reflected sky radiance, for which allowance must be made. Because of absorption by gases in the atmosphere T_{bn} is cooler than T_s by an amount which varies in time and place, mainly with the amount of atmospheric water vapor. It is the task of the atmospheric correction procedure to estimate T_s given top-of-atmosphere measurements of T_{bn}.

A well-established method of atmospheric correction is to make use of differential attenuation in different wavebands. This is illustrated schematically in Figure 2.15. In case A, water vapor is assumed to be less than in case B, causing less absorption in both channels i and j for case A than for B, and requiring a larger correction to be applied to B than to A, although we cannot say directly how big either correction should be. However, if channels i and j respond differently to water vapor they will record a difference $\Delta_{i,j}(T_b)$ between the top-of-atmosphere brightness temperatures T_{bi} and T_{bj} measured simultaneously by each channel. This spectral difference is also related to the number of absorbing gases in the atmospheric path and so $\Delta_{i,j}(T_b)$ will be smaller for case A than for B. Thus the size of $\Delta_{i,j}(T_b)$ provides a measure of the degree of atmospheric attenuation. This is utilized by

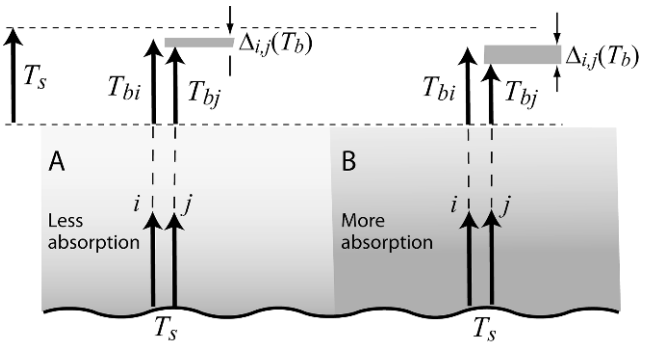


Figure 2.15. Schematic to illustrate the principle of using band-differential response to the atmosphere as the basis of atmospheric correction algorithms. The length of the arrows indicates the magnitude of the temperature represented.

algorithms of the form:

$$T_s = aT_{bi} + b(T_{bi} - T_{bj}) + c, \tag{2.5}$$

where a, b, and c are coefficients to be determined. During the day, the algorithm uses only split window channels, while at night the 3.7 μm channel can also be used. The latter is corrupted by reflected solar radiation and so cannot be used in the daytime. A number of nonlinear variants of this basic form have also been developed (Barton, 1995).

Common to these algorithms for AVHRR is the requirement for coefficients to be determined by a best fit between satellite predictions and coincident observations of SST from a number of drifting buoys. The match between buoys and satellites has a variance of more than 0.5 K, applicable only to regions populated by buoys. The same algorithms are assumed to apply to parts of the ocean where there are no buoys, although the validity of this assumption needs to be quantified. Regional algorithms matched to local data can achieve greater accuracy.

Although the instantaneous distribution of water vapor and aerosols in the atmosphere are not known, the radiation transfer physics of the atmosphere is well understood and can be modeled with some confidence in fine spectral detail. It is therefore possible to simulate T_b for a given combination of T_s, atmospheric profile and viewing angle for the spectral characteristics, and viewing geometry of each channel of a particular sensor. This offers an alternative strategy for atmospheric correction in which an artificial dataset of matching T_s and T_{bi}, T_{bj}, etc. is created using a wide variety of typical atmospheric water vapor and temperature profiles. The coefficients for an equation of form similar to Equation (2.5) are generated by a regression fit to the artificial dataset. The resulting algorithm should be applicable to all atmospheric circumstances similar to those included in the modeled dataset, leading to an estimate of the skin SST. It is independent of coincident *in situ* measurements, although those are needed for validation. This is the approach adopted for the ATSR.

The ATSR scans conically to observe a forward view at about 60° incidence angle and a near-nadir view of the same patch of sea about 2 min to 3 min later (as illustrated in Figure 2.16). Using the same three spectral channels as AVHRR for

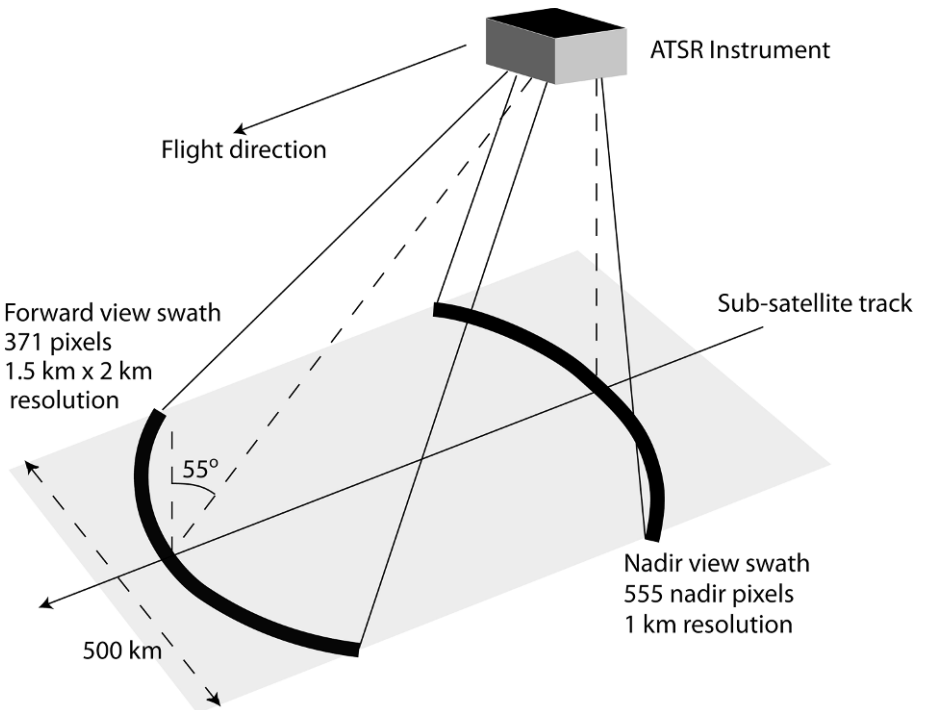


Figure 2.16. The conical scanning arrangement for the ATSR.

each of the views, it thus acquires six measures of brightness temperature (four during the daytime). The different pathlengths for forward and nadir views provide extra information leading to a more robust algorithm. The single-view approach failed temporarily when large volumes of volcanic dust were suddenly injected into the stratosphere by the eruption of Mt. Pinatubo in 1991 (Reynolds, 1993), whereas for ATSR a reworking of the semiphysical model by including stratospheric aerosols in the radiation model led to algorithms which cope well with volcanic ash or similar problems (Merchant and Harris, 1999; Merchant *et al.*, 1999).

Cloud detection

Atmospheric correction algorithms produce maps of SST at fine resolution (about 1.1 km) for each overpass. However, atmospheric correction methods cannot retrieve SST when cloud wholly or partly obstructs the field of view. Therefore at this stage cloud must be detected using a variety of tests (Saunders and Kriebel, 1988), so that only cloud-free pixels are retained for oceanographic applications, such as assimilation into models. The most difficult cloud contamination to identify is that by subpixel-size clouds, thin cirrus, or sea fog where only small deviations of temperature occur. Failure to detect cloud leads to underestimation of the SST and can produce cool biases of order 0.5 K. Thus confidence in the cloud detection procedure

is just as important as atmospheric correction for achieving accurate SST. Where uncertainty remains in cloud detection, this should be flagged in the error estimate fields attached to SST products. Cloud detection is generally more successful during daytime, when visible and near-IR image data can be used, than at night.

Which sea surface temperature does a satellite measure?

There is an additional pitfall for interpreting satellite-derived SST data, consequent upon the thermal structure in the top few meters of the ocean. Two distinct factors create near-surface vertical temperature gradients, illustrated schematically in Figure 2.17. First, on sunny calm days a diurnal thermocline tends to develop above which an upper layer is found, a meter or so thick and up to about 1 K warmer than below (although exceptionally it can be several Kelvins warmer). At night the warm layer collapses. Second (and independently of the first effect), the top skin layer of the sea, a fraction of a millimeter thick, tends to be a few tenths of a Kelvin cooler than the water immediately below.

The problem for understanding remotely sensed SST is that there are various methods of measuring SST samples at different levels of the near-surface thermal structure (shown in Figure 2.17). Thus the term "SST" means different things for the thermometer on a buoy's hull, for a sensor in a ship's cooling water intake, for an infrared radiometer, for a microwave radiometer, and for an ocean model. These differences are important when accuracies of a few tenths of a Kelvin are required. They may also vary considerably during the day so that if a single daily sample is

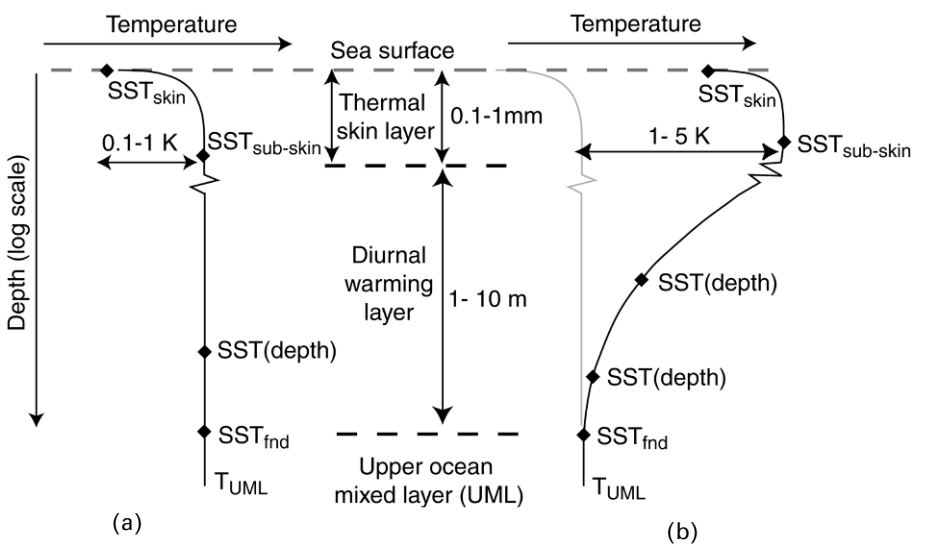


Figure 2.17. Schematic diagram showing characteristic temperature profiles at the sea surface for (a) nighttime conditions or daytime with moderate to strong winds and (b) daytime calm to light wind conditions and direct solar heating.

used to characterize SST the result may be aliased depending on the time in the diurnal cycle at which it is sampled.

Therefore the practice is now being adopted of distinguishing between *skin SST*, the temperature in the top few microns, and *subskin SST* a short distance (of order 1 mm) below the surface. These are separated by the thermal skin layer where heat transport is restricted to molecular conductivity because of the suppression of turbulence close to the surface. The subskin is typically a few tenths of a degree warmer than the actual skin. Infrared radiometers measure skin SST whereas microwave radiometers, penetrating deeper, approximately measure subskin SST.

In addition a new term, *foundation SST*, has been coined to describe the temperature on which diurnal warming (if any) is built each day. It is most clearly specified as the temperature of the well-mixed layer found just below the skin layer at dawn, when any diurnal thermocline structure from the previous day has collapsed. At this time of day it is equivalent to subskin SST. Foundation SST is defined on a daily basis. It corresponds to what oceanographers generally mean when they refer to the temperature of the upper mixed layer of the ocean. When *in situ* measurements are made from buoys or ships they sample the structure shown in Figure 2.17 at an indeterminate depth, typically somewhere between subskin and foundation SST. This creates additional uncertainty for using *in situ* observations to calibrate SST atmospheric correction algorithms.

SST data products

SST level 2 image datasets, at a resolution of about 1 km and arranged in satellite scan co-ordinates, are produced within an hour of AVHRR overpasses and made available for a variety of applications. Because raw AVHRR data are broadcast directly, local receiving stations around the world can obtain highest resolution data and produce SST products in near-real time. Because of the complexity of processing ATSR data, having to resample the forward and nadir views from their curved scan lines onto a rectangular grid before applying the atmospheric correction algorithm, the SST products from ATSR are produced only by ESA, but are still made available within, at most, a few hours of the overpass. Infrared scanners on geostationary satellites generate image datasets every 15 min or 30 min. The most recent geostationary sensors are well calibrated and use multiband radiometry for accurate atmospheric correction, capable of generating SST maps of comparable accuracy to AVHRR and ATSR. The disadvantage of these products is that their coverage does not extend to high latitudes and their spatial resolution is poorer, typically 4 km or 5 km pixels. A summary of oceanographically important infrared sensors is given in Table 2.5 (to be found in Section 2.5).

Until recently, oceanographers looking for a satellite-based record of global or ocean-wide SST distribution and how it evolves with time would go to one of the global composite datasets constructed from a particular infrared sensor (as outlined in Section 2.3.6). A number of different composite SST products have been produced, with lengthscales varying from 1/2° to 1/6° (about 50 km to 16 km at the Equator) or finer, and time steps from 4 days to 1 week or 1 month. From these,

annual climatologies were created allowing SST anomalies to be quantified relative to the seasonal norm (see Section 6.2.1) The most widely used were those derived from AVHRR, such as multichannel sea surface temperature (MCSST) (Walton *et al*., 1998) or Pathfinder SST (Vazquez *et al*., 1998), which is a reprocessing of the archived pixel-level AVHRR data with algorithms incorporating the best knowledge of sensor calibration drift and making full use of the available drifting buoy dataset (Kilpatrick *et al*., 2001). The global composite product from ATSR is called averaged SST (ASST) which was reprocessed using the more robust atmospheric algorithms (Merchant and Harris, 1999; Merchant *et al*., 1999).

However, most oceanographic users looked for the best possible estimate of SST, rather than a product from a particular sensor. Reynolds and Smith (1994) developed a level 4 analysis product that was an optimal interpolation of both *in situ* and satellite data, and should therefore provide better climatological continuity with pre-satellite SST records from before 1980. The most significant development in SST monitoring since 2002 has been an international collaborative initiative to exploit the complementarity between different sensors rather than leaving users to choose one of the competing SST data products. This is the GODAE High Resolution SST Pilot Project—GHRSST-PP (Donlon *et al*., 2007). Its main achievement has been to persuade Earth observation agencies to produce level 2 maps of SST in a common format with essential ancillary data, which has enabled new level 4 SST analyses to be created that draw from all available level 2 products. The primary aim is to improve the usefulness of satellite data for assimilation into ocean-forecasting models (as discussed further in Chapter 14), and for creating robust SST records for climate time series. In Section 2.6 pointers are given for accessing the datasets mentioned above.

2.4.4 Microwave radiometry

It was noted in the previous subsection that radiation in the microwave as well as the thermal infrared parts of the spectrum is thermally emitted by the sea surface and can be used directly to measure the radiation temperature of surfaces. However, unlike the infrared, the microwave brightness temperature of a surface is not related so directly to its physical temperature and depends also on other properties of the surface. This has created a different set of challenges for the development of ocean measurement capabilities using microwave radiometry. Although in many ways the use of passive-microwave radiometers for measuring SST is inferior to using infrared sensors, the technique does have the very significant advantage of being able to see through cloud.

Physical principles of microwave radiometry

The physics of microwave emission from a surface of temperature T is deceptively simple. Ideal black-body emission is expressed as a linear dependence on

temperature in a simplified form of the Planck equation:

$$B_f = \frac{2kT}{c^2}, \tag{2.6}$$

where B_f is spectral radiance expressed as radiance per unit frequency interval. Because of this linear relationship, microwave radiation is itself often referred to as *brightness* temperature (i.e., the temperature of the black-body source that would generate the measured radiance).

However, the use of passive-microwave radiometers to measure SST is more complex than Equation (2.6) might lead us to believe. Unlike thermal IR wavebands, where $\varepsilon \approx 0.98$ and the emitted radiation is dominated by skin temperature, in the microwave region the $\varepsilon < 0.5$ for the sea surface. It turns out that emissivity depends on the viewing incidence angle relative to the local surface slope, and the dielectric constant of seawater as well as temperature. Since the dielectric constant also depends on temperature, the dependence of microwave brightness temperature on subskin SST is not linear. Moreover radiation may change if the mean square surface slope or surface salinity changes, even if the SST remains constant. While this is clearly a drawback for measuring SST, it does offer the possibility of using microwave radiometers for detecting sea surface *roughness*, another of satellite oceanography's primary observable quantities. It also implies that *salinity* may be considered to be a fifth primary observable quantity, and so it has been added to the third row of Figure 2.11 although, at the time of writing, this has still to be demonstrated from a satellite sensor. Figure 2.18 is a cartoon summarizing the various different environmental factors which influence microwave emission from the sea surface and its passage through the atmosphere.

Elements of microwave radiometers

Microwave radiometers are passive devices. Unlike radars they do not create their own coherent source of energy and so are incapable of many of the complex signal-processing techniques used to enhance the resolution of active instruments. Radiometers measure the power of the continuous, incoherent, electromagnetic radiation incident upon their detectors. They sample within specific narrow-frequency bands and some radiometers are capable of differentiating between the power in different polarization orientations. Radiometers are restricted to specific microwave wavebands, not because of atmospheric windows as for infrared and visible radiometers, but because the microwave spectrum is used extensively by modern telecommunications and broadcasting infrastructure. These signals would swamp background radiation from natural sources, but by international agreement certain bands are not permitted for use by radio sources in order to preserve them for passive radiometry.

Some instrument designs use a parabolic reflector to focus the ground view onto the detector. By rotating the reflector about a vertical axis the field of view scans across the ground in a circular arc, maintaining the same incidence angle for all samples. Calibration is achieved by pointing at a calibration source on the satellite.

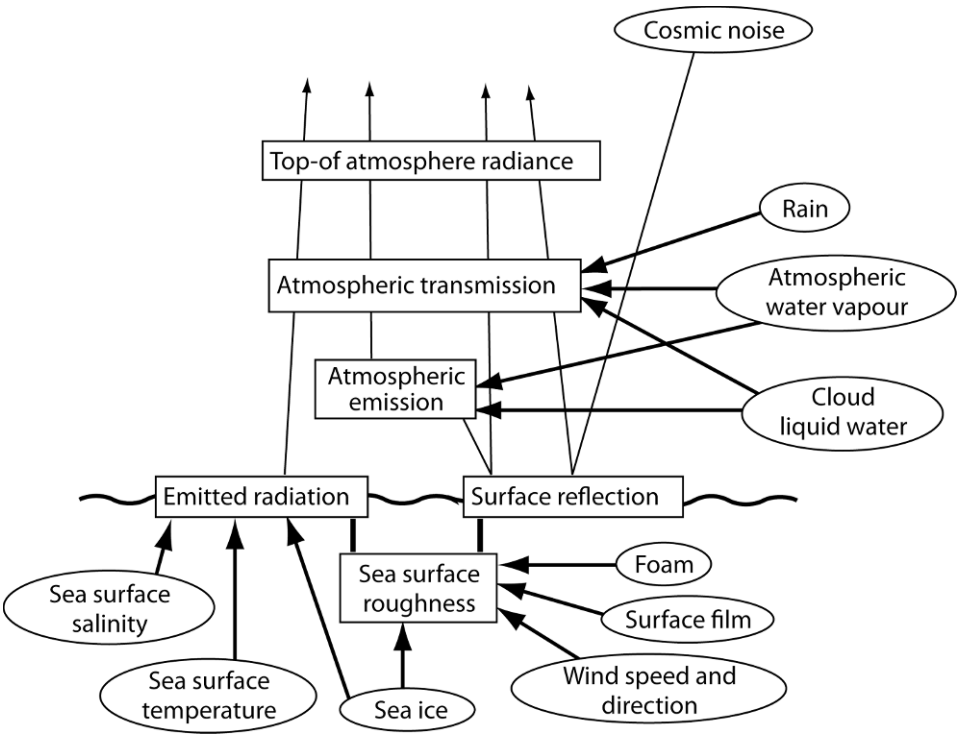


Figure 2.18. Schematic of the physical dependences that determine microwave radiation measured above the atmosphere when viewing the open sea.

In a different type of radiometer design the signal is recorded from each element in an array of unfocused detectors. By integrating these signals with different timelags the effective field of view can be steered to obtain a spatially resolved field of brightness temperature. However, whether mechanical or electrical focusing is used, the spatial resolution is between one and two orders of magnitude poorer than for an IR radiometer. Because of poor focusing, microwave radiometers are not reliable within about 100 km of land, because of the occurrence of stray signals and brightness contrasts between land and sea which leak through sidelobes in the power pattern of antennas.

Retrieving geophysical quantities from microwave radiometers

It is possible to distinguish between different contributions to the brightness temperature of SST, surface roughness, and salinity, as well as to identify atmospheric contamination by liquid water, because each factor differentially affects different microwave frequencies. For example, SST strongly affects wavebands between 6 GHz and 11 GHz whereas the effects of salinity are found only at

frequencies below about 3 GHz. Surface roughness effects influence frequencies at 10 GHz and above, and are also polarization-specific. Thus a multifrequency and multipolarization radiometer can, in principle, be used to measure SST, surface wind, and precipitation (see chapter 8 of Robinson, 2004). The retrieval of useful oceanographic measurements is based mainly on using empirical algorithms, developed from matchups between *in situ* observations and satellite data.

Microwave radiometry missions useful for oceanography

Table 2.6 (see Section 2.5) lists the microwave radiometers used by oceanographers and the data products that they have delivered. Since the mid-1980s a series of Special Sensor Microwave Imagers (SSM/I) have flown on a U.S. defense satellite program, delivering mainly meteorological products (Wentz, 1997), and used by oceanographers to detect sea ice or for wind speed measurements. However, serious consideration of microwave measurements of SST from space started only when a sensor having a 10.7 GHz channel was flown on the Japanese–U.S. Tropical Rainfall Measuring Mission. Called the TRMM microwave imager (TMI) it has a spatial resolution of 0.5° (about 50 km) and because it oversamples it is capable of mapping mesoscale eddies quite effectively using a gridscale of 25 km. It lacks the preferred SST waveband of 6.6 GHz, but its 10.7 GHz channel is sensitive to SST in tropical water temperatures. It covers only latitudes lower than 40°.

In 2002 the Japanese Advanced Microwave Scanning Radiometer (AMSR-E) was launched into a near-polar orbit on the NASA Aqua satellite. This sensor includes a channel at 6.6 GHz, which is sensitive over the full range of sea temperatures, and has opened the way for routine, high-quality, global mapping of SST by microwave radiometry. AMSR-E is now providing global cloud-free SST to an accuracy of ~0.4 K derived from oversampled 76 km resolution data. Composite daily, weekly, and monthly SST products are supplied on a 1/4° grid.

In the context of the GHRSST Pilot Project mentioned in Section 2.4.3, SST retrievals from microwave radiometry are now routinely used to complement infrared radiometers for producing SST analysis (level 4) products, their strong contribution coming from their ability to sample daily with few data dropouts caused by atmospheric conditions. It is therefore of considerable concern to operational oceanographers that the continuity should be maintained of microwave radiometers with 6.6 GHz capability when the existing sensors reach the end of their life.

There are, however, experimental sensors planned to test the capability of ocean salinity measurement from satellites. The European Space Agency's Soil Moisture and Ocean Salinity (SMOS) satellite was launched in late 2009. Its main payload is an electronically focusing synthetic aperture microwave radiometer operating in L band (1.4 GHz). NASA are preparing another L-band radiometer for launch on their Aquarius mission in 2010, focused on measuring global ocean surface salinity.

2.4.5 Altimetry for measuring surface slope, currents, and wave height

A more comprehensive introduction to the scientific principles and the detailed methods of satellite altimetry, and a description of 30 years of evolving altimeters, can be found in chapter 11 of *MTOFS*.

The principle of altimetry over the ocean

A satellite altimeter is a nadir-viewing radar which emits regular pulses and records the travel time, the magnitude, and the shape of each return signal after reflection from the Earth's surface. Travel time is the essential altimetric measurement, leading to determination of the ocean surface topography at lengthscales longer than about 100 km. Ocean surface topography contains information about ocean dynamical and geophysical phenomena. If the travel time can be measured to a precision of 6×10^{-11} s then, knowing the speed of light, the distance can be calculated to a resolution of 1 cm. Corrections have to be made to allow for the changed speed of light through the ionosphere and the atmosphere, and for delays associated with reflection from a rough sea surface (Chelton *et al.*, 2001). It is generally agreed that for these corrections to approach the target accuracy of 1 cm a dual-frequency altimeter must be used (to determine ionospheric refraction), and a three-channel microwave radiometer is needed to sound water vapor in the atmosphere.

The altimeter is not an imaging sensor. Viewing only the nadir point below the satellite, it simply records measurements of distance between the satellite and the sea surface along the ground track. As discussed in Section 2.2, the spatial and temporal-sampling characteristics therefore depend entirely on the exact orbit repeat cycle of the satellite. This was chosen to be about 10 days for the TOPEX/Poseidon (T/P) and Jason altimeters which fly on platforms dedicated to the altimetric mission, although for other altimeters it has ranged between 3 days, 17 days, 35 days, and longer. The longer the revisit interval the finer the spatial-sampling grid. Typically, ocean topography data are interpolated onto a geographical grid and composited over the period of an exact repeat cycle, to produce "images" which are comparable with global SST or ocean chlorophyll composite images although produced in a completely different way.

By itself, knowing the distance R_{alt} between the ocean surface and a satellite is of limited value. Figure 2.19 shows what else needs to be defined or measured for this to yield an oceanographically useful property. First of all, when the height of the satellite, H_{sat}, is known relative to a reference level, then the height, h, of the sea above the reference level can be determined. The reference level is a regular ellipsoid-shaped surface defined within a frame of reference fixed in the rotating Earth. It is chosen to match approximately the shape of the Earth at sea level, and provides a convenient datum from which to measure all other heights.

Several physical factors contribute to h, which is called the ocean surface topography. The first is the distribution of gravity over the Earth, as represented by the geoid, at height h_{geoid} above the reference ellipsoid in Figure 2.19. The geoid is the equipotential surface, at mean sea level, of the effective gravitational field of the Earth which incorporates Earth rotation forces and the gravitation of the solid

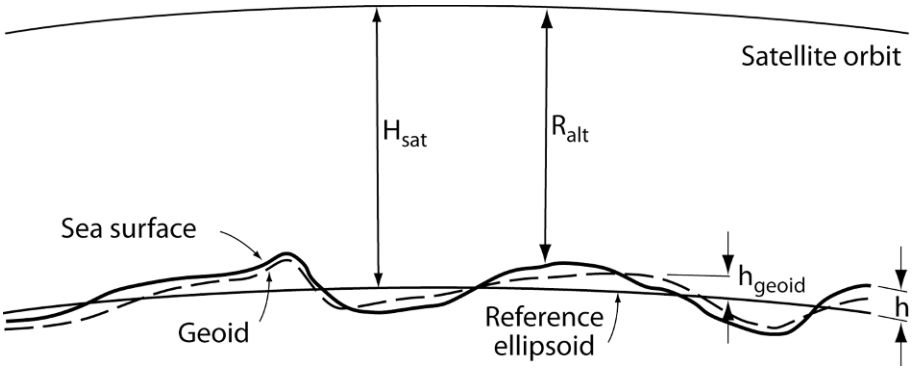


Figure 2.19. The relationship between different distance quantities used in altimetry.

Earth, the ocean itself, and the atmosphere. By definition it is normal to the local effective gravity force, and if the ocean were everywhere in stationary equilibrium relative to the Earth, its surface would define the geoid.

Another factor which contributes to h is h_{tide}, the instantaneous tidal displacement of the sea surface relative to its tidally averaged mean position, including the contribution of the Earth tide. A third is the local response, h_{atm}, of the ocean to the atmospheric pressure distribution over the ocean, approximated by the inverse barometer effect in which an increased pressure of 1 mbar lowers sea level by 1 cm. The remaining factor is displacement of the sea surface associated with the motion of the sea, called the ocean dynamic topography h_{dyn}. Thus:

$$h = h_{dyn} + h_{geoid} + h_{tide} + h_{atm}. \tag{2.7}$$

The dynamic topography is the property which is of most relevance for ocean modeling since it contains information about ocean circulation. Rearranging (2.7) and substituting $h = H_{sat} - R_{alt}$ yields:

$$h_{dyn} = H_{sat} - R_{alt} - h_{geoid} - h_{tide} - h_{atm}. \tag{2.8}$$

The accuracy and precision of the estimated ocean dynamic height depends not only on the altimetric measurement itself but also on the other four terms in (2.8). For dedicated altimetry missions flying at a height of about 1,340 km where atmospheric drag is minimal, the height of the satellite in orbit, H_{sat}, can now be predicted to a precision of 2 cm using a combination of laser and microwave-tracking devices and an orbit model using precise gravity fields. The tidal contribution has been evaluated along the repeat orbit track by tidal analysis of the altimeter record spanning several years. Because tidal frequencies are very precisely known the response to each constituent can be evaluated to an accuracy better than 2 cm in the open ocean, even though the sampling interval of about 10 days is longer than most tidal periods. This is only possible when the precise period of the repeat cycle is chosen to avoid any serious aliasing with one of the major tidal constituents. For this reason a Sun-synchronous orbit, which aliases the S_2 (solar semidiurnal) tidal signal, should not be

used. Over shelf seas where tides are very high and can vary rapidly over short distances it is not so easy to remove the tides and so the estimate of dynamic height is less accurate. Atmospheric pressure correction is based on the output of atmospheric circulation models.

Evaluating sea surface height anomaly

At the time of writing, the geoid has been measured independently, but not yet to a very high precision, and so oceanographers must be content with measuring the combined $h_{dyn} + h_{geoid}$. Of these, the typical magnitude of the spatial variability of h_{geoid} is measured in tens of meters, about 10 times greater than that of h_{dyn}, which is why until recently the time-mean ocean topography from altimeters provided geophysicists with the best measure of the geoid. However, h_{geoid} does not vary with time, at least not sufficiently to be detected by an altimeter over tens of years, whereas the time-variable part of h_{dyn} is comparable in magnitude with the mean component, of order meters over a few months.

Therefore the time variable part of h_{dyn}, called the sea surface height anomaly, SSHA, can be separated from the measured $h_{dyn} + h_{geoid}$ by simply subtracting from it the time-mean sea surface height over many orbit cycles ($\mathrm{MSS} = \mathrm{Mean}\{h_{dyn} + h_{geoid}\}$). To enable a time-mean to be produced, the orbit track must be precisely repeated to within a kilometer and the data must be accumulated from several years of a 10-day cycle. For this reason it is essential to fly a new altimeter in precisely the same orbit as its predecessor so that the mean surface topography of the earlier mission can be used straight away. Then SSHA can be calculated from the first orbit cycle of the new altimeter, without having to wait another few years to build up a new mean topography for a different orbit track.

It is important to remember that the SSHA, which is widely used for oceanographic analysis and is assimilated into dynamical ocean models, does not contain any information about the dynamic height of the ocean associated with the mean circulation. Global maps of SSHA do not display the dynamic topography signatures of the strong ocean currents at all, apart from the fact that the eddy-like activity is strongest where the major currents tend to meander.

There are presently three families of altimeters in operation, as listed in Table 2.7 in Section 2.5, with details of their attitude, orbit repeat, and approximate accuracy (root mean square) of an averaged SSHA product. The T/P–Jason family is a joint French/U.S.-dedicated altimetry mission in a high non–Sun-synchronous orbit. In contrast the Geosat and ERS series are on lower Sun-synchronous platforms for which orbit prediction accuracy would be, on their own, much poorer. However, because these satellites cross over each other's orbit tracks it is possible, over an extended timespan, to significantly improve their orbit definitions by cross-referencing to the better known T/P or Jason orbits (Le Traon *et al.*, 1995; Le Traon and Ogor, 1998). The accuracy quoted for the SSHA applies only after this procedure has been performed, and would otherwise be much worse for the ERS and Geosat families. The specification of errors for an altimeter must be handled with care because the error magnitude relates very much to the time and spacescale over

which it is being averaged. The lower error attached to larger scale/longer period averaging must be offset against the lesser utility of the averaged SSHA field, especially in the context of operational oceanography.

The data products from altimeters are presented first as along-track values of SSHA, wind speed (determined from the peak height of the echo), and significant wave height (from the pulse shape—see below and Chapter 8). These level 2 products are sampled every second along track, and are contained in the Geophysical Data Record (GDR), which also includes ancillary information about the various corrections applied. Whilst each agency publishes the GDR for their own altimeter, scientific users of the data may find it most helpful to work with data where cross-referencing between different altimeters has been performed in a consistent way, referred to as the Data Unification and Combination System (DUACS), ensuring that there should be little if any bias between the SSHA from different satellites. Data are also resampled onto a $1/3^\circ \times 1/3^\circ$ Mercator grid, integrated over a period of time that relates to the orbit repeat interval. Details of these vary between the different agencies producing products (see Table 2.10).

Variable currents from sea surface height anomaly

To determine an estimate of the time-variable part of ocean surface currents, geostrophic equations are used:

$$\left.\begin{aligned} fv &= g\frac{\partial h_{SSHA}}{\partial x} \\ fu &= -g\frac{\partial h_{SSHA}}{\partial y} \end{aligned}\right\}, \tag{2.9}$$

where (u, v) are the east and north components of geostrophic velocity; f is the Coriolis parameter; g is the acceleration due to gravity; and x and y are distances in the east and north direction, respectively.

From a single overpass, only the component of current in a direction across the altimeter track can be determined, but where ascending and descending tracks cross each other the full vector velocity can be estimated. Because Equation (2.9) assumes geostrophic balance, if there is any ageostrophic surface displacement it will lead to errors in (u, v). However, ageostrophic currents should not persist for longer than half a pendulum day $(1/f)$ before adjusting to geostrophy. Thus the spatially and temporally averaged SSHA maps produced from all the tracks acquired during a single repeat cycle (10, 17, or 35 days depending on the altimeter) should represent a good approximation to a geostrophic surface that can be inverted to produce surface geostrophic currents.

Close to the Equator the SSHA cannot be interpreted directly in terms of surface currents since here f is very small and the geostrophic equations (2.9) cannot be applied.

New altimeter data products containing independent geoid data

In the relatively near future it is hoped that the lack of knowledge about the geoid can be remedied. What is needed is a means of measuring h_{geoid} without using altimetry, and this is provided by measurement of the gravity field above the Earth from satellites. Both the presently operating Gravity Recovery And Climate Experiment (GRACE) and the Gravity and Ocean Circulation Explorer (GOCE) mission which was launched in mid-2009 (although first results are still awaited as this book goes to press) measure elements of the gravity field from which it is possible to recreate the sea level geoid. At the required accuracy of about 1 cm GRACE can achieve this only at a lengthscale longer than several hundred kilometers, but it is hoped that GOCE can do so once and for all down to a lengthscale of about 100 km. This will allow steady-state ocean currents to be derived from archived altimetric data and greatly improve the capacity to utilize altimetric data in near-real time.

In anticipation of the eventual availability of a high-quality, independent geoid from GOCE, a hybrid mean dynamic topography (MDT) was produced (Rio and Hernandez, 2004) using the following approach. The absolute dynamic topography of the sea surface, h_{dyn}, is the sum of the SSHA and MDT. Eventually MDT should be determined precisely by subtracting an independent measure of h_{geoid} from MSS. This was done approximately using the EIGEN-GRACE03S geoid, evaluated to spherical harmonic degree 30, which implies that it contains little useful information on geoid variability at lengths less than 400 km but is quite well defined for lengthscales above 660 km. To improve the accuracy of MDT at shorter lengthscales it was fitted to the dynamic height associated with *in situ* measurements of steady currents using an inverse technique. The *in situ* data were buoy velocities from the WOCE–TOGA program, corrected for mesoscale variability using coincident SSHA. Comparison with independent velocity observations show differences to be globally less than 13 cm/s r.m.s. From MDT a new altimetry product is produced (Table 2.10) called the absolute dynamic topography (ADT = MDT + SSHA) from which absolute currents can be estimated using standard geostrophic retrieval, following Equation (2.9) with h_{ADT} replacing h_{SSHA}.

Measuring significant wave height from altimeters

When an altimeter measures the time for an emitted pulse to return, it tracks in detail the shape of the leading edge of the echo, from which it is possible to make a very good estimate of significant wave height, $H_{1/3}$, within the pulse-limited footprint illuminated by the altimeter. For a perfectly flat, calm surface the return echo has a very sharp edge. If there are large waves, several meters in height from trough to crest, then the return signal starts to rise earlier, as the first echoes are received from the crests, but takes longer to reach its maximum, when the first echoes are received from the wave troughs. The rising edge of the echo is modeled by a function in terms of the root mean square ocean wave height, so that by matching the observed shape to the model function it is easy to gain an estimate of $H_{1/3}$. This method has delivered robustly accurate measurements of $H_{1/3}$ for more than 20 years from

different altimeters (Cotton and Carter, 1994) and comparison with buoys shows root mean square differences of only 0.3 m (Gower, 1996), which is the limit of buoy accuracy. Applications of this method to wave monitoring and forecasting are discussed further in Chapter 8.

2.4.6 Oblique-viewing radars for measuring sea surface roughness

Active microwave devices provide their own energy, in the form of radar pulses which are emitted from spacecraft, reflected from the sea surface, and received back at the sensor again. Since the amount of energy reflected depends largely on the short-scale profile of the surface at lengthscales comparable with the radar wavelength, most radars provide information about sea surface *roughness* as the primary observable quantity.

Interpreting radar backscatter measurements

The magnitude of the radar echo reflected from the sea surface is expressed as a variable called the normalized radar backscatter cross-section, usually referred to by the symbol σ_0. After calibration, the data from a radar yield estimates of σ_0, either as single averages for a given field of view or as an array of many samples mapped over the sea surface. The size of σ_0 depends on surface roughness and in particular on the amplitude of short waves on the sea surface propagating in the radar ground range direction and having a wavelength of $n\lambda/(2 \sin \theta)$—where n is 1, 2, etc.; λ is the radar wavelength; and θ is the radar incidence angle. This is the Bragg resonance mechanism, as a consequence of which different radar frequencies produce different magnitudes of echoes from the same sea surface and same incidence angle.

Figure 2.20 illustrates broadly how σ_0 varies with incidence angle under different wind conditions or sea states. The behavior of σ_0 can be separately characterized in three ranges of incidence angle. At low incidence angles (a) specular reflection appears to be the dominant process. For a very calm sea there is a very narrow angular response giving a very high return at 0° incidence which rapidly drops off as the incidence angle increases. For a somewhat rougher surface under moderate winds the nadir-viewing response is weaker, but does not decay so rapidly with increasing viewing angle, so that within a few degrees from normal incidence it is reflecting more power than the flat surface. The very high sea state continues the trend, with an even lower σ_0 at 0° but very little dropoff with incidence angle. This is the way in which the magnitude of altimeter pulses responds to surface roughness.

In a central region of the diagram (b) at incidence angles between about 20° and 70°, appropriate for most oblique-viewing radars, the behavior of σ_0 is quite simply described. At a given θ it increases with sea state, while there is an approximately linear reduction with increasing viewing angle, except for a calm sea that is already very low. Finally at incidence angles greater than 70° (c) the value of σ_0 appears to drop off more rapidly with θ, reaching very low values at grazing incidence approaching 90°. The broad dependence on sea state, albeit different in different bands of viewing angle, is what makes σ_0 such a useful parameter for marine remote

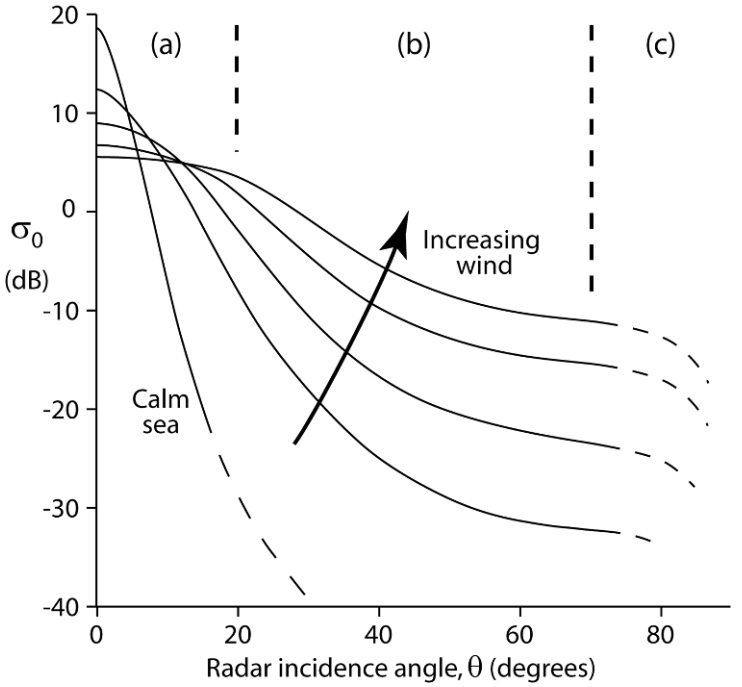


Figure 2.20. Sketch of typical measurements of σ_0 as a function of incidence angle and sea state. The curves show the responses to different winds.

sensing. Note that in practice σ_0 depends also on other parameters such as frequency and polarization, so Figure 2.20 is not intended to be precise. Sections 9.3 and 9.4 of *MTOFS* (Robinson, 2004) provide a much fuller discussion of radar backscatter from the ocean, with many references to the wide literature on this subject.

Scatterometers

A scatterometer is the simplest type of radar used for remote sensing. It is an oblique-viewing radar pointed towards the sea from aircraft or satellites at incidence angles normally between 20° and 70°. The receiver simply measures the backscattered power from the field of view in order to determine σ_0. There is no attempt to preserve phase information after demodulation of the microwave signal. Therefore it does not resolve variations of σ_0 in range or azimuth in a detailed way and cannot generate a high-resolution image. By measuring the average σ_0 over a wide area of sea (with a spatial resolution typically 20–50 km) it uses this to estimate the wind speed.

The interpretation of scatterometer measurements of backscatter relies on an empirically derived model of the relationship between σ_0, wind speed, incidence angle, and the direction of the wind relative to the radar azimuth. As long as σ_0 at each point on the ground is measured at least twice in close succession, viewing from different directions, there is in principle enough information to be able to retrieve an estimate of wind speed and direction using this model. Scatterometers

deployed to provide operational meteorological measurements have swaths spanning about 1,500 km and can view the global ocean surface twice in two days. A much fuller discussion of scatterometers and the principles of scatterometry can be found in sections 9.6 and 9.7 of *MTOFS* (Robinson, 2004). Chapter 9 of the current book shows how wind fields measured by scatterometers are contributing to oceanographic applications.

Imaging radars

With an active device, there is scope to measure not only the energy flux of the reflected signal, but also its detailed amplitude and phase, depending on how complex a measuring device is used, and how much data can be sampled and transmitted back to the ground station. Thus the timing of the return signal can be used to resolve between patches of sea surface at different distances from the radar. Moreover, the detailed shape of the return pulse can be compared with the pulse that was originally transmitted and, for example, Doppler shifts can be detected. When suitably analyzed, such information can be made to yield further information about the sea surface, and in particular to improve the spatial resolution of detection making it possible to generate detailed images of surface roughness. Instruments that collect such detailed information are known as imaging radars. Most imaging radars on satellites belong to a class known as synthetic aperture radars (SARs) because of the way they process data to recover detailed spatial resolution in the azimuth direction. These are described in detail in chapter 10 of *MTOFS* (Robinson, 2004).

While short-scale surface roughness may not seem at first to be a very important oceanographic parameter, much important oceanographic information can be derived from it in addition to sea surface wind strength and direction which drives it in the first place. There are a number of upper-ocean phenomena and processes which modulate wind-driven, short surface ripples. These become much more evident using SARs that can resolve details down to 30 m. One source of modulation is variable surface tension caused by the presence of surfactant material, which allows SARs to reveal the presence of oil spills and sea surface slicks. Another major cause of modulation is by small-scale patterns of convergent and divergent currents at the sea surface. Driven by phenomena such as long-surface swell waves, internal waves, flow over undulating shallow topography, or ocean fronts and eddies, the hydrodynamic interaction between variable surface currents and the energy of Bragg ripples is able to generate signatures in the σ_0 field. Thus processes whose center of action may be tens of meters below the sea surface are "painted" on the radar images, providing an unexpected opportunity to gain new scientific understanding of subsurface phenomena. For example, Chapter 12 in this volume shows how much new knowledge about internal waves has come from the analysis of SAR image data. The capacity of SAR to contribute a unique perspective to the measurement of long-surface gravity waves is demonstrated in Chapter 8, while other, sometimes unexpected, oceanographic applications of SAR image data will emerge in other chapters.

2.5 PLATFORMS AND SENSORS FOR SATELLITE OCEANOGRAPHY

This section summarizes, mainly in tabular form, information about the main satellites and sensors that have been used for ocean remote sensing. Table 2.3 lists the more important satellites, indicating their type of orbit and the ocean-viewing sensors they have carried. The list is by no means complete, since many agencies from many countries have launched satellites that have delivered some useful ocean measurements. However, this is intended to serve as a pointer to those satellite series which today deliver the majority of data for routine ocean monitoring, to previous satellites whose data form a valuable body of archived ocean observations spanning 30 years, and also to those pioneering satellites that proved the concepts underlying the methods of ocean remote sensing. Where indications are given of ocean-viewing satellites and sensors later than 2009, which is when the tables were compiled, they are based on firm plans of space agencies but readers must confirm for themselves whether they have been successfully launched. A fuller list of most satellites with any relevance to ocean observation up to 2004 can be found in table 3.2 of *MTOFS*.

The rest of the section contains details of the main sensors used for the different ocean remote-sensing methods that were summarized in Section 2.4. Table 2.4 lists the ocean color sensors whose data are most widely used by ocean scientists. Table 2.5 lists the infrared radiometers used for measuring sea surface temperature. In both cases it must be pointed out that several other sensors have been, or still are, delivering data, but not routinely or with wide and fairly open access to the data. The sensors listed here are those that the reader is most likely to encounter in the scientific literature on ocean applications of remote sensing, or whose data can be readily obtained. Table 2.6 lists the third class of passive ocean-observing sensor, microwave radiometers. Note in this case the variety of different data products available from microwave radiometry.

The remaining tables identify the more important active-microwave sensors. Table 2.7 lists the altimeters, Table 2.8 the high-resolution imaging radars (synthetic aperture radars—SARs), and Table 2.9 the scatterometers. Note that more information can be found in Chapter 8 about the various sensors used for measuring ocean surface waves, and in Chapter 9 about sensors used to measure the wind over the sea, while Table 7.1 lists some high-resolution visible and near-infrared sensors used for land mapping but which have potential for seabed mapping in tropical coastal ecosystems.

2.6 SATELLITE OCEAN DATA PRODUCTS

In this final section of the chapter, our attention focuses on the ocean data products that are now readily available for users wishing to apply them to some specific operational monitoring or forecasting task, for ocean scientists to use as observational research data, or simply for the curious who are fascinated to discover how the ocean is behaving. The current trend in the dissemination of satellite data is towards providing users with fully processed end products, validated estimates of an actual

Table 2.3. Satellites carrying important ocean-viewing sensors. Entries in bold refer to series.

Satellite	*Period of useful operation*	*Agency*	*Orbit type*	*Ocean observing sensors or sensor types deployed*
Landsat-1 to -7	1972-present	NASA/ USGS	Leo, P, SS	MSS, TM, ATM, ETM +
Meteosat-1 to -7	1977-present	ESA/Eumetsat	Geo	VISSR
Seasat	July-Sep, 1978	NASA	Leo, NSSS	MMR, Scat, SAR, RA
Nimbus-7	1978–1986	NASA	Leo, P, SS	CZCS, SMMR
TIROS-N	1978–1981	NOAA	Leo, P, SS	AVHRR
NOAA-7 to -18	1981–present	NOAA, U.S.A.	Leo, P, SS	AVHRR/2, AVHRR/3
Geosat	1985(86)–90	NASA	Leo, P, 17-day ERM	RA altimeter
SPOT-1 to -4	1986–present	CNES, France	Leo, P, SS	DORIS, HRVIR, Vegetation
DMSP-F8 to -F15	1987–present	DoD, U.S.A.	Leo, P, SS	SSM/I ocean meteorology
ERS-1	1991–1999	ESA, Europe	Leo, P, SS	RA, AMI (SAR-Scat), ATSR
TOPEX-Poseidon	1992–2005	NASA/CNES	1,336 km, 10-day ERM non-SS	DORIS, Poseidon-1, TOPEX altimeter
GOES-8 to -12	1994–present	NOAA, U.S.A.	Geo	GOES I-M Imager
ERS-2	1995–present	ESA, Europe	Leo, P, SS	RA, AMI, ATSR-2, PRARE
Radarsat-1	1995–present	Canada	Leo, P, SS	SAR

(*continued*)

Table 2.3 (*cont.*)

Satellite	*Period of useful operation*	*Agency*	*Orbit type*	*Ocean observing sensors or sensor types deployed*
ADEOS	1996-1997	NASDA, Japan	Leo, P, SS	OCTS, Scat
Seastar	1997–present	NASA	Leo, P, SS	SeaWiFS ocean color sensor
TRMM	1997–present	NASDA, NASA	Leo, non SS	TMI
Geosat FO	1998–present	U.S. Navy	Leo, P, 17-day ERM	RA altimeter
Quikscat	1999–present	NASA	Leo, P, SS	SeaWinds Scatterometer
Terra	1999–present	NASA	Leo, P, SS	MODIS imaging spectrometer
Jason-1, -2	2001–present	CNES/NASA	Leo, 10-day ERM non-SS	Poseidon-2 altimeter
Envisat	2002–present	ESA	Leo, P, SS	ASAR, RA-2, AATSR, MERIS
Aqua	2002–present	NASA	Leo, P, SS	AMSR-E, MODIS
MSG	2002–present	Eumetsat ESA	Geo	SEVIRI radiometer
ADEOS-2	2002–2003	JAXA, Japan	Leo, P, SS	AMSR, GLI, POLDER, SeaWinds
Coriolis	2003–present	DoD, U.S.A.	Leo, P, SS	Windsat M/w radiometer
METOP-1	2006–present	Eumetsat/ESA	Leo, P, SS	AVHRR/3, HIRS/4ASCAT

Leo = low Earth orbit, Geo = geostationary orbit, P = near polar, SS = sun-synchronous, X d ERM = X-day exact repeat mission.

Table 2.4. Details of major satellite ocean color sensors.

Sensor acronym	*Platform*	*Full name of sensor*	*Agency*	*Dates*	*No. of narrow spectral channels*	
					Visible	*Near-IR*
CZCS	Nimbus-7	Coastal zone color scanner	NASA	1978–1986	4	—
OCTS	Adeos	Ocean color and thermal sensor	NASDA	1996–1997	6	2
SeaWiFs	Sea Star	Sea-viewing wide-field-of-view sensor	NASA	1997–present[a]	6	2
MODIS	TERRA	Moderate-resolution imaging spectrometer	NASA	2000–present[a]	7	2
MERIS	Envisat	Medium-resolution imaging spectrometer	ESA	2002–present[a]	8	3
MODIS	AQUA	Moderate-resolution imaging spectrometer	NASA	2002–present[a]	7	2
GLI	MIDORI	Global imager	NASDA	2002–2003	12	3

[a] October 2009.

ocean variable accompanied by measures of accuracy and reliability. This widens access to satellite-derived data compared with the first two decades of ocean remote sensing when only raw or half-processed data were available, which required the user to acquire skills in remote-sensing data processing before she or he could draw useful information from them. While this trend is making for much greater usability of satellite ocean data, it is still important for users to understand the limitations of the data and how to get the best from them. This section guides the reader towards selecting ocean image datasets from the more reliable sources of ocean satellite data, and identifies some software tools for manipulating these image data.

As outlined in Section 2.3, published satellite data are normally categorized as being processed to a certain level. Most of the useful sources provide ocean data products at level 2 or above (see Table 2.1) and users must consider which level is appropriate for their needs. If they are interested in observing fairly high–resolution ocean phenomena as they occur in a particular region (e.g., tracking the position of a front, or monitoring the development of a localized phytoplankton bloom), then level 2 products would be appropriate. To avoid being swamped by data they should restrict their geographical search area. At level 2 the data may be provided in satellite co-ordinates (with axes along-track and across-track) and so each image from a time sequence of the same nominal area may not necessarily overlay the others precisely. Ideally the geographical locations of each pixel will be specified in another array within the dataset. At level 2 there are likely to be gaps in each image caused by

Table 2.5. Recent and current series of high-quality satellite infrared radiometers.

Sensor acronym	*Platform(s)*	*Full name of sensor*	*Agency*	*Dates (for series)*	*Main[a] IR spectral bands* (μm)
AVHRR/2	NOAA-7, -9, -11, -12, -14	Advanced very high–resolution radiometer, version 2	NASA/NOAA	June 1981–Mar 2001	0.725–1.10 3.55–3.93 10.3–11.3 11.5–12.5
AVHRR/3	NOAA-15, -16, -17, -18, METOP	Advanced very high–resolution radiometer, version 3	NASA/NOAA Eumetsat	May 1998–present[b]	0.725–1.10 1.58–1.64 3.55–3.93 10.3–11.3 11.5–12.5
ATSR-1, ATSR-2, AATSR	ERS-1[e] ERS-2[e] Envisat[e]	Along-track scanning radiometer	ESA	Sep 2001–present[b]	1.45–1.75[c] 3.55–3.85[d] 10.3–11.3 11.5-12.5
MODIS	TERRA, AQUA	Moderate resolution imaging spectrometer	NASA	Feb 2000–present[b]	3.660–3.840 3.929–3.989 4.020–4.080 10.780–11.280 11.770–12.270
SEVIRI	Meteosat second generation	Spinning enhanced visible and infrared imager	Eumetsat	Sep 2002–present[b]	1.50–1.78 3.48–4.36 8.30–9.10 9.80–11.80 11.00–13.00

[a] Most sensors have additional visible band(s) used for daytime cloud detection.
[b] April 2008.
[c] Day.
[d] Night.
[e] Each has a forward and a nadir view

cloud or other factors depending on the type of data. Agencies that provide access to data typically offer a search and selection Internet interface so that users can specify a location, a period of time, and possibly a filter to eliminate images that have too few cloud-free pixels.

Those users looking at larger scale phenomena and global maps are more likely to prefer level 3 or level 4 data. Invariably these will already be gridded in a latitude and longitude array. Although data dropout because of cloud may occur at level 3 it will be less problematic than at level 2, while level 4 data are analyzed products from

Table 2.6. Recent and current series of satellite microwave radiometers.

Sensor acronym	*Platform(s)*	*Agency*	*Dates (for series)*	*Channels*		*Main data products*
				Center frequency (MHz)	*Polarization*	
SSM/I (Special Sensor Microwave Imager)	DMSP: F8, F10, F11, F13, F14, F15	U.S. Dept. of Defense (DoD)	Sep 1987–present[b]	19.35 22.235 37.0 85.5	V, H V V, H V, H	Wind speed[a] Water vapor Cloud water Rain rate Sea ice
TMI (TRMM Microwave Imager)	TRMM	NASA/ JAXA	Nov 1997–present[b]	10.7 19.4 21.3 37.0 85.5	V, H V, H H V, H V, H	SST Wind speed[a] Water vapor Cloud liquid water Rain rate
AMSR-E (Advanced Microwave Scanning Radiometer)	Aqua	JAXA/ NASA	May 2002–present[b]	6.925 10.65 18.7 23.8 36.5 89.0	V, H V, H V, H V, H V, H V, H	SST Wind speed[a] Atmospheric water vapor Cloud liquid water Rain rate Sea ice
WindSat	Coriolis	U.S. DoD	Jan 2003–present[b]	6.8 10.7 18.7 23.8 37.0	V, H FP[c] FP[c] V, H FP[c]	SST, wind speed and direction

[a] Wind speed of 10 m.
[b] September 2009.
[c] FP = fully polarimetric (see sections 8.2.6 and 8.4.7 of *MTOFS*).

several sources and should consist of completely full fields. Once a sensor, or a series of sensors, has been producing data for several years, climatologies ought to be available, as should anomaly maps (the production of anomaly datasets is explained in Section 6.2.1). Users should consider which of these they require for a particular application. The rest of this book provides examples of many ocean applications of satellite data and points towards the most appropriate dataset to use.

Oceanographers requiring satellite data for a particular application ought to approach the selection of a suitable source with care. They should consider

Table 2.7. Recent and current series of satellite altimeters.

Altimeter	*Agency*	*Dates*	*Height*	*Orbit*	*SSHA r.m.s. accuracy*
TOPEX/Poseidon	NASA/CNES	1992–2005	1,336 km	9.92 day repeat non–Sun-synchronous	2–3 cm
Poseidon-2 on Jason-1	NASA/CNES	2001–present	1,336 km	9.92 day repeat non–Sun-synchronous	~2 cm
Poseidon-3 on Jason-2	NOAA/ NASA/CNES Eumetsat	June 2008–present	1,336 km	9.92 day repeat non–Sun-synchronous	~2 cm
Radar altimeter on ERS-1	ESA	1991–2000	780 km	3 & 35 day repeat Sun-synchronous(RA)	~5–6 cm
RA on ERS-2	ESA	1995–2003	780 km	35 day repeat Sun-synchronous	~5–6 cm
RA2 on Envisat	ESA	2002–present	800 km	35 day repeat Sun-synchronous	3 cm
Geosat	U.S. Navy	1986–1989	800 km	17.05 day repeat Sun-synchronous	10 cm reanalysis
Geosat Follow-on	U.S. Navy	2000–present	880 km	17.05 day repeat Sun-synchronous	~10 cm

- The provider of the data. Is this a recognized agency with a reputation for high standards of quality and a professional reputation to maintain?
- The completeness of the data provided. Are there error estimates and/or confidence flags attached? If the datasets are not in a geographically gridded form can the latitude and longitude of every pixel be readily identified?
- Is the data format readily accessible? Normally the attachment of ancillary information on accuracy, location, etc. in addition to the primary data files requires the use of a structured but flexible format such as NetCDF or HDF for which data viewers are widely available. If the format is a proprietary one used by a particular agency then are data viewers available?
- Is there adequate information to be able to interpret the data? For example, are the units and the scale of the data clear—these should be included among the attributes in a NetCDF or HDF file.
- Is the process for generating ocean data products transparent? This means that it should be clear what algorithm or procedure has been used to generate the products. Ideally there should be a link with an algorithm theoretical basis document (ATBD) for each product. For level 3 data the rules for constructing the composite should be available, and for level 4 there should be a clear

Table 2.8. Recent and current satellite synthetic aperture radars.

Sensor	*ERS-1 SAR*	*Radarsat*	*ERS-2 SAR*	*Envisat ASAR*	*Radarsat-2*
Agency	ESA	CSA	ESA	ESA	CSA
Altitude, km	780	800	780	700	800
Radar band	C	C	C	C	C
Polarization	VV	HH	VV	HH, VV	Multiple
Wavelength, cm	5.7	5.7	5.7	5.7	5.5
Incidence angle, deg	23°	20–50°	23°	17–50°	10–60°
Resolution, m	25	10–100	25	25–1,000	3–100
Swath, km	100	10–500	100	100–400	20–500
Period of operation	Jul 1991–Mar 2000	Nov 1995–present	Jul 1995–present	May 2002–present	Jan 2008–present

Table 2.9. Recent and current satellite scatterometers measuring wind speed and direction.

Sensor	AMI	NSCAT	SeaWinds	ASCAT
Agency	ESA	NASA/JAXA	NASA/JAXA	ESA/Eumetsat
Satellite(s)	ERS-1, ERS-2	ADEOS-1	QuikScat, Midori-2	METOP
Period of operation	Sep 1991–Jun 2003	Aug 1996–Jun 1997	Jun 1999–present Apr–Oct 2003	Oct 2006–present
Altitude, km	780	805	803	837
Radar band	C	K_u	K_u	C
Polarization	VV	VV and HH	VV and HH	VV
Frequency, GHz	5.3	13.995	13.4	5.255
Mode of operation	3 beams on one side	3 beams on both sides	Twin rotating beams	3 beams on both sides
Swath, km	500 km	2×600 km separated by 400 km	1,800 km includes 2 prime swaths of 450 km	2×500 km
Resolution, km	45 km	25–50 km	25 km	25 km

Table 2.10. Access to useful sources of satellite-derived ocean data products from the Web, in binary digital form.

Ocean data product	*Agency*	*Internet URL*	*For access to data select*
Sea surface temperature			
From several sensors including AVHRR, GOES	NASA-JPL	*http://podaac.jpl.nasa.gov*	Tools & Services/ftp Tools & Services/POET
Various SST products in standard GHRSST form	GDAC at JPL Archive at NODC	*http://ghrsst.jpl.nasa.gov/data_access.html* http://ghrsst.nodc.noaa.gov/accessdata.html	Select "Data"
From AMSR-E, TMI microwave radiometry	REMSS	*http://www.ssmi.com/*	Select AMSR *or* TMI Browse Data. Then select FTP or Download
Pathfinder (from AVHRR)	NOAA	*http://www.nodc.noaa.gov/sog/pathfinder4km/*	Select "Available Data"
Ocean color and related products			
MERIS	Global gridded	*http://envisat.esa.int/level3/meris/*	Select year from products table
MODIS, SeaWiFS, etc.	Global 4 and 9 km	*http://oceancolor.gsfc.nasa.gov/*	Level 3 Browser
MODIS, SeaWiFs etc	Local full resolution	*http://oceancolor.gsfc.nasa.gov/*	Level 1 & 2 Browser

		Altimetry: surface height and related products	
SSHA Along track	AVISO–DUACS	*http://las.aviso.oceanobs.com/las/servlets/dataset*	
SSHA Gridded	AVISO–DUACS	*http://www.aviso.oceanobs.com/en/data/data-access-services/index.html*	
Merged using DUACS	AVISO–DUACS	*http://atoll-motu.aviso.oceanobs.com/*	
SSHA	PO-DAAC (JPL)	*http://podaac.jpl.nasa.gov*	Tools & Services/ftp
		Winds over the sea	
SSMI, Q-Scat, AMSR, TMI	REMSS	*http://www.ssmi.com/*	Select required sensor
http://www.osi-saf.org/	PO-DAAC (JPL)	*http://podaac.jpl.nasa.gov*	Tools & Services / ftp
Various	Ifremer-Cersat	*http://www.ifremer.fr/cersat/en/index.htm*	Select "Data", then "Download"
Various	Eumetsat O&SI-SAF	*http://www.osi-saf.org/*	

Table 2.11. Access to useful image data-viewing and manipulation tools.

Name	*Description*	*Sponsor*	*Policy on availability*	*Website URL*
BILKO	Bilko is a complete system for learning and teaching remote-sensing image analysis skills. Current lessons teach the application of remote sensing to oceanography and coastal management. PC-based image analysis software supports the main image data formats	UNESCO	Free to all who register their email address	*http://www.unesco.bilko.org*
BEAM	A software system for viewing, analyzing, and processing remote-sensing data. Originally named as the Basic ERS & Envisat (A)ATSR and MERIS toolbox, BEAM now supports a growing number of other sensors such as MODIS, AVNIR, PRISM, and CHRIS/Proba. Available for several platforms.	ESA/ Brockmann Consult	Free access under the terms of the GNU General Public License	*http://141.4.215.13/index.html*
BEST	BEST (Basic Envisat SAR Toolbox) is a collection of executable software tools designed to handle ESA data products from both the Envisat ASAR instrument and the AMIs (Active Microwave Instruments) on ERS 1&2. The latest version 4.2.0 solves the problem in handling ERS PGS products both in CEOS and in Envisat format, and also has ASAR WSS new functionalities	ESA		*http://envisat.esa.int/resources/softwaretools/*

specification of the different sources of data, with explanations of the rules for prioritizing them, bias adjustments, interpolation procedures, etc.

- Is the version of the dataset unambiguous? Some datasets differ between the near–real time product, which may lack some ancillary inputs, and a ''consolidated'' product that has been reprocessed once all the necessary information for geolocation and corrections has been acquired. In subsequent years the scientific and technical understanding of the processes may improve so that newer versions of the algorithm are introduced, leading to updated, more reliable, products. Similarly level 3 and 4 products may also be reprocessed or reanalyzed, using the reprocessed level 2 as input. It is vital for some applications that the user knows which version of data is being used.
- Are the products credible? Has there been adequate calibration of the sensor and the processes used to generate the products? Have the products themselves been validated against independent observations of the same ocean variable? Are there peer-reviewed publications which define the sensor calibration and processing algorithms and report the validation of the products?

This amounts to quite an extensive list of issues that need to be considered. Obviously, common sense should be applied. The casual user is keen to have a look at some ocean image data and will be most influenced by how easy it is to gain access to the data. In contrast the climate scientist, whose own reputation for careful analysis of climate trends is at stake, must be able to critically assess the quality of the data and justify her or his assessment by reference to peer-reviewed publications. All who use any type of data for scientific purposes need to consider these issues before committing their work to a dependence on data provided by a third party. They owe it to themselves to understand how the products they use have been generated so that they can ascertain whether the exciting ''discoveries'' from their analysis are genuine environmental phenomena or disappointingly an artifact of the data-processing chain! Moreover it is the user's responsibility to keep track of product versions so as to avoid the embarrassment of publishing a paper about a sudden change in an ocean variable that turns out to correspond to a version switch of data products!

As a guide, Table 2.10 identifies by data product type some producers to whom users can look for data that should match up to many of the requirements listed above. Although the stated web addresses may go out of date, the agencies themselves should still be traceable for several years from the publication of this book. Most of these agencies provide data through the Internet at no cost to scientific users. In some cases the data are downloadable instantly. Others allow users to select subsets of data from an archive which are subsequently placed in an ftp site for download. Some agencies require users to register before allowing access to the data, so that they can differentiate between commercial usage that is charged and private scientific use that is not.

For the student new to satellite data on the web, it is worth pointing out the distinction between image datasets and digital pictures. A digital image dataset stores the data in a format that allows a user with suitable software to readily

access the measured ocean properties as digital values that can be extracted and analyzed. The file format often has provision for ancillary data to be attached to the file with the primary image data. It also allows the image to be displayed and enhanced without losing track of actual scientific values. Examples of such formats are *.hdf* and *.cdf* files, or proprietary formats such as ESA's *.N1* data files. In contrast a digital picture file format such as *.jpg*, *.png*, *.gif*, *.bmp* contains digital values allowing a picture of the image data to be reproduced, but in general there is no way of reaching the true scientific values behind the picture. Such pictures may be satisfactory for illustrating a particular phenomenon, in which case they must be printed with the correct color scale, but cannot form the basis of any scientific analysis or manipulation. Some of the agencies that provide ocean data products have websites where images can be browsed at low or full resolution, and allow download by clicking on the pictures. However, this typically saves the files as a picture format. Scientific users need to explicitly download a *.cdf* or *.hdf* version if they wish to do more than simply copy the picture.

There are now a number of useful image display and image-processing software systems available for enhancing image data and performing more elaborate analytical procedures. Table 2.11 lists some of these. The BILKO system is a general purpose image analysis system for PCs which was developed by UNESCO as a basis for training in satellite data analysis. It is freely available and comes with tutorials that allow the user to discover its capabilities through examples of analyzing a variety of typical satellite ocean data products. It can handle standard image data formats like NetCDF and HDF, and has been extended to read ESA's N1 file formats. The BEAM software has been developed specifically to allow scientific manipulation of ESA's data products. Other agencies are also starting to provide software tools for selecting, enhancing, and manipulating their data products, sometimes online before the products are downloaded.

2.7 REFERENCES

Barton, I. J. (1995), Satellite-derived sea surface temperatures: Current status. *J. Geophys. Res.*, **100**, 8777–8790.

Chelton, D. B., J. C. Ries, B. J. Haines, L.-L. Fu, and P. S. Callahan (2001), Satellite altimetry. In: L.-L. Fu and A. Cazenave (Eds.), *Satellite Altimetry and Earth Sciences* (pp. 1–131). Academic Press, San Diego.

Cotton, P. D., and D. J. T. Carter (1994), Cross calibration of TOPEX ERS-1 and Geosat wave heights. *J. Geophys. Res.*, **99**, 25025–25033.

Donlon, C. J., I. S. Robinson, K. S. Casey, J. Vazquez, E. Armstrong, O. Arino, C. L. Gentemann, D. May, P. Le Borgne, J.-F. Piollé, and 16 others (2007) The Global Ocean Data Assimilation Experiment (GODAE) High Resolution Sea Surface Temperature Pilot Project (GHRSST-PP), *Bull. Am. Meteorol. Soc.*, **88**(8), 1197–1213, doi: 10.1175/BAMS-88-8-1197.

Gower, J. F. R. (1996), Intercalibration of wave and winds data from TOPEX/Poseidon and moored buoys off the west coast of Canada. *J. Geophys. Res.*, **101**, 3817–3829.

Kilpatrick, K. A., G. P. Podesta, and R. Evans (2001), Overview of the NOAA/NASA advanced very high resolution Pathfinder algorithm for sea surface temperature and associated matchup database. *J. Geophys. Res.*, **106**(C5), 9179–9197.

Le Traon, P.-Y., and F. Ogor (1998), ERS-1/2 orbit improvement using Topex/Poseidon: The 2 cm challenge. *J Geophys. Res.*, **103**(C4), 8045–8057.

Le Traon, P.-Y., P. Gaspar, F. Bouyssel, and H. Makhmara (1995), Using Topex/Poseidon data to enhance ERS-1 orbit. *J. Atmos. Oceanic Tech.*, **12**, 161–170.

Merchant, C. J., and A. R. Harris (1999), Toward the elimination of bias in satellite retrievals of sea surface temperature: 2, Comparison with in situ measurements. *J. Geophys. Res.*, **104**(C10), 23579–23590.

Merchant, C. J., A. R. Harris, M. J. Murray, and A. M. Závody (1999), Toward the elimination of bias in satellite retrievals of sea surface temperature: 1, Theory, modelling and interalgorithm comparison. *J. Geophys. Res.*, **104**(C10), 23565–23578.

Reynolds, R. W. (1993), Impact of Mt. Pinatubo aerosols on satellite-derived sea surface temperatures. *J. Climate*, **6**, 768–774.

Reynolds, R. W., and T. S. Smith (1994), Improved global sea surface temperature analyses. *J. Climate*, **7**, 928–948.

Rio, M.-H., and F. Hernandez (2004), A mean dynamic topography computed over the world ocean from altimetry, in situ measurements, and a geoid model. *J Geophys. Res.*, **109**(C12032), doi: 10.1029/2003JC002226.

Robinson, I. S. (2004), *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Saunders, R. W., and K. T. Kriebel (1988), An improved method for detecting clear sky radiances from AVHRR data. *Int. J. Remote Sensing*, **9**, 123–150, 1393–1394.

Vazquez, J., K. Perry, and K. A. Kilpatrick (1998), *NOAA/NASA AVHRR Oceans Pathfinder Sea Surface Temperature Data Set: User's Reference Manual* (Version 4.0). Jet Propulsion Laboratory, Pasadena, CA.

Walton, C. C., W. G. Pichel, J. F. Sapper, and D. A. May (1998), The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. *J. Geophys. Res.*, **103**(C12), 27999–28012.

Wentz, F. J. (1997), A well calibrated ocean algorithm for special sensor microwave/imager. *J. Geophys. Res.*, **102**, 8703–8718.

3

Mesoscale ocean features: Eddies

3.1 DISCOVERING MESOSCALE VARIABILITY FROM SPACE

One of the earliest impacts of satellite data in mainstream oceanography, persuading marine scientists nearly four decades ago that here was a potentially useful new tool, was to reveal the ubiquitous occurrence of mesoscale variability in the open ocean. The first sight of turbulent eddy-like patterns in the images from early infrared scanners, or in photographs from manned spacecraft, made sense of the perturbations at frequencies of days to weeks that had been recorded by long-term moorings. It confirmed that the apparently random variability in the surface temperature recorded by ships along some transect lines was a real phenomenon, part of coherent spatial structures, and not simply a manifestation of instrument noise.

Long before satellite remote sensing of the ocean became the precise measurement technique it is today, pictures such as those in Figure 3.1 helped to transform the perception of physical oceanographers. By displaying qualitatively the meanders of major ocean fronts such as the Gulf Stream, they must surely have helped to stimulate the research effort of the 1970s and 1980s towards measuring mesoscale variability using conventional instruments from ships. In the 1980s and 1990s satellite data, from infrared and visible imagers and from altimeters, became supplementary measurement tools used by physical oceanographers to improve their understanding of mesoscale dynamics. Now they have become an almost essential element of monitoring aspects of mesoscale variability. By capturing a synoptic view of the ocean, satellite images can readily provide spatial data about the extent, the shape, and the variability in lengthscales of certain ocean processes, information that is otherwise hard to obtain from conventional oceanographic experiments.

This chapter and the two that follow it take a broad look at the variety of mesoscale ocean phenomena that can readily be seen by satellite sensors. They describe the various methods of remote sensing used for observing ocean features at lengthscales between 10 km and 1,000 km. Their underlying aim is to demonstrate

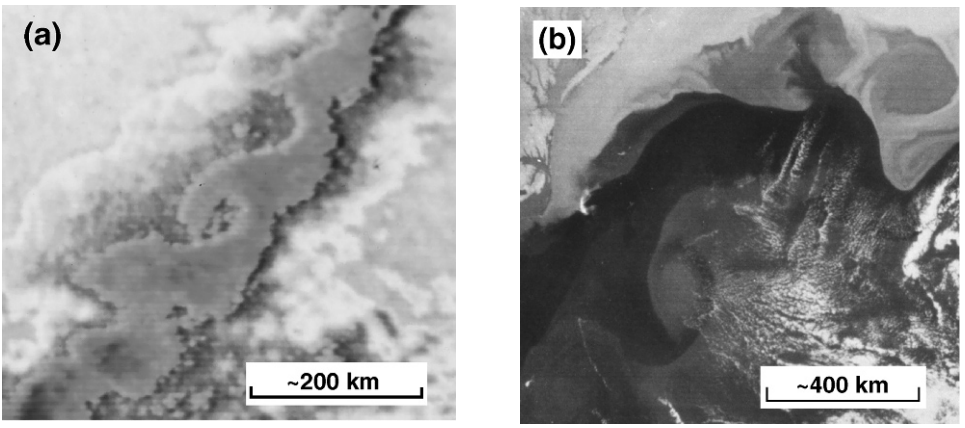


Figure 3.1. Early infrared images of the ocean. (a) Gray-tone rendering of a color-enhanced TIROS television picture of 1968, showing Gulf Stream eddies off South Carolina. (b) Gray-tone infrared image (warmer is darker) acquired by the Very High Resolution Radiometer on the NOAA-3 satellite, on April 28, 1974, showing the Gulf Stream off North Carolina (based on images from the NOAA Photo Library at (a) *http://www.photolib.noaa.gov/htmls/spac0088.htm* and (b) *http://www.photolib.noaa.gov/htmls/spac0301.htm*).

the unique role satellites have played in increasing our scientific understanding of mesoscale variability throughout the ocean.

Figure 3.2 illustrates the type of mesoscale features which are the subject of Chapters 3–5. It is a sea surface temperature image derived from infrared radiometry and shows the ocean surrounding South Africa. This is a region where some important ocean currents meet and interact. The warm Agulhas Current flows southward down the east coast of Africa and starts to follow the coast round to the west. At about 22°E it appears in this image to bend back on itself, towards south and east in what is called the Agulhas retroflection, although it is known that some of the warm water does in fact manage to pass round Cape of Good Hope into the Atlantic where it contributes to the northward-flowing Benguela Current. The edge of this reflected, warm current shows up clearly as a strong, surface thermal front, especially to the east of the retroflection region, where it meanders to form warm and cold core eddies. Meanwhile south of about 40°S part of the Antarctic Circumpolar Current (ACC) flows eastward along the subpolar front which also meanders to produce a complex of mesoscale eddies. Despite the patches of cloud which persist through the 8 days during which this composite of several satellite datasets was created, the patterns of these eddies have been captured quite well by the image. On the west coast of Africa, the cold water close inshore is evidence of upwelling.

Thus in a single image can be seen eddies, fronts, and upwelling, which are the phenomena that are discussed separately in Chapters 3, 4, and 5, respectively. However, Figure 3.2 is a reminder that in practice these three phenomena are often dynamically interwoven. Fronts are rarely stationary and their deformations

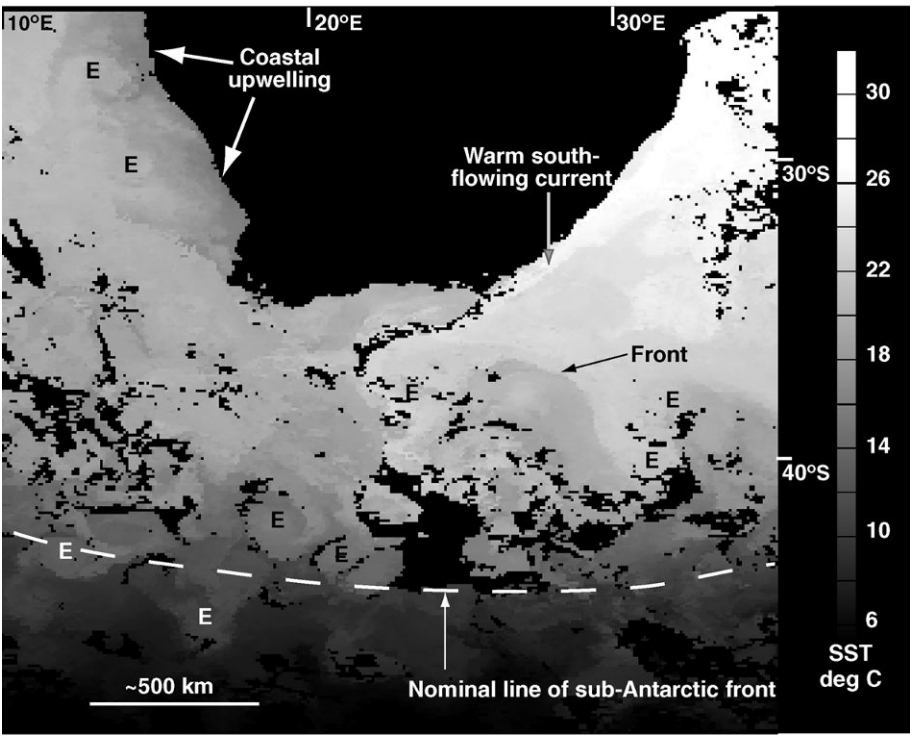


Figure 3.2. Sea surface temperature measured by the AVHRR sensor (8-day average Pathfinder v4.1 product at 9 km resolution) on March 15–22, 2001. Black gaps are land or cloud. "E" denotes a probable eddy.

can grow and pinch off to form isolated eddies. Upwelling events can create sharp fronts between the cold upwelled water and the warmer surface water it is replacing. Upwelling fronts may deform into offshore jets which wind up to form eddies. Oceanographers now take such turbulent processes and phenomena for granted, yet 30 years ago very little was known about them. Looking back it seems clear that satellite images, capturing snapshots of continuously evolving spatial structures, have often provided the key to unlock our understanding of these important phenomena.

These three chapters will probe beyond qualitative pictures to present the systematic analytical methods by which remote-sensing techniques can measure aspects of mesoscale dynamics at lengthscales of 10 km to 1,000 km in the ocean. While infrared images are the starting point, it is interesting to see how views of the same phenomenon from other types of sensor can enrich the information available from satellites. Chapter 5 will also mention additional mesoscale ocean features, other than fronts, eddies, and upwelling, that can be observed by satellite oceanography, such as wind-driven or island-induced mixing.

3.2 MESOSCALE OCEAN EDDIES

3.2.1 Eddies—ubiquitous phenomena in a turbulent ocean

Figure 3.3 is a rather spectacular example of mesoscale variability found in the southwest Indian Ocean. It was observed by the MODIS sensor which revealed the phenomenon in both the SST image from the infrared channels and the chlorophyll image derived from the visible channels. Conveniently located in a clear area between two regions of cloud, there is a large elliptical eddy which, when set in its wider context, can be recognized as part of the intense eddy activity that occurs in the Antarctic Circumpolar Current.

The temperature at the center of this cold core eddy is about 14°C and the warmest temperature to the north of the eddy is about 19°C. The chlorophyll concentration at the center is about $0.5\,\mathrm{mg\,m^{-3}}$, surrounded by even lower concentrations of $0.3\,\mathrm{mg\,m^{-3}}$ in a horseshoe shape of diameter 30 km to 40 km. The concentration outside the eddy is about $0.8\,\mathrm{mg\,m^{-3}}$, while the band of light tone southeast of the eddy corresponds to about $1.5\,\mathrm{mg\,m^{-3}}$.

What stands out most impressively in this particular example is the ring of smaller scale shear tongues forming into small eddies around the circumference of the main eddy, in a regular and almost symmetrical pattern. These appear most clearly in the SST image but are also evident in the chlorophyll image. Here is a graphic example where only satellite data are capable of revealing such spatial detail; *in situ* observations could never be sampled at sufficient spatial frequency in a short enough time interval to capture a synoptic view of this convoluted and beautiful feature.

Such dynamical complexity comes as no surprise to fluid dynamicists. After all, at the lengthscale of one to thousands of kilometers, flows in the ocean have a very high Reynolds number[1] which implies that they are also very turbulent. Mesoscale variability is a manifestation of that turbulence. What is fascinating to discover in images such as Figure 3.3 is that the flows resulting from turbulent energy are not always random but can be constrained, at least temporarily, into coherent patterns. It is instructive to consider why that should be the case at certain lengthscales.

3.2.2 Lengthscales of mesoscale eddies—the Rossby radius

The lengthscales of variability evident in Figure 3.3 span from a minimum of about 10 km to 15 km corresponding to the small shear tongues on the main eddy circumference or the width of the streaks, to a maximum of about 150 km which is the east–west size of the main eddy. It is instructive to ask what determines the size of the patterns in images like this? What are the forces that preserve the order and regularity of the flows evident in Figure 3.3 and control the horizontal scale of

[1] The Reynolds number is a dimensionless quantity, $R_e = VL/v$—where V is the characteristic velocity scale of the flow; L is the lengthscale; v is kinematic viscosity; and R_e represents the ratio between inertial forces and viscous damping forces in fluid flow.

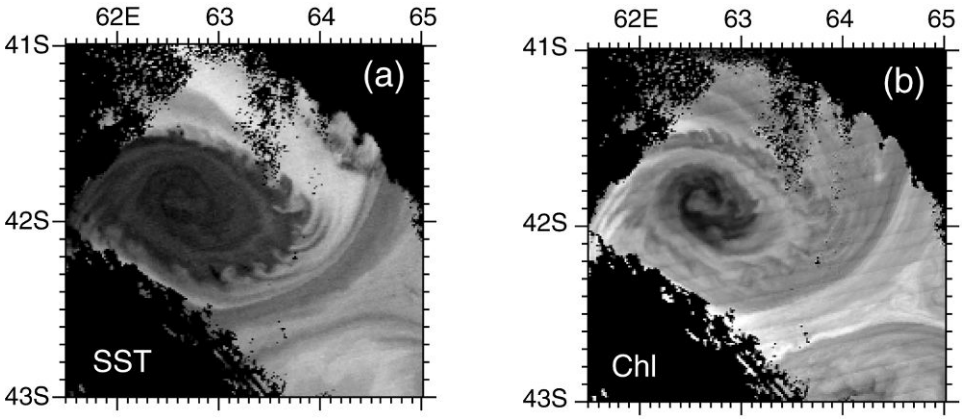


Figure 3.3. Extracts from images of data products derived from MODIS on the Terra satellite, January 27, 2005. (a) Sea surface temperature (the darker tones are cooler in this image). (b) Chlorophyll (dark is lower concentration) (original image data from NASA Ocean Color website.)

mesoscale variability? The answers to these questions are crucial for fully understanding satellite observations of mesoscale phenomena.

We have noted that the sea, like the atmosphere, is naturally turbulent at all but the smallest lengthscales, but the mesoscale turbulence which can produce these beautiful flow patterns is constrained by geostrophy (to learn more about this see, for example, Vallis, 2006; Stewart, 2008). We make the assumption that, in order to maintain this quasi-stable flow, there must be a balance between the horizontal pressure gradients created by a sloping sea surface and/or density gradients in the water, and the Coriolis force. The Coriolis force occurs because the ocean rotates with the Earth: it acts at 90° to the direction of flow, to the right in the northern hemisphere and to the left in the south, and is proportional to flow speed. Figure 3.4

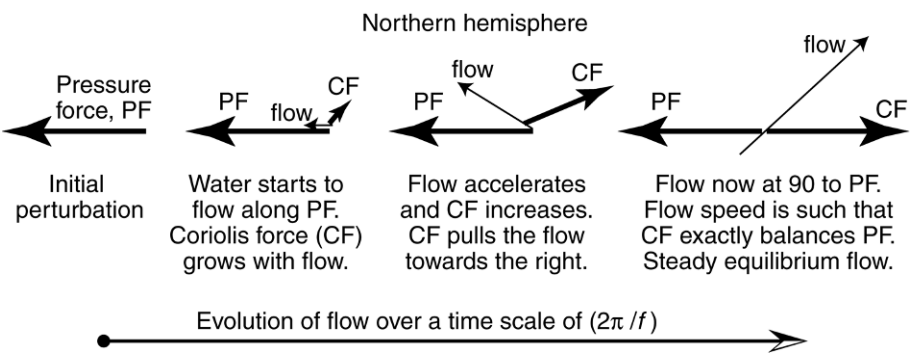


Figure 3.4. Spinup to geostrophic equilibrium from the onset of a pressure gradient force.

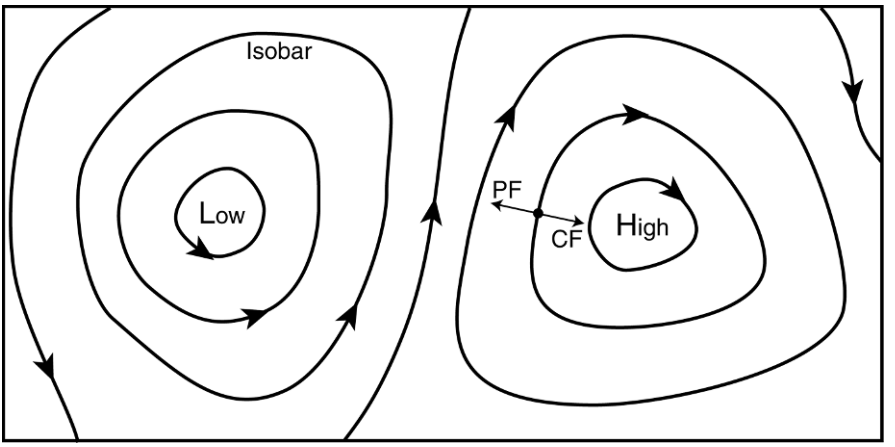


Figure 3.5. Part of a geostrophic flow field with steady flow along isobars (northern hemisphere).

illustrates how, after the onset of a pressure force in the ocean, the flow develops in response until it reaches geostrophic equilibrium.

In this state the flow is steady and at right angles to the pressure gradient, parallel to the contours of constant pressure, called isobars. Thus water can circulate steadily around centers of high and low pressure without changing mass distribution (as shown in Figure 3.5). If there were no Earth rotation effect, the water would flow from high-pressure zones to low-pressure ones, thus leveling out the pressure and mean flow would tend to zero. In contrast, on the rotating Earth the geostrophic flow patterns can be sustained over many days and they change only gradually in response to other factors in the wider flow field.

If geostrophic balance is the mechanism controlling the spatial patterns of mesoscale variability in the ocean, what determines their size? To answer this we must first consider the ocean's gravitational response to an initial disturbance (e.g., a storm with winds which blow for a time and then decline again). Wind stress tilts the sea surface, but when the wind declines, the sloping surface is left with nothing to balance the gravity force. This drives the water down the slope, tending to level out the surface again. In fact, because water gains momentum in the process, the surface does not simply flatten out but a wavelike motion is created. This propagates at the speed of barotropic long waves (i.e., $\sqrt{(gh)}$ in water of depth h). In an ocean of depth 4 km this is a speed of 200 m s^{-1}, meaning that within a few hours any disturbance of the sea surface travels thousands of kilometers. However, it is not perturbations of the sea surface that are relevant here, but disturbances of the layered density structure in the upper ocean, and the resulting baroclinic response. At its simplest we can think of the ocean as two layers, a less dense mixed layer of depth h_1 floating on top of a slightly denser deep layer below the thermocline, as shown in Figure 3.6. The wind event described above would also cause the boundary between the upper and lower layer to slope, and this perturbation would propagate

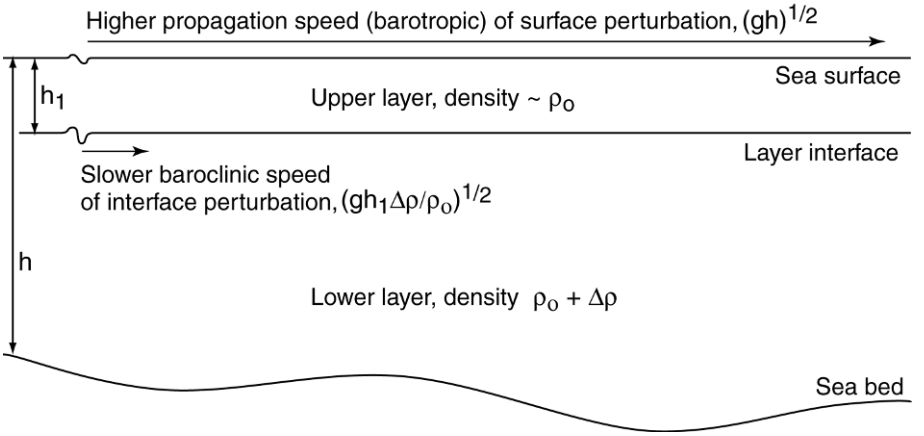


Figure 3.6. A simplified two-layer ocean and the difference between barotropic and baroclinic gravity wave speeds.

at the baroclinic wave speed. If the density contrast between the layers is $\Delta\rho$ and the mean density is ρ_0 then the baroclinic wave speed is $\sqrt{(gh_1\Delta\rho/\rho_0)}$, very much slower than the barotropic speed. With h_1 about 20 m to 50 m and $\Delta\rho/\rho_0$ about 0.001 m to 0.01 m, the speed is typically a few kilometers per hour.

Baroclinic flows are much slower and gentler than barotropic flows because of the effectively reduced gravity force. Thus a baroclinic disturbance does not propagate far during the time it takes for the effect of Earth rotation to modify the baroclinic response to the perturbation. The natural timescale of the Coriolis effect is the order of one pendulum day (i.e., f^{-1}—where $f = 2\Omega \sin\phi$ is the Coriolis parameter; Ω is the Earth rotation rate; and ϕ is the latitude). This characterizes the time it takes for the Coriolis force to redirect the flow caused by the perturbation so as to reach a geostrophic balance flowing along the contours of the sloping surface instead of down the slope (as in Figure 3.4). This allows pressure perturbations to be trapped instead of propagating rapidly away, and it is these trapped perturbations that show up as mesoscale eddies and perturbations of frontal structures. There must be a minimum size for these structures because it takes a period of time for the Earth rotation to control them, and during that period they are propagating as gravity waves. Thus the lengthscale is determined as the speed of baroclinic gravity wave propagation multiplied by the half-pendulum day. This lengthscale is called the baroclinic Rossby radius of deformation, L_{Rb}, and is defined as

$$L_{Rb} = \frac{\sqrt{gh_1\Delta\rho/\rho_0}}{f}. \tag{3.1}$$

The actual value of this varies with the depth of the mixed layer and the density contrast across the thermocline, but it varies typically from about 10 km at high

latitudes to more than 80 km in the tropics (Chelton *et al.*, 1998). Close to the Equator it is larger still where geostrophy does not constrain the flow.

Any dynamical structures that are smaller than the Rossby radius cannot be in geostrophic balance. They will either evolve rapidly into larger structures that are in equilibrium or else propagate away; in either case such small-scale patterns are ephemeral. Patterns of flow in the ocean with lengthscales equal to or greater than the Rossby radius, such as those evident in the images already shown in this chapter, are expected to be in geostrophic balance and therefore stable and subject to only gradual change.

Consequently turbulent energy tends to be trapped in the mesoscale, bounded at the low end by a lengthscale of L_{Rb} and a timescale of f^{-1}. Structures much larger than L_{Rb} with longer timescales are also in geostrophic balance. The mesoscale is bounded at the upper end by the fact that f varies with latitude, and so large structures experience variation of the magnitude of the Coriolis force across their north–south extent. This is sometimes referred to as the β effect because the latitudinal variation of f is often approximated to $f = \beta\phi$. This introduces additional complexity into the geostrophic response of the ocean to large-scale perturbations, leading to a different class of large-scale dynamical phenomena, which are also observable from satellites (they are discussed in Chapter 6).

3.2.3 The dynamical structure of rings and eddies

We may conclude from the above subsection that much of the ocean's turbulent energy occurs in the mesoscale part of the size spectrum, at lengthscales of tens to hundreds of kilometers and timescales from days to weeks. This is confirmed by the satellite images of various types presented throughout this chapter. If followed through a time sequence, some images paint a picture of an ever-changing gradual folding and unfolding of mesoscale patterns of turbulence. However, there are some very well–defined eddies and rings which appear to be more permanent than most of the turbulent eddies. It is instructive to examine the basic dynamics of these in order to interpret their remote-sensing signatures which are discussed in the rest of the chapter. Because satellites tend to see only the surface signatures of eddies it is important to be able to relate this to what may be going on at depth.

Figure 3.7 shows a schematic diagram of vertical slices through the diameters of two typical eddies, one cyclonic and the other anticyclonic, drawn for the northern hemisphere (the sense of rotation would reverse in the southern hemisphere but everything else would be the same). The basic mechanism is geostrophic balance, at all depths, between the horizontal pressure gradient (indicated by the thick straight arrows) and the horizontal circular flow producing a geostrophic force radially outwards from the core of the anticyclonic eddy (inwards towards the core of the cyclonic eddy). Note that in this very simple conceptual model we are ignoring the inertial (centrifugal) forces associated with flow around curved streamlines. The pressure gradient is maintained by the sea surface slope which rises (falls) towards the eddy center. If the ocean were not stratified, the same horizontal pressure gradient would exist all the way to the sea bed, and the eddy would be a spinning

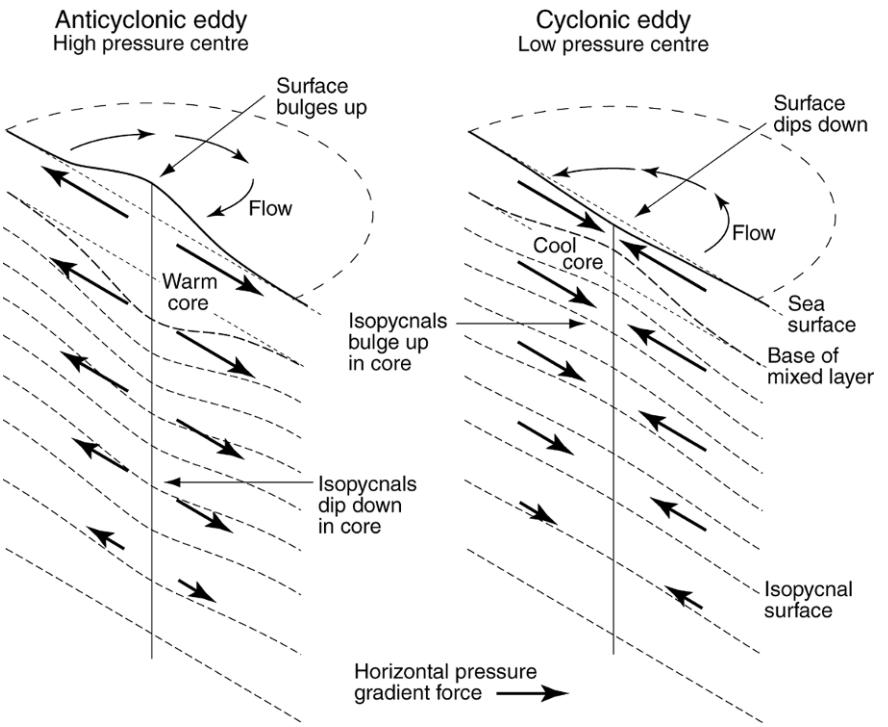


Figure 3.7. Schematic of the basic structure of a simple eddy in the northern hemisphere. On the left is an anticyclonic eddy (warm core) and on the right is a cyclonic eddy (cold core).

column of water extending through the whole depth of the sea. However, in a typical ocean situation density stratification is also perturbed by the eddy in such a way that the horizontal pressure gradient reduces with depth down to a level at which no eddy motion is present.

As Figure 3.7 shows, in a stable eddy this requires surfaces of equal density (isopycnals) to dip down in the center of an anticyclonic eddy and upwards in a cyclonic eddy. With increasing depth, the pressure gradient in one direction caused by the surface slope is increasingly offset by the pressure gradient in the opposite direction associated with the horizontal density gradient. The balancing geostrophic horizontal flow reduces with depth—what meteorologists call the thermal wind effect. The horizontal density gradient is mostly associated with a horizontal temperature gradient. Temperature must increase towards the center of an anticyclonic eddy which has a warm core. At the stage when an embryonic eddy is first spinning up there is downwelling at the center which draws down the isotherms, so warmer water penetrates deeper than normal. When the eddy is established, warm water is maintained dynamically in this position. Conversely in the center of a cyclonic eddy there is cooler, denser water than normal at a given depth.

While this describes a framework for understanding the primary geostrophic balance achieved in an ocean eddy, there are other factors which determine how

an eddy changes over time. Secondary nonlinear dynamical processes such as advective or centripetal accelerations allow perturbations from the simple, circular, symmetrical flow to grow as instabilities, changing gradual density gradients into steep fronts and producing secondary vertical circulations which redistribute the density structure until the eddy is transformed into a completely different feature. The simple model of an isolated eddy also breaks down when two eddies interact, causing a redistribution of vorticity which can create filaments with strong density gradients and horizontal shear. The secondary flow in eddies is important because the associated upwelling and downwelling promotes vertical mixing, raising nutrients to the sea surface, and helping to ventilate the deeper ocean. A full dynamical treatment of ocean eddies can be found in ocean dynamics textbooks (Vallis, 2006; Stewart, 2008). The emphasis here is on how the eddies and their complex evolution over time can be observed from satellites. The idealized structure illustrated in Figure 3.7 provides an adequate basis for discussing the primary remote-sensing signatures of ocean eddies.

3.3 DETECTING EDDIES FROM SATELLITES

In common with all other applications of satellite oceanography, the observation of mesoscale eddies requires that they produce a surface signature in the field of the primary quantity being observed by a particular sensor. What fundamentally defines the eddies is the curved, and sometimes closed, streamlines of the horizontal velocity field in the upper layers of the ocean. Ideally we would like to detect the flow field directly, but this is not yet reliably achievable with satellite remote sensing. However, surface geostrophic velocity can be inferred from the field of sea surface height anomaly (SSHA) observed by radar altimeters because the small displacement of sea surface elevation associated with mesoscale eddies is detected directly by altimetry. This provides the most reliable way of observing eddies from space; even when the dynamical core of the eddy is well below the surface it is still expected to have a signature in surface height. Section 3.4 will discuss in more detail how eddies are revealed in the altimeter SSHA record, and what characteristics of eddies can be measured in this way.

As already demonstrated in Figures 3.1 to 3.3, eddies can also be observed in satellite-measured sea surface temperature (SST). This is not surprising, since the simple model of Figure 3.7 shows that a cyclonic (anticyclonic) eddy requires a core of denser (less dense) water, and this is normally a consequence of water that is cooler (warmer) than normal in the eddy core at each depth. Thus we may expect to see cold and warm core eddies in cloud-free SST images. However, further consideration of Figure 3.7 shows that the eddy mechanism does not require any surface density gradient as long as there are horizontal density gradients below the surface down to the depth of penetration of the eddy. While in many cases there is a corresponding change in surface layer temperature, an eddy can still exist without this.

Some types of eddy are centered at some considerable depth below the surface. For example, coherent vortices containing 200 m to 1,000 m thick cores of warm and

salty Mediterranean water are found in the North Atlantic, centered at a depth of about 1,000 m. Called "meddies" (McDowell and Rossby, 1978) because of their dependence on deep Mediterranean outflow, they are important means for transporting heat and salt through the North Atlantic (Bower *et al.*, 1997), but do not generally have a detectable signature in surface temperature.

Moreover, even when the basic structure of an eddy does produce a measurable difference in temperature of the surface mixed layer, it is possible for this to be temporarily hidden by short-term, meteorologically related events, such as diurnal warming producing a very shallow warm layer (see section 7.3.3 of *MTOFS*), or a storm producing a patch of cooler surface water from local mixed layer deepening. Thus care must be taken when inferring information about eddy fields from the evidence of the SST record. By no mean all ocean eddies have a measurable SST signature. Nonetheless, as shown in Section 3.5 and Chapters 4 and 5, there have been many studies of the character of ocean mesoscale features based on satellite SST measurements.

A third way in which eddies may be revealed by remote sensing is when observable tracers are advected by a slowly evolving eddy field and are drawn out into patterns that describe the underlying structure of the eddy. Both temperature and color are candidates as tracers for this type of eddy signature. For example, in Figure 3.3 the SST field shows spirals of warmer water being wound up into the main eddy core, serving as a tracer of the current field. Moreover, at the periphery of the main eddy in Figure 3.3, where small eddies are being spun up by current shear, it is the advected SST field which makes them clearly visible.

The color of the sea sometimes provides a tracer of mesoscale motion that can clearly be detected from space. Optically reflective material in the water, such as phytoplankton or suspended particulates, are drawn out into patterns that are evident in visible waveband image data. In Figure 3.3b it is the phytoplankton distribution, as defined by the chlorophyll concentration derived from MODIS ocean color data, that serves to trace the outline of the eddy. However, the retrieval of information about mesoscale eddy dynamics from ocean color images must also be approached with caution. Sometimes the mesoscale motion is revealed by patches of colored material caught up and advected by the flow as a passive tracer, such as the small spiral arms of the peripheral eddies in Figure 3.3b. In other cases, the eddy's color signature is associated with the different water masses which make up the inside and the outside of an eddy. For example, a cold core eddy in the North Atlantic typically provides sufficient nutrients by upwelling into the surface layer to support phytoplankton growth and enhanced levels of chlorophyll, which on a satellite image paint the eddy core greener than the blue, less productive water outside the eddy. In cases like this, it would be wrong to use ideas about horizontal advection to interpret the image; instead the higher concentration of chlorophyll can be used as evidence of local upwelling. Section 3.6 will explore further how ocean color is used to study eddies.

A fourth imaging mechanism by which eddies have been detected is their fine-scale surface roughness signature on synthetic aperture radar images (explained in Section 3.7). Although SAR images are not hindered by cloud, it turns out that the

success of this method is just as dependent on the weather as are the SST and color methods, because radar-imaging mechanisms are sensitive to the sea state. Thus, while it might seem that four independent methods for detecting eddies is something of a luxury, in fact only altimetry can be considered to provide a reliable, routine, monitoring capability for mesoscale eddies, and then only for larger scale eddies. In the following four sections each of the primary ocean quantities measured by remote sensing, height, temperature, color, and roughness is explored in relation to how it has been used to advance knowledge and understanding of ocean mesoscale turbulent flows. It will become apparent that each method reveals different aspects of the underlying phenomenon, and retrieves information at different length and timescales. Each has particular strengths and limitations, and together they build up a significant body of knowledge about ocean eddies, which would not have been possible without the use of satellite remote sensing.

3.4 USING SSHA FROM ALTIMETRY TO OBSERVE EDDIES

3.4.1 Revealing ocean eddies in altimeter SSHA data

By measuring the departure of sea surface height from its long-term mean level at that location, the sea surface height anomaly (SSHA) product obtained from satellite radar altimeters is capable of directly detecting the presence of ocean eddies. The principles of radar altimetry are outlined in Section 2.4.5, elaborated in chapter 11 of *MTOFS* and comprehensively described by Fu and Cazenave (2001). Oceanographic users of altimeters typically obtain data in the form of the along-track record at one-second intervals. The SSHA is the along-track surface height record from which have been removed the long-term mean height at each sampled point along the given orbit track, the tidal height signal estimated from tidal analysis of the long-term altimeter record, and the instantaneous effect of barometric pressure. This requires that the data have been obtained from a mature altimetric program such as TOPEX/Poseidon (T/P) plus Jason-1 and Jason-2, or ERS-1 plus ERS-2 and Envisat, for which a long-term record already exists over a precisely defined orbit track repeated by subsequent satellites in the sequence.

Figure. 3.8a shows an example of the SSHA along-track record from T/P over the Arabian Sea where the satellite track crosses two eddies called the Southern Gyre (SG) and the Great Whirl (GW). Near the Equator the orbit track is almost north–south and so this record, plotted against latitude, spans about 1,800 km. The plot shows both the one-second SSHA samples and a smoothed profile of these samples. The height of SG, as intersected by this track, is about 40 cm and of GW about 25 cm. These are large eddies with diameters of at least 300 km to 400 km, which are expected at low latitudes where the Rossby radius is large. However, a single along-track cut through an eddy does not necessarily pass through its center and so these may be underestimates of both height and diameter.

From the along-track gradient of SSHA the cross-track geostrophic velocity can be calculated using Equation (2.9). This estimates the time-variable part of the

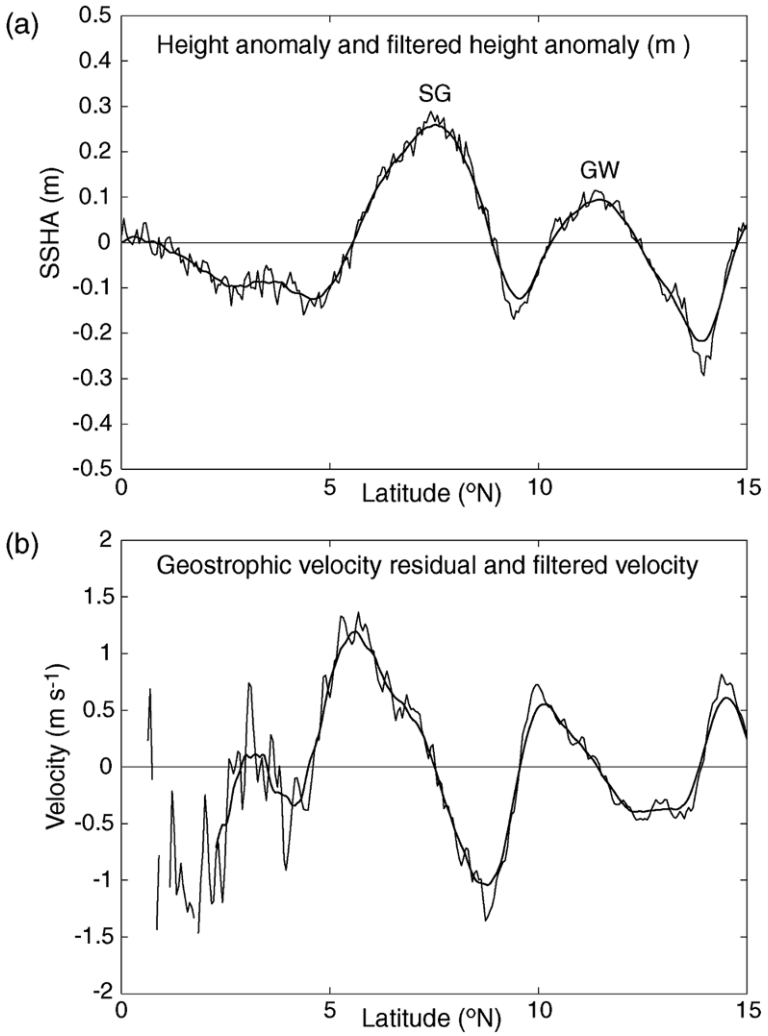


Figure 3.8. Along-track altimeter records from TOPEX/Poseidon, showing a portion of Track 233 over the Arabian Sea during Cycle 36 in August 1993. (a) Sea surface height anomaly. SG is the Southern Gyre eddy and GW the Great Whirl eddy. (b) Derived geostrophic velocity (data courtesy of B. Subrahmanyam).

velocity, relative to mean circulation. Figure 3.8b shows the velocity evaluated from the smoothed SSHA in Figure 3.8a. Because the velocity is derived from the along-track gradient of the altimeter-measured height, the smoothed version of SSHA is preferable to avoid an extremely noisy velocity profile with many reversals of direction. Since most high-frequency SSHA variability can be considered as measurement noise or resulting from wind-driven ageostrophic flows, it is reasonable to derive the geostrophic flow from the smoothed SSHA profile. Even so the resulting cross-track velocity profile also has high-frequency fluctuations. While this fine-scale velocity structure may represent actual shear flows within the eddy, its sensitivity to very small fluctuations in SSHA implies that it also contains measurement noise. More confidence can be placed in the smoothed velocity profile drawn in Figure 3.8b. It is

evident from this that the altimeter is better at detecting variability at the larger end of the mesoscale range than at the smaller end defined by the baroclinic Rossby radius. Note also in this example that velocity cannot be retrieved within one or two degrees of the equator where the Coriolis parameter f is zero and (2.9) does not apply.

Individual orbit records have an along-track spatial resolution of better than 10 km; for T/P and Jason it is 7 km. Providing a continuous record of more than 17 years, sampled every 10 days independently of weather, the T/P-Jason along-track record of SSHA has become a widely used and respected data resource for many fields of ocean dynamics research (Fu and Cazenave, 2001). However, the primary along-track SSHA record does not provide visualization of the two-dimensional horizontal field of ocean eddy activity. A simple way to do this would be to plot all the along-track data acquired during an orbit repeat cycle (~10 days for T/P–Jason) onto a two-dimensional map but they would then appear as a set of point values, closely spaced along a set of regular, intersecting tracks, with wide areas between them in which there are no observations. In order to produce a complete two-dimensional field with no gaps, the data from individual altimeter tracks are interpolated onto a grid that is typically $1/3^\circ \times 1/3^\circ$, resulting in a global image like Figure 3.9.

This was produced entirely from one complete orbit repeat cycle of Jason-1 and so it represents an integration of all the orbits over 10 days from a single sensor. At the highest latitudes (it does not extend poleward of 66°N and 66°S because of orbit inclination) the tracks overlap each other closely, justifying the map resolution of $1/3^\circ$. However, near the Equator the tracks are spaced about 250 km apart and so a number of the grid points are filled by values interpolated from samples up to 100 km distant, which must be borne in mind when interpreting these maps. In practice an

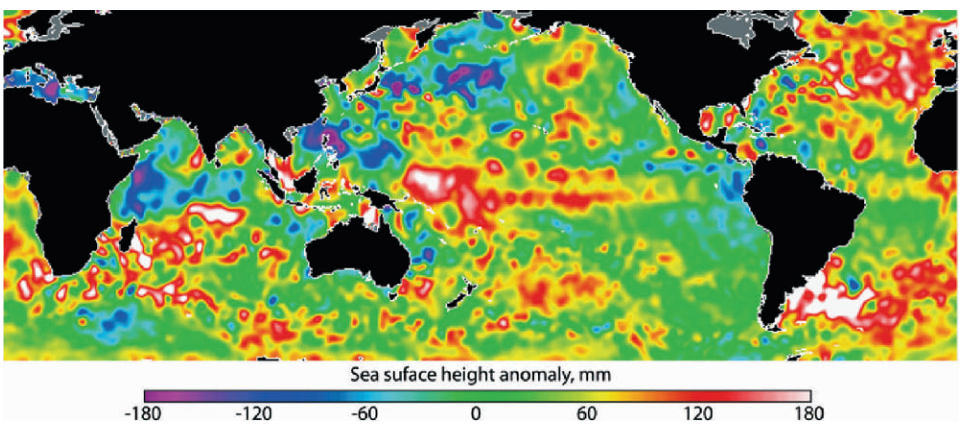


Figure 3.9. Example of the two-dimensionally smoothed SSHA field from a single orbit repeat cycle of Jason-1, between January 10 and 20, 2005 (from the PO.DAAC website at NASA Jet Propulsion Laboratory).

image like Figure 3.9 contains such a wealth of information about the dynamic variability of the ocean, at lengthscales from 100 km to 10,000 km, that the need to ignore small-scale variability of less than 100 km near the Equator is not a serious issue.

One of the attractions of gridded maps of SSHA data is that the contours of constant SSHA correspond to streamlines of the geostrophic flow anomaly. This offers us an immediate visualization of the time-variable part of the surface ocean velocity field. The direction of flow along the contours is with the higher values of SSHA to the right and lower to the left in the northern hemisphere and *vice versa* in the southern.

In Figure 3.9 we readily see the heterogeneity of eddy activity which is most evident where there are major currents like the Gulf Stream, the Kuroshio, the confluence of the Brazil and Malvinas Currents and the Antarctic Circumpolar Current, on either side of the Equator in the Pacific and Atlantic Oceans and round the south of Africa. Some parts of the ocean, such as the southeastern Pacific have very little activity. It is also apparent that the lengthscale of variability tends to be larger at low than at high latitudes. There also appear to be very large–scale perturbations of sea level with a tendency to negative anomalies in the equatorial Indian and northwest Pacific oceans and positive anomalies in the southeast Pacific and throughout the Atlantic Ocean.

If a time sequence of such images is viewed it shows how the eddies grow and decline in magnitude, change their shape and size, and some propagate. Large-scale disturbances also evolve but typically do so more slowly than smaller features. A time sequence of data over several years yields information about relationships between eddies of different scales, their dependence on the strength of ocean currents, on larger climatic signals such as El Niño, and seasonal dependences. For example, Adamec (2000) characterized eddy and mean flow behavior in the North Pacific Ocean related to the behavior of the Kuroshio Current and Fu (2007) made a similar study in the Argentine Basin of the South Atlantic.

Since 1992 there have been at least two and sometimes as many as four different altimeters in orbit at the same time. When the orbits of these are harmonized the along-track data from each of them can be used in the interpolation to achieve a much higher quality, gridded product (Le Traon *et al.*, 1998; Ducet *et al.*, 2000). Figure 3.10a shows an example of a 7-day integrated product on a $1/3^{\circ}$ grid which merges data from Topex and ERS-1. It matches the region and approximately the time period of the single-track record shown in Figure 3.8. The two major eddies, the Great Whirl and Southern Gyre, have been labeled. Figure 3.10b plots the percentage error estimated formally by the optimal interpolation algorithm used to fill gaps between the tracks and integrate data from several different sources. In the region of intersecting gyres it is shown to be everywhere less than 10%, because the different orbit tracks have filled each other's gaps effectively. Elsewhere on the image there are regions where the error approaches 30%.

Comparison between Figures 3.8 and 3.10 shows the different character and utility of the two forms in which altimeter data are supplied, the along-track record and gridded maps. For precise knowledge of actual SSH variability and

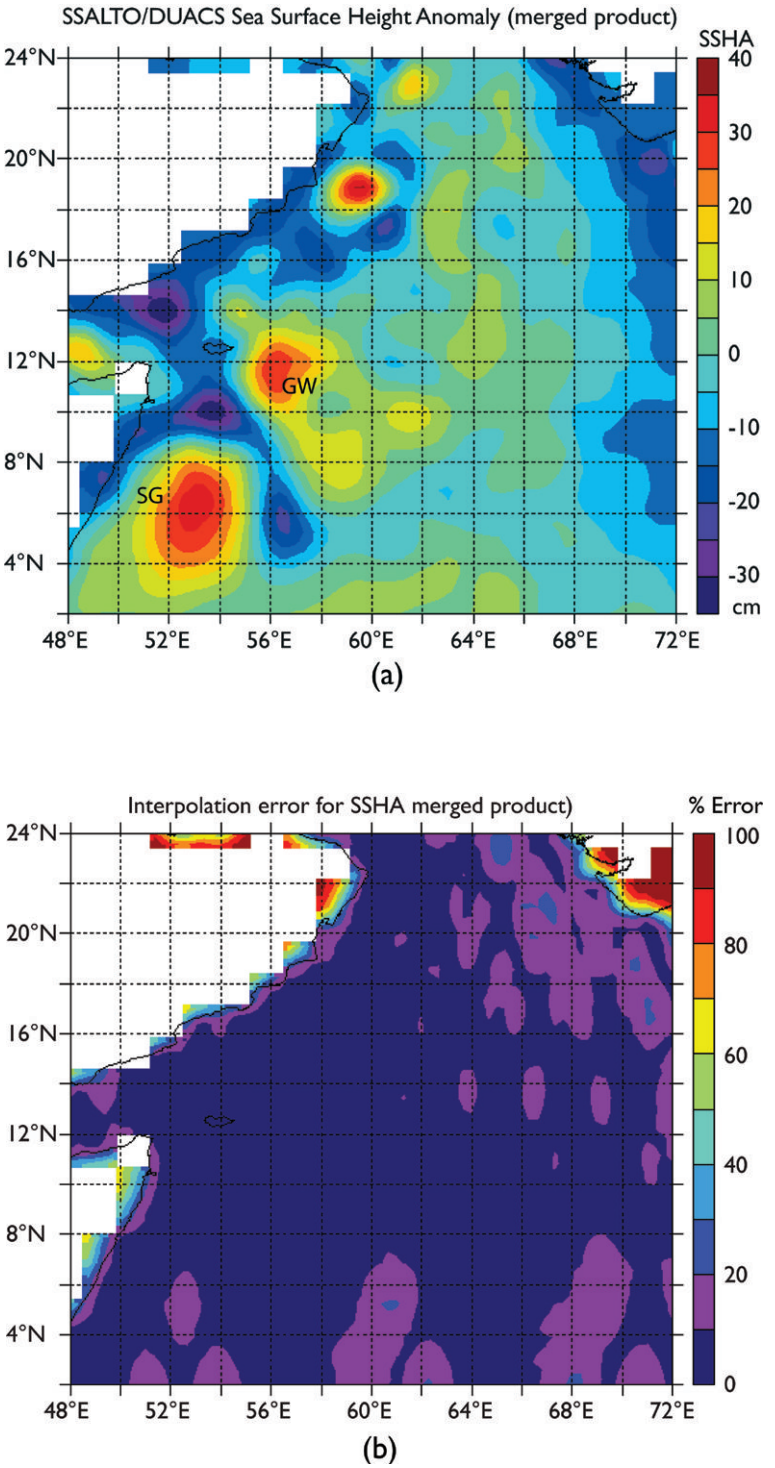


Figure 3.10. (a) Sea surface height anomaly over Arabian Sea on August 4, 1993, produced as a SSALTO/DUACS merged product from all available altimetry records, mapped onto a 1/3° grid. SG is the Southern Gyre eddy and GW the Great Whirl eddy. (b) Map of the formal error estimate for (a), associated with interpolation from along-track records on which it is based (gridded data obtained from the AVISO website).

research studies concerned with investigating altimeter performance, the along-track record is essential. On the other hand, for a regional study of where and how often eddies are to be found, their scale, their movement, and their relationship to each other, mapped data are more useful. The availability of high-quality altimeter data since 1991 has underpinned detailed analyses of specific, regional, mesoscale variability systems in a variety of locations; for example, in the Gulf of Alaska where eddies track westwards at around 500 km/yr along corridors related to boundary currents of the Alaskan Gyre (Okkonen *et al.*, 2001), or in the Labrador Sea where spatial patterns were detected for seasonal and interannual variations of eddy energy (Brandt *et al.*, 2004). In the Indian Ocean at 25°S, eddy behavior is shown (Palastanga *et al.*, 2007) to be related to instabilities in shear between the eastward-flowing South Indian Ocean Countercurrent and the deeper, westward-flowing South Equatorial Current. Because mesoscale variability impacts on a variety of other oceanographic processes, the results of such studies have led to a better understanding of what is influencing variability in local ocean variables that might otherwise appear to be random.

The capacity for using SSHA maps to identify individual eddies and follow their evolution, transport, and decay has allowed an inventory to be made of the Agulhas rings, shed from the Agulhas retroflection into the Atlantic Ocean over several years (Schonten *et al.*, 2000). This has wider implications for estimating the role of these rings as a source of warm and salty Indian Ocean water flowing into the Atlantic overturning circulation (van Leeuwen *et al.*, 2000).

An increasing number of mainstream oceanographers are also using readily available altimeter data to provide the spatial and temporal context which can link together other measurements from ships and buoys when studying regional dynamical, chemical and biological processes, such as in the Newfoundland Basin (Caniaux *et al.*, 2001), the North Atlantic (Mouriño *et al.*, 2002), southeast of New Zealand (Stanton and Morris, 2004), and the Bay of Bengal (Gopalan *et al.*, 2000).

SSHA maps of finer spatial resolution, with $1/8^\circ \times 1/8^\circ$ cells, are produced for the Mediterranean Sea (as shown in Figure 3.11). Clearly evident are a series of

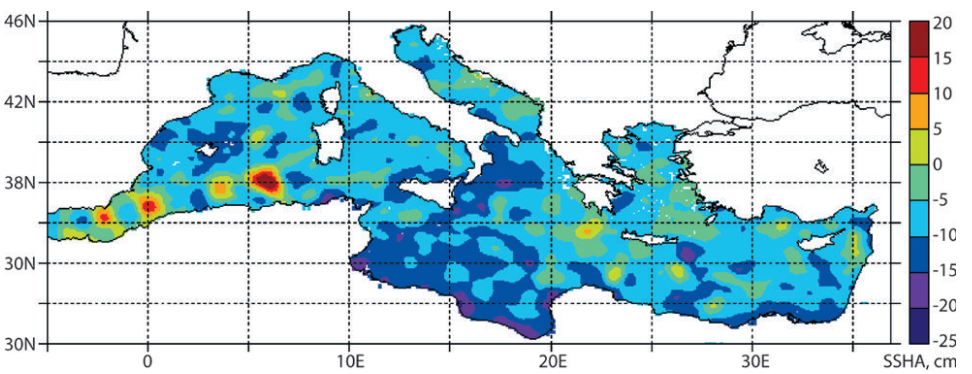


Figure 3.11. Sea surface height anomaly over Mediterranean Sea on May 10, 2006, produced as a SSALTO/DUACS merged product from all available altimetry records, mapped onto a $1/8^\circ$ grid (from the AVISO website).

strong rings in the western Basin off the North African coast, as well as mesoscale activity throughout the sea. In a detailed analysis of a 7-year span of such data, individual eddies were identified and characterized (Isern-Fontanet *et al.*, 2006). Trajectories of the most persistent eddies were found to have complex but well-defined patterns.

While each individual map of SSHA can be useful it is in the time series of SSHA data that most information is to be found about the ongoing evolution of the geostrophic current field. Much can be learned simply by viewing a sequence of SSHA maps as a movie. The eye soon identifies those highs and lows that persist, indicating that a feature is not simply random noise but is representative of a developing dynamical feature. At short lengthscales up to a few 100 km, mesoscale eddies are readily identified and their motion, if any, can be tracked.

3.4.2 Present limitations of satellite altimetry

The use of altimetry to observe ocean eddies is first of all constrained by the accuracy of around 2 cm to 3 cm with which SSHA can be measured, preventing small-amplitude mesoscale signals from being detected. The challenge facing researchers is to demonstrate the significance of results when the signals being detected are marginal in relation to this level of accuracy.

Because the spacing between adjacent tracks is over 250 km at the Equator, gridded maps have a much coarser spatial resolution than along-track records, and this remains an unavoidable limitation of altimetry until wide-swath altimeters are flown (see *MTOFS*, section 11.5.5). It is also consistent to spatially smooth the data when averaging over a period of 7 or 10 days, since finest scale variability detected in the along-track record is unlikely to persist unchanged over the averaging period. Therefore only the mid-size ($\sim$150 km) and larger mesoscale eddies can be confidently resolved in two dimensions by single-sensor altimetry, study of the higher spatial frequencies being restricted to the analysis of linear tracks. Therefore the continued deployment of several altimeters at the same time is essential to achieve finer spatial resolution that can identify mesoscale variability down to sizes of 50 km.

The other fundamental limitation of altimetry today is that derived velocity measurements are relative to steady ocean circulation. While it might be argued that this should not be a fundamental problem for studying variability, it can nonetheless be difficult to interpret the streamlines in variability flow when there is a strong mean flow that is being omitted. For example, the eddy fields in the major ocean currents like the Gulf Stream appear to have closed streamlines (the contours of SSHA are equivalent to geostrophic flow lines), whereas if the steady Gulf Stream flow were added many of these eddies would be revealed as meanders of the main current.

However, there are approximate estimates of the absolute dynamic topography (ADT) already being produced. These are based on the most recently produced geoid using GRACE data and *in situ* current measurements to estimate the steady part of dynamic topography (as mentioned in Section 2.4.5). An example of one of these is shown in Figure 3.12a. It is from the Agulhas retroflection zone, where strong steady

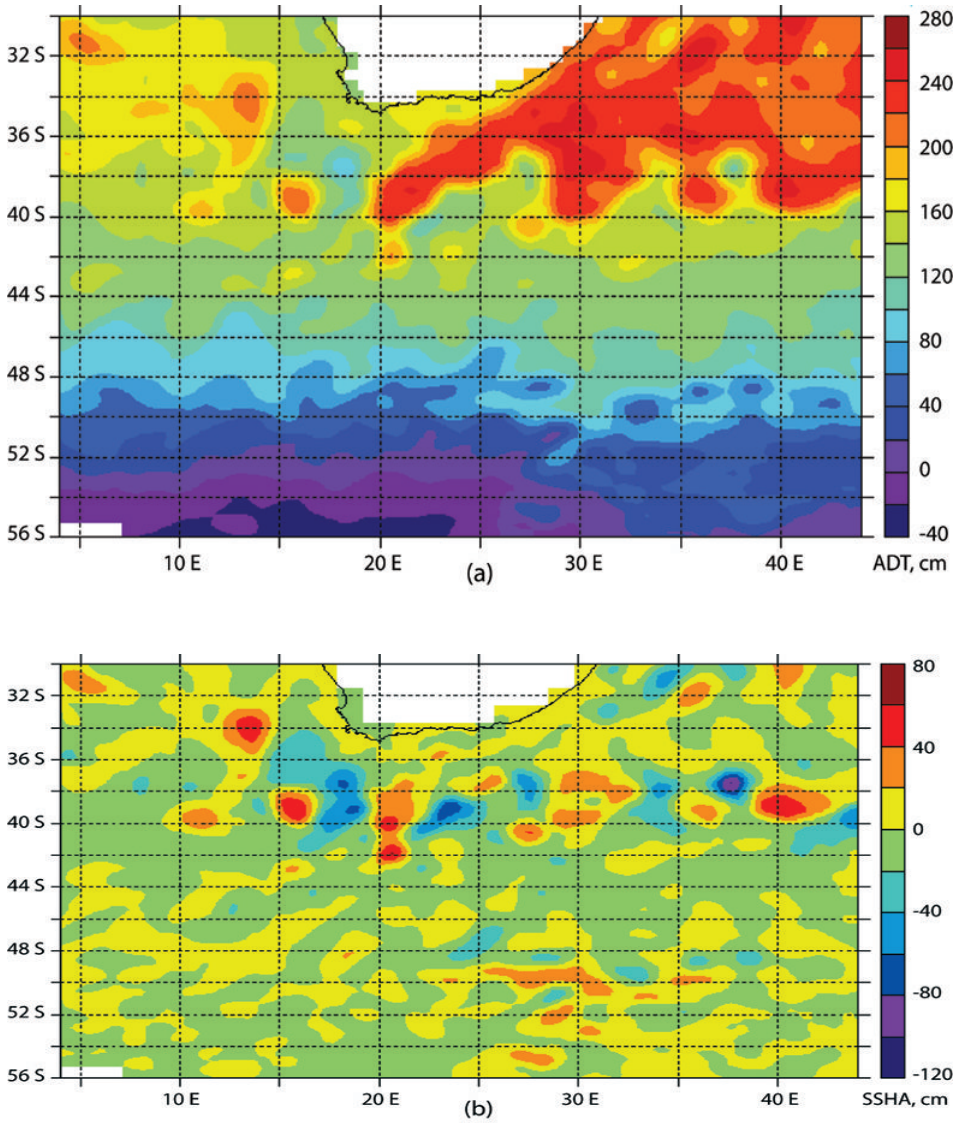


Figure 3.12. Sea surface altimetry data products for the Southern Ocean off South Africa on August 25, 1993, produced as a SSALTO/DUACS merged product from all available altimetry records at the time, mapped onto a 1/3° grid. (a) Approximate absolute dynamic topography (ADT). (b) Sea surface height anomaly (SSHA) (from the AVISO website).

flows in the Agulhas and Antarctic Circumpolar Currents underpin the variable current field. Compared with the corresponding SSHA map in Figure 3.12b, it is much easier to interpret Figure 3.12a because the streamlines of the absolute geostrophic flow field are contours of the ADT. Although these provisional ADT

datasets are still of questionable accuracy, applications have already emerged where they make an important contribution, such as the study of nutrient pumping for mesoscale phytoplankton variability (described in Section 5.6).

3.4.3 Kinematic measurements from altimetric SSHA fields

The ability of altimetry to redefine every 7–10 days the two-dimensional distribution of the stream function of the surface ocean current field is a remarkable achievement which provides dynamical oceanographers with a very powerful tool for monitoring ocean circulation. The essential value of such information when assimilated into ocean-forecasting systems is discussed in Chapter 14. Knowledge of the nearly instantaneous flow field also permits the kinematic properties of the flow to be examined, thus defining the relative movement of particles in different parts of the flow and providing insight into the two-dimensional transport and dispersion of tracers in the upper ocean, within the limits of the geostrophic assumption and the accuracy and sampling resolution of SSHA determination.

Applying the geostrophic equations (2.9) to determine the east and north components—u and v, respectively, of the ocean variability current—distributed across the ocean surface, allows us to calculate three kinematic properties:

$$\omega = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \frac{g}{f}\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right) \quad \text{is the vorticity (about the vertical axis),} \tag{3.2}$$

$$S_n = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = -\frac{2g}{f}\frac{\partial^2 h}{\partial x\,\partial y} \quad \text{is the normal component of strain, and} \tag{3.3}$$

$$S_s = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = \frac{g}{f}\left(\frac{\partial^2 h}{\partial x^2} - \frac{\partial^2 h}{\partial y^2}\right) \quad \text{is the shear component of strain.} \tag{3.4}$$

where h is the sea surface height anomaly.

Fluid-dynamical studies (Okubo, 1970; Weiss, 1991) have explored the way in which particles, which were initially adjacent, diverge when passing through a complex flow field. In some parts of the flow, straining of the surface causes them to diverge, implying strong dispersion. In other parts of the field their positions are "frozen" relative to each other which implies that they remain trapped in that region. These and other studies (e.g., Bracco *et al.*, 2000; Pasquero *et al.*, 2001) showed that, among other things, the dispersion behavior of the field can be characterized by a single parameter:

$$W = S_n^2 + S_s^2 - \omega^2, \tag{3.5}$$

which is now called the Okubo–Weiss parameter. It is evident from Equation (3.5) that where this is negative then vorticity dominates and where it is positive strain dominates. The studies confirmed that where $W > 0$, shear or normal strain disperses adjacent particles, and the reverse is the case where $W < 0$.

In a series of interesting papers (Isern-Fontanet *et al.*, 2003, 2004, 2006) this theoretical framework was applied to altimetry data in the Mediterranean Sea.

SSHA fields were converted into fields of W which revealed strong features matching the eddy patterns in the SSHA field. Large eddies contained a core of negative W, an order of magnitude greater than the background perturbations of W measured in parts of the sea where there is little eddy activity. The presence of a strong negative W peak indicates a ring where water particles remain trapped within the vorticity-dominated core. The negative core is surrounded by a ring of steep radial gradients outside which W is strongly positive in the vicinity of the eddy. This denotes the region where eddy velocity is high and there is strong shear that ensures vigorous stirring of water particles resulting in strong mixing with the water outside the eddy. In modeling studies, tracked particles are shown not to cross the contour where $W = 0$. A closed zero contour of W surrounding a strong negative core therefore indicates the circumference of an ocean ring inside which the water mass is trapped. If the movement of rings can be tracked, use of the Okubo–Weiss parameter to define the size of the core allows the water transported by the ring to be estimated approximately. This diagnostic approach for analyzing the SSHA field to explore mixing, stirring, dispersion, and transport in ocean eddy fields is already becoming more widely adopted (Waugh *et al.*, 2006; Chelton *et al.*, 2007; Henson and Thomas, 2007).

3.4.4 The distribution of mesoscale turbulent energy

Instead of looking explicitly at individual eddies or defining the flow fields, the SSHA from altimetry can also be used to measure the spatial distribution of the strength of mesoscale turbulence. As explained in section 11.6.2 of *MTOFS* the simplest way to do this is by evaluating the variance or the root mean square (r.m.s.) of the SSHA at each sampling point along the altimeter ground track over an extended period of time containing many overpasses (Ducet *et al.*, 2000). A high r.m.s. indicates that there are considerable changes in geostrophic velocity at a variety of length and timescales. Regions of high-SSHA r.m.s. on global maps (see figure 11.26 of *MTOFS*) tend to follow the major ocean currents because in general it is steady ocean currents and forcing which drive them to provide the energy that sustains mesoscale turbulent flow. The Gulf Stream north and east of Cape Hatteras, the Kuroshio Current to the east of Japan, and the Antarctic Circumpolar Current at certain longitudes are good examples.

Figure 3.13 shows a local map of r.m.s. SSHA presented at much finer resolution for the region of the Brazil–Malvinas Current confluence. At this scale it becomes clear that, despite their apparently random location on individual SSHA maps, the eddies occupy preferred regions and avoid others. By overlaying mean positions of the front at the edge of the Brazil Current and the Sub-Antarctic Front (SAF) Saraceno *et al.* (2004) showed how the eddies are strongest on the warm side of where the two fronts coincide near the Argentine coast, and never penetrate south of the SAF. Towards the east where the two fronts are separated by 8° line of latitude, there is a zone of low variance between higher variance adjacent to each of the fronts. The solid line represents a distinct province in relation to primary production, the Zapiola Rise, identified from ocean color data as a zone where the annual

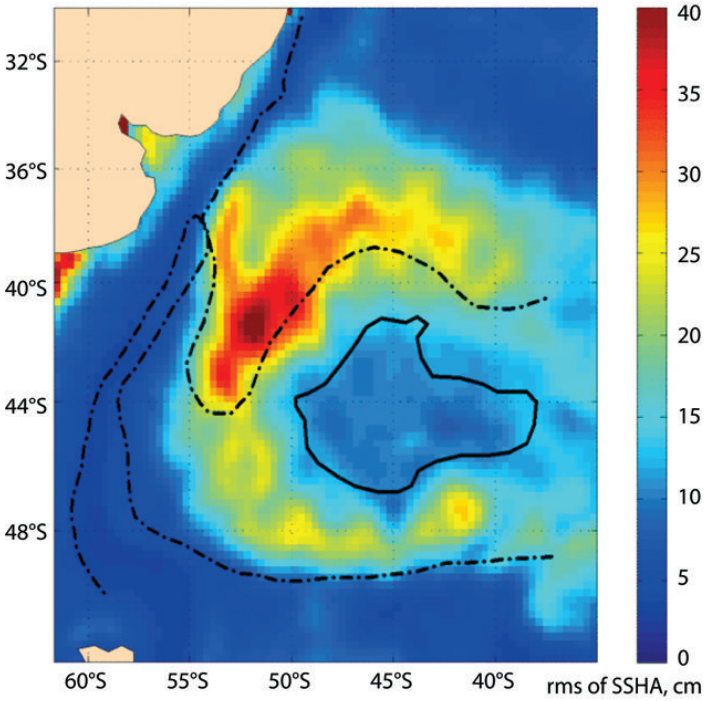


Figure 3.13. Root mean square (r.m.s.) of sea level anomaly obtained from 11 years of sea surface height from ERS + TOPEX/Poseidon + Jason-1. Dash-dot lines are mean position of Brazil Current Front (north) and Subantarctic Front (SAF) (south). The solid line encloses a region of relatively low variance (from the AVISO website; originally published in Saraceno *et al.*, 2005).

chlorophyll peak occurs 3 months later than in the surrounding waters (Saraceno *et al.*, 2005). Correspondence between this province and relatively low eddy activity points to a dynamical control of biological behavior.

The eddy kinetic energy (EKE) is defined as $\frac{1}{2}(u^2 + v^2)$ where u and v are, respectively, the meridional and zonal time-varying geostrophic currents associated with the SSHA. Although for each overpass only the cross-track component of the eddy velocity field is measured, the EKE can be based on the square of this value, averaged over the resolution cell of the final EKE product, as long as the eddies are assumed to be isotropic (having no preferred orientation). As shown in figure 11.27 of *MTOFS*, a long-term average map of EKE global distribution shows similar patterns to those of the r.m.s. SSHA.

EKE distribution can also be evaluated almost instantaneously, based on data from only a few days. A sequence of these reveals the time evolution of the EKE field. This promises an important way to use altimeter data for understanding what may be driving mixing processes or upwelling at a local scale (e.g., in relation to patchiness of primary production). Figure 3.14 shows an example of a snapshot of EKE acquired from merged altimeter data in the Argentine Basin of the southwest Atlantic. Comparing this with time-averaged r.m.s. SSHA in Figure 3.13 it is evident that the near-instantaneous energy distribution is spatially more variable and finely granulated than the impression given by the long-term mean.

To be reliable, snapshots of eddy energy require several altimeters to be operating at once. A single satellite is limited in its sampling of the full spectrum

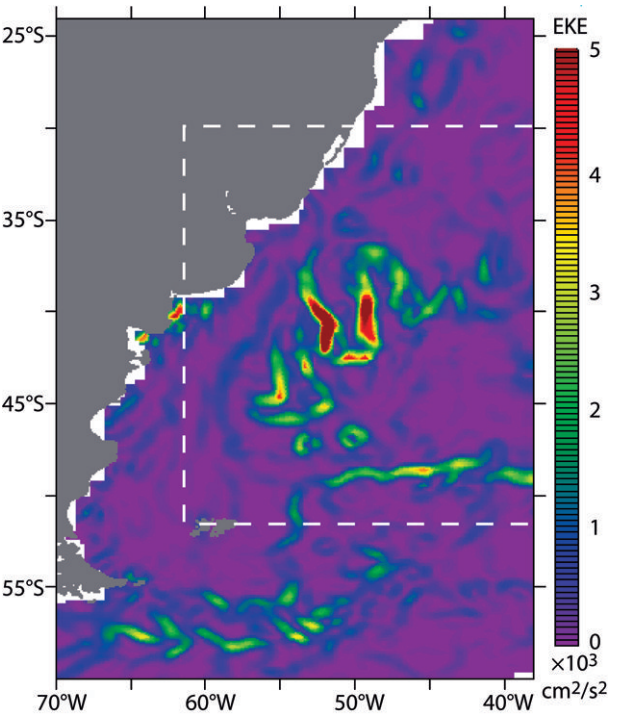


Figure 3.14. Eddy kinetic energy (EKE) in cm^2/s^2 evaluated from merged altimeter SSHA records on November 5, 2005 in the Brazil Malvinas confluence zone in the southwest Atlantic Ocean. The white dashed box shows the region represented in Figure 3.13 (adapted from an image acquired from the AVISO website at *http://www.aviso.oceanobs.com*).

of mesoscale variability because its fixed repeat aliases the higher frequencies and thus tends to underestimate total EKE. By averaging EKE over just a few days from overpasses of several different sensors, it is possible to monitor more confidently how EKE and its distribution varies with time (Le Traon and Dibarboure, 1999). Figure 3.15 illustrates the difference between a map of EKE based on data from a single altimeter and from four. Figures such as this provide clear evidence of how important it is to maintain a sufficient number of altimeters in operation at all times if we are to continue to monitor the ocean accurately. They also serve as a warning to users not to rely on maps based on sparse altimetry data without making a formal assessment of the confidence that can be placed in them.

3.5 OBSERVATION OF EDDIES AND MESOSCALE TURBULENCE IN THE SST FIELD

3.5.1 SST signatures of eddies in infrared imagery

Figure 3.16 shows a classic example of an SST image, derived from infrared radiometry, illuminating a wealth of mesoscale variability and revealing a familiar feature, the Gulf Stream, to be a richly complex dynamical phenomenon. Because most viewers of this image will come to it with some background knowledge of the

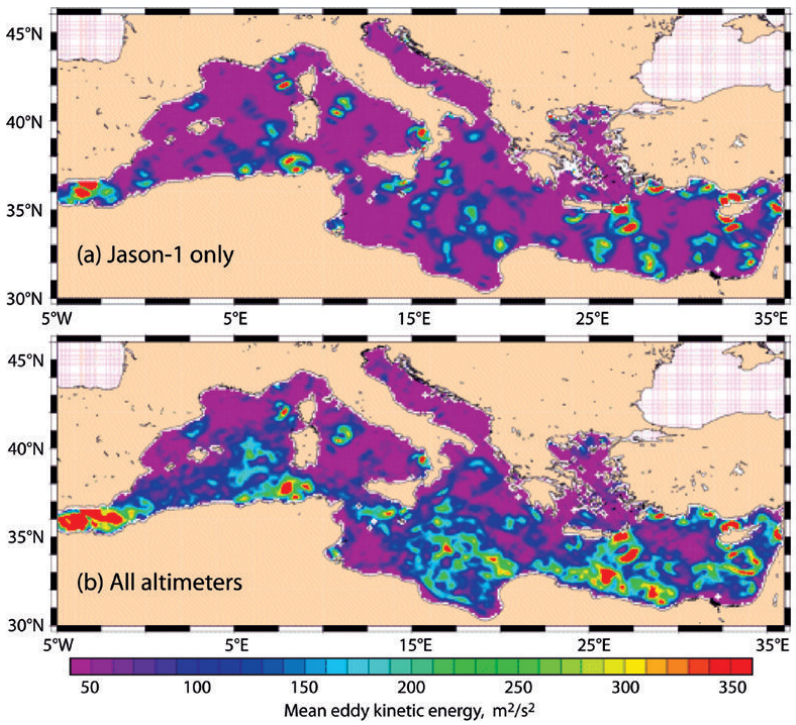


Figure 3.15. Mean eddy kinetic energy in the Mediterranean, as evaluated from SSHA measurements (a) by Jason-1 alone and (b) by Jason-1, ERS-2, TOPEX/Poseidon, and Geosat follow-on combined. The magnitude of observed kinetic energy is higher with the combination of satellites, and shows more detailed structure (adapted from images displayed on the AVISO website and produced by MFS/CLS).

Gulf Stream as a major ocean boundary current it is relatively easy to interpret dynamically. The band of warm (but not the warmest) water meandering east-northeast away from the U.S. coast is obviously the main Gulf Stream, flowing along the front between the warmer Atlantic water in the central Atlantic Gyre and the cooler water along the U.S. coast and to the north. We shall consider the signature of fronts in Chapter 4, but here the Gulf Stream sets the context in which to interpret eddy structures. For example, the largest eddies are clearly in the process of being created by the wide meandering of the front, pinching off isolated cores of cold water to the right of the main flow direction, to produce cyclonic cold core eddies, and injecting warm core anticyclonic eddies into the cooler ocean to the north. This is entirely consistent with the simple model of eddy construction illustrated in Figure 3.7, in which the eddy has a thermal signature visible from space which corresponds to its density structure. The same is true of the large central eddy shown in Figure 3.3a.

However, the smaller mesoscale structures in Figures 3.16 and 3.3a reveal variability in the ocean in a different way. Many of the smaller scale features seem

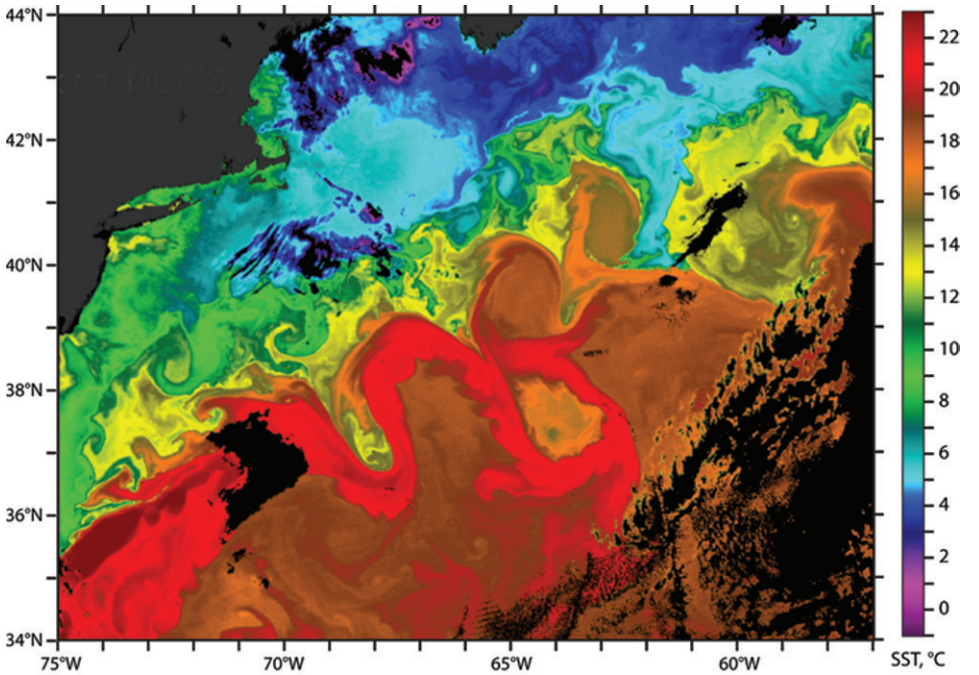


Figure 3.16. Sea surface temperature field derived by MODIS on Aqua, April 18, 2005, showing the meanders of the Gulf Stream (from NASA Ocean Color website at *http://oceancolor.gsfc.nasa.gov/*).

to show up because pre-existing temperature gradients in surface waters are being stretched or squeezed by two-dimensional turbulence, leading to hook-like shapes where there is a shear zone between flows of different magnitude and hammerhead patterns where two eddies interact. Both these distinctive isotherm patterns seem to be characteristic signatures of an energetic field of two-dimensional mesoscale turbulence. Figure 3.17 shows similar features in a thermal image of mesoscale variability at the finest resolution (1 km pixels) available from satellites. It confirms that down to scales of a few kilometers infrared images can provide a lot of fine-scale detail about the temperature structure. However, at short scales up to a few tens of kilometers, as distinct from the large eddies in Figures 3.16 and 3.3a, what can such images tell us about the underlying dynamical processes?

In practice it is difficult to derive firm information about eddy dynamics from these smaller scale structures. As long as we can see the thermal signature of large eddies then it is reasonable to assume that the ocean current flow is aligned locally with the smoothed isotherms of these features, in geostrophic balance as it is along a major front. However, as soon as the isotherms start to be advected by two-dimensional turbulence they become deformed by stretching and squashing, advection disturbs the geostrophic balance, and it is no longer appropriate to assume that the isothermals are parallel to the streamlines at smaller deformation scales. We are

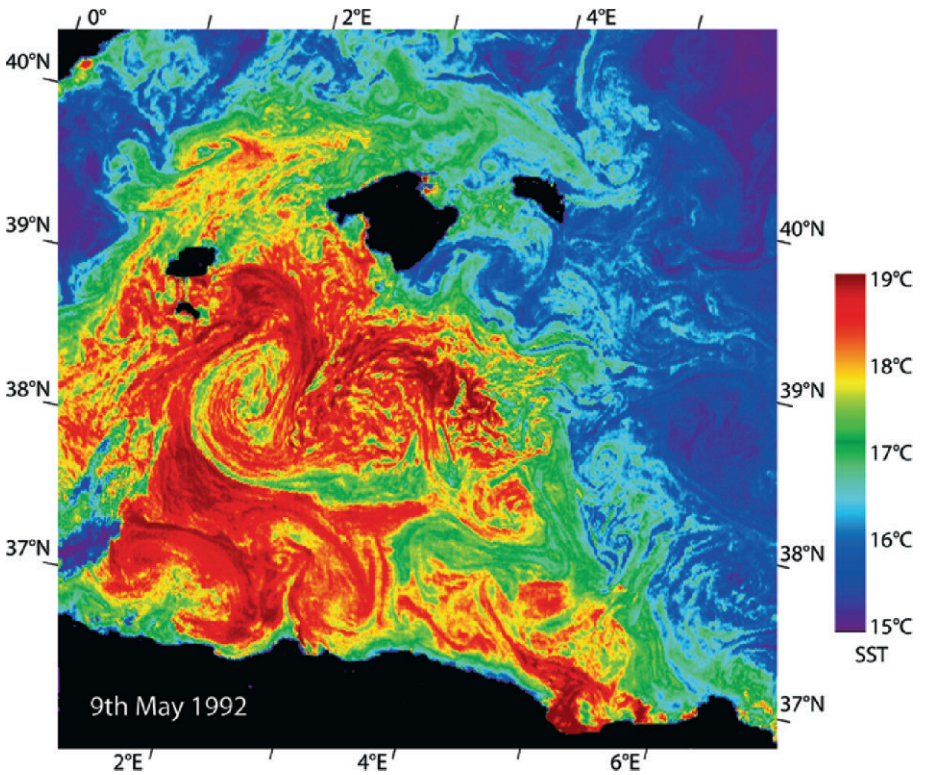


Figure 3.17. Map of level 2 SST data from ATSR nighttime image on May 9, 1992 over the Balearic Islands in the western Mediterranean Sea, showing fine-scale thermal structures viewed using a spatial resolution of 1 km. This image is 500 km across.

left with a confusing set of patterns, repeated at smaller and smaller scales. A time sequence of images would show how ephemeral the smaller features are. They appear and disappear in response to the flow fields of larger features. This is not surprising since it is a consequence of the cascade of turbulent energy from larger to smaller scales. At small enough scales, thermal gradients disappear as local mixing tends to homogenize the temperature.

It is reasonable to conclude that a rich texture of mesoscale structure in an infrared image implies that horizontal stirring is occurring. However, it is not straightforward to quantify that stirring, or eddy kinetic energy, simply by measuring the variability of the SST field. Moreover, once the surface temperature field has been homogenized by stirring and mixing it can no longer serve as a tracer even though there may still be considerable mesoscale eddy activity occurring. To serve as a tracer for any mesoscale mixing process, large-scale gradients on the temperature field need to be regularly refreshed. This is precisely what is happening in the Gulf Stream case, where the main current continuously reinforces the thermal contrast between the current and the surrounding waters past which it flows.

It is also important to note that not all open-sky images contain such rich spatial detail as some (even for the same area a few days apart). This may be a consequence of atmospheric effects when large amounts of water vapor attenuate surface signatures in atmospherically corrected data products. It may also result from the signatures of mesoscale variability being erased from the SST by other processes in the mixed layer. For example, at very low wind speeds, spatially patchy diurnal thermoclines may develop during the day. Alternatively under windy conditions the upper ocean is stirred more vigorously which destroys the rich mesoscale patterns that may depend on ocean processes below the wind-mixed layer. At intermediate wind speeds roll vortices can occur in the atmospheric boundary layer which induce a response in the patterns of sea surface roughness (as discussed in Section 3.7), especially if there are surface-active materials like organic film present. By affecting local air–sea heat transfer such surface phenomena can induce parallel streaks of warmer and cooler surface skin SST. When these are advected they could account for some of the curved streaky patterns at lengthscales of a few kilometers that are visible in Figure 3.17.

Under these circumstances, SST signatures may tell us more about local air–sea interaction effects than the underlying mesoscale dynamics. It is important to remember that the main center of action of mesoscale dynamics is at the depth of the maximum gradient of density (i.e., the thermocline). Thermodynamic processes driven from the sea surface may, if strong enough, "paint their own story" onto the SST field, hiding or distorting the thermal patterns which are driven by the dynamical activities focused at the thermocline. This marks an important distinction between SST and SSHA when used as remote-sensing signatures of mesoscale dynamics, because the SSHA field is hardly affected by small-scale thermodynamic processes in the surface layer.

Because infrared radiometry can resolve spatially down to 1 km, and radiometrically to a few tenths of a Kelvin, it is possible to acquire some spectacular images (there are more examples in Chapters 4 and 5) which qualitatively are exciting and interesting, although difficult to interpret quantitatively. However, it is rare to obtain more than isolated examples of such a cloud-free expanse (as shown in Figure 3.16). Even that image contains a number of patches of cloud obscuring the temperature field. Given the ubiquitous occurrence of cloud over most parts of the ocean, the prospect of obtaining a time sequence of infrared images allowing us to follow the evolution of the SST field in fine detail is generally a vain hope. A realistic approach when using thermal imagery at 1 km resolution is to look for occasional snapshots of stunning imagery which can inform us about the type of fine-scale processes occurring, but we should not expect to follow these processes dynamically. To obtain a time sequence of SST fields from satellite infrared data that are less sensitive to clouds requires the use of either level 3 composite images, with a reduction of temporal resolution to several days, or level 4 analyses which interpolate between several data sources to obtain daily images with cloud gaps filled in. Neither of these options reliably preserves the spatial resolution and fine-scale detail of individual cloud-free images.

3.5.2 Microwave radiometry for viewing ocean eddies

This brings us to an alternative for observing the thermal signature of mesoscale eddies from space, that of using SST from microwave radiometers (MWRs). The characteristics of this methodology are introduced in Section 2.3. and treated in more detail in chapter 8 of *MTOFS*. MWRs are unaffected by the presence of cloud and only when there is heavy rainfall in the field of view of the radiometer do they fail to measure SST. Their most serious drawback is their coarse spatial resolution, presently no better than about 60 km. This means they cannot contain the fine spatial detail of SST that is seen in some infrared imagery, but it is essentially no poorer than altimetry-gridded fields, and Section 3.4 shows what a rich source of information they are for monitoring mesoscale dynamics. Moreover by sampling at a spatial interval much less than the resolution (typically about 20–30 km) it is possible to identify structures smaller than 50 km. The radiometric sensitivity and absolute accuracy of MWRs are somewhat poorer than the best infrared radiometers, although this can be expected to improve in future.

Figure 3.18 shows a time sequence of AMSR-E SST fields over the ocean off South Africa, enclosing the region covered by Figure 3.2. The images are 3-day means and are presented at intervals of 6 days on a spatial grid of 0.25° of latitude and longitude. Perhaps contrary to theoretical expectations this turns out to be a very effective way of following the complex dynamics of this interesting region. Pulsing of the Agulhas Current becomes apparent as it produces rings in the retroflection region (van Leeuwen *et al.*, 2000). With some enhancement it is possible to see some eddies with fairly weak thermal signatures drift into the Atlantic Ocean, while the warm front that flows eastward from the retroflection is seen to develop instabilities. Farther south the generation of meanders on the ACC is clearly visible, despite the general fuzziness of the features which is an inevitable consequence of spatial oversampling. Although it is not possible to show this on the printed page, animation of the images presented in Figure 3.18 provides the oceanographer with a sense of experiencing the character of mesoscale turbulence in just the way that an aerodynamic designer might study video loops of flow past objects in a wind tunnel. Thus it brings to life ocean flows varying at lengthscales of order 100 km over a span of thousands of kilometers. Moreover, this is not a specially selected example of optimal conditions, but is repeatable throughout the year.

Comparison with Figure 3.12a, the absolute dynamic topography of the same regions although from a different time, shows that the altimeter and the MWRs provide a similar view of the larger mesoscale eddies and appear to complement each other well. However, it is interesting to note from this comparison that the Agulhas rings that propagate into the Atlantic appear to have a clear and unambiguous signature in the altimetry compared with the SST where they are difficult to detect. This reminds us that even when the cloud cover problem is largely solved by the use of MWRs, mesoscale eddies may not necessarily have a signature in SST. Nonetheless, for monitoring the propagation and evolution of those large-mesoscale eddies that do perturb the SST sufficiently to be detectable from space, microwave radiometry appears to be a better method than infrared. That judgment is based on

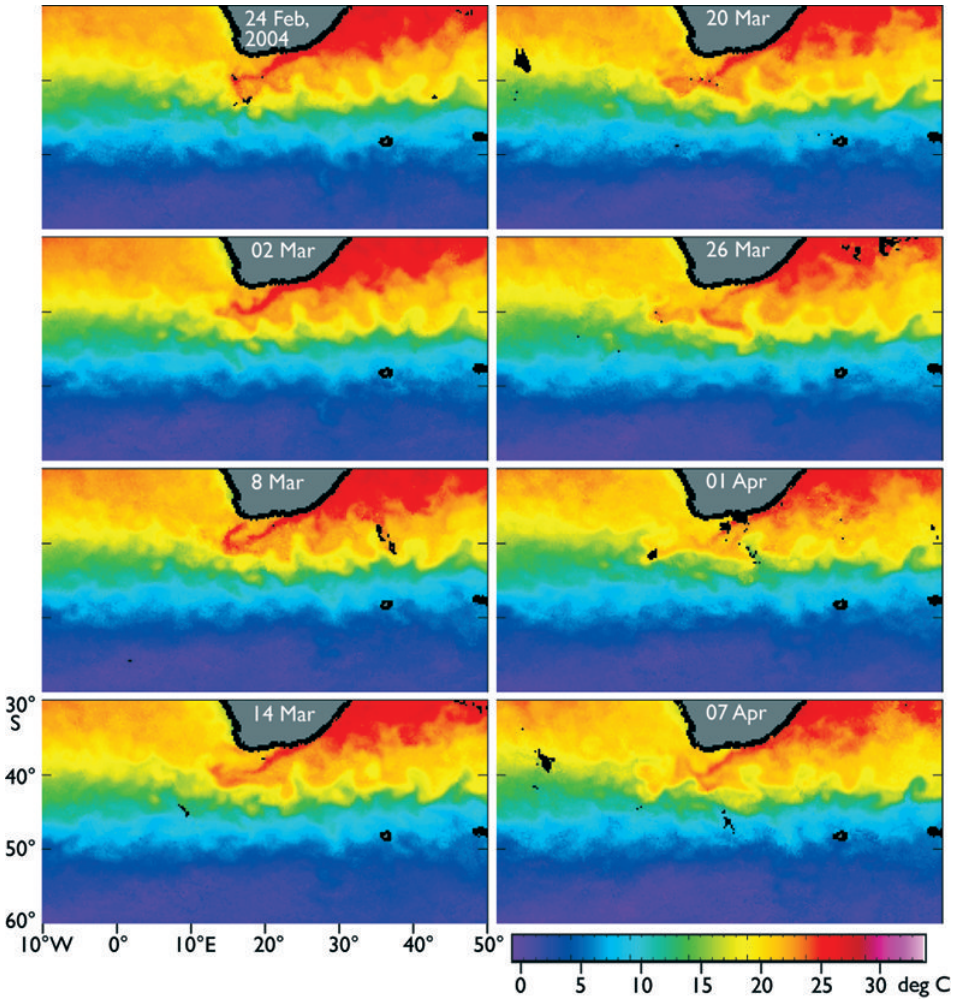


Figure 3.18. SST maps of the Southern Ocean poleward of South Africa, as measured by the AMSR-E microwave radiometer between February 22 and April 7, 2004. Each image is a composite from all overpasses in the 3 days up to and including the date given. The sequence shows an image every 6th day to reveal the clear and unambiguous evolution of mesoscale features in a clear field with few data dropouts (except for the loss of data within 100 km of land). The pixel size is 1/4° lat. and long. (AMSR-E data are produced by Remote Sensing Systems and sponsored by the NASA Earth Science REASoN DISCOVER Project and the AMSR-E Science Team. Data are available at *www.remss.com*.)

the capacity of MWRs to deliver time sequences of unblocked fields of SST data every two or three days. In the light of the discussion in Section 3.5.1 about the uncertainty of interpreting the dynamical significance of high-resolution SST fields from 1 km to 50 km, the inability of MWRs to sample at such fine resolution

is not a drawback, although it is for many other applications of satellite SST measurements.

3.6 VIEWS OF MESOSCALE TURBULENCE FROM OCEAN COLOR

There are two different ways in which the apparent color of the sea viewed from space reveals mesoscale variability in the near-surface layers of the ocean. First, color can be used to characterize distinct water masses. Figure 3.19 shows an example of this for the Gulf Stream, using an image of chlorophyll concentration derived from the ratio of the blue and green visible wavebands of MODIS at the same time as the SST map in Figure 3.16. In the large eddies shed by the Gulf Stream in the North Atlantic it is very evident that warm-core rings contain nutrient-depleted water with only small concentrations of chlorophyll while cold-core rings contain more nutrients, are more productive, and have a higher chlorophyll concentration. Viewed on the broad scale, an image like this readily shows how mesoscale turbulence has begun to redistribute the water masses that started distinctly on either side of the main Gulf Stream front.

The second way of approaching the interpretation of remotely sensed patterns of

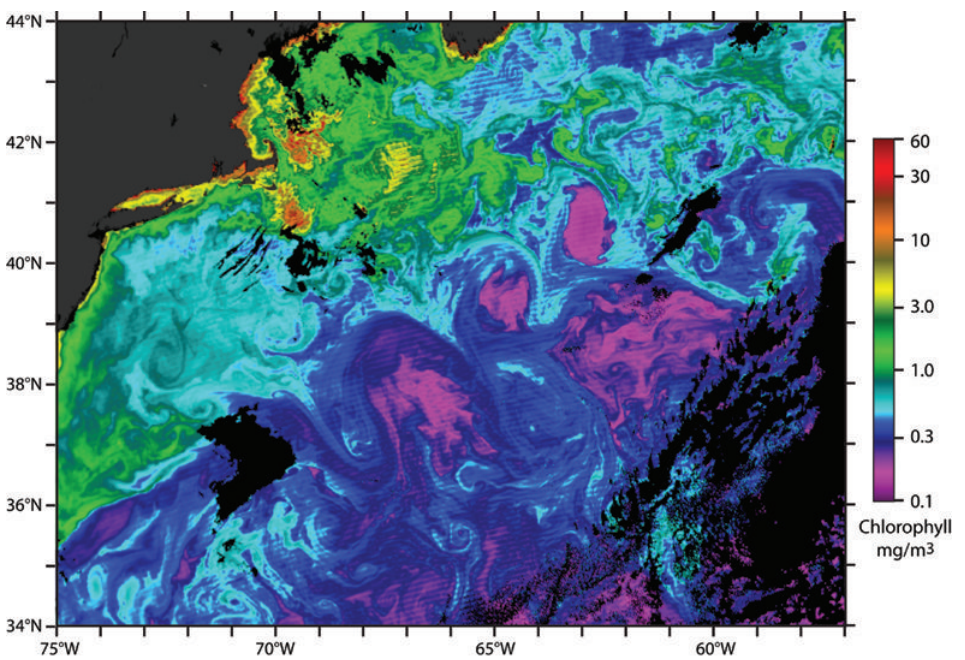


Figure 3.19. Map of chlorophyll concentration derived from MODIS on Aqua, April 18, 2005, showing the Gulf Stream. This corresponds to the temperature map in Figure 3.15 (from NASA Ocean Color website at *http://oceancolor.gsfc.nasa.gov/*).

Figure 3.20. A break in the clouds over the Barents Sea on August 1, 2007 reveals a large coccolithophore bloom. This image is created from the level 2 atmospherically corrected data acquired by the 531 nm waveband of MODIS on NASA's Terra satellite, at 09:30 on August 1, 2007. The image is about 600 km wide, 430 km high, and located about 300 km north of Murmansk (data accessed and downloaded through the "Level 1 and 2 browser" on the NASA Ocean Color website at *http://oceancolor.gsfc.nasa.gov/*).

color is to consider them as tracers of the water flow and in particular the stirring effect of turbulent flow. This is evident to some extent in the fine-scale structures in Figure 3.19. It is even clearer in Figure 3.20, which shows a single-band reflectance image from the green (531 nm) waveband of the MODIS on Terra, in the Barents Sea north of Norway during what appears to be a coccolithophore bloom, and Figure 3.21, which shows chlorophyll distribution in the Gulf of Aden, derived from SeaWiFS ocean color data.

Considering first what these images can tell us about the mesoscale dynamics of the regions, it appears that localized high concentrations of colored material, clearly visible from space, have been drawn out into filaments by the shear flow of larger scale eddies, and then folded at smaller scales into swirls and hammerhead patterns by dynamical processes with lengthscales of less than ten up to hundreds of kilometers. These are very similar to the patterns revealed in the high-resolution thermal infrared images discussed in Section 3.5.1. As with thermal patterns it would be misleading to identify the streaks of color with streamlines in the flow,

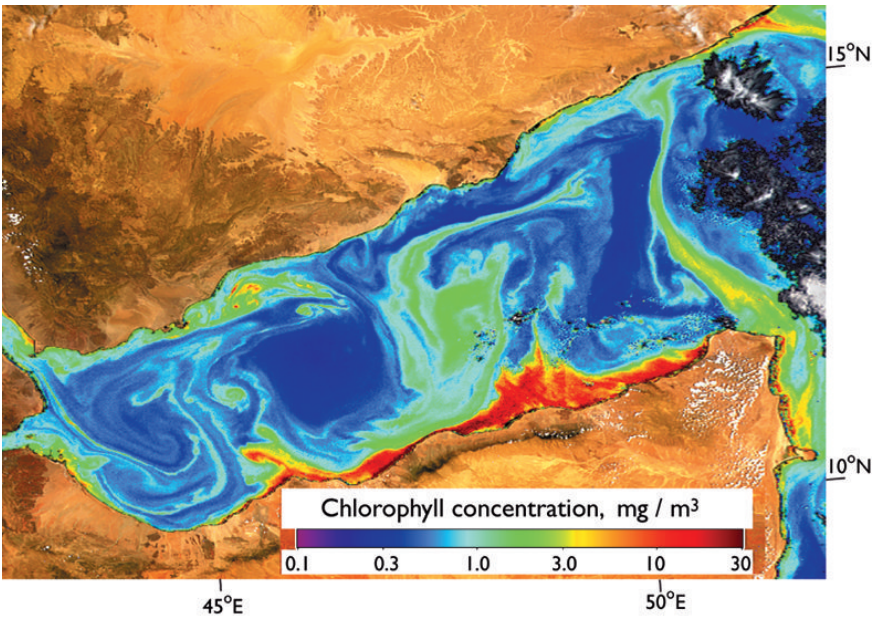


Figure 3.21. Map of chlorophyll concentration derived from the SeaWiFS overpass of the Gulf of Aden, between the Indian Ocean and the Red Sea, on November 1, 2003.

since this is not a steady flow. It is difficult to derive quantitative measures of the turbulence from color streak line patterns. Nonetheless, such an image does give a good qualitative impression of the complexity of turbulent motion. However, the absence of patterns should not be used to infer that there is no turbulent flow. Rather it suggests either that there has been no source of differentiation into patches of higher and lower concentrations of the visible tracer that "seed" the turbulence visualization process, or else that turbulent stirring has succeeded in completely mixing such patches into homogeneity again.

Most readers will be familiar with the everyday analogies of stirring colorant into a can of white paint, mixing cocoa into a chocolate cake mixture, or adding milk to a cup of coffee. There is a short timespan during which the tracer is spread throughout the rest of the fluid but remains in streaks that are sufficiently coherent to mark the flow in the main fluid. Then, as soon as the coherent threads are broken and the color is more completely dispersed, the capacity of the tracer to visualize the flow field is lost. This implies that the highlighting of turbulent mesoscale structures in the ocean by a colored tracer is an ephemeral event lasting a few days at most. To offer a sustained tracer capability of revealing mesoscale eddies there must be regular renewals of the source of colored material, which may be from phytoplankton blooms, from river inputs, from upwelling events, from fronts between waters of different optical properties or, in shallow seas, from resuspended sea bed material.

Moreover different tracer materials may have different natural decay times, through, for example, settling of particulate material, death of phytoplankton

cells, or decomposition of organic material. Ideally the characteristic decay timescale of the material that serves as a tracer needs to be significantly longer than the dispersive timescale of the eddy field. It is notable that where complex tracer fields have been revealed most clearly and tracked for several days they have often been associated with blooms of the coccolithophorid species *Emiliania huxleyi* (Holligan *et al.*, 1983; Holligan and Groom, 1986; Brown and Yoder, 1994). In such cases the visible tracer consists of the highly reflective liths that surround each plant cell, which continue to change the water's optical properties long after the phytoplankton cells which created them have died. Eventually these particles drop out of suspension in the upper layer and are no longer visible from satellites, but their lifetime as quasi-conservative tracers is likely to be longer than that of other types of phytoplankton blooms.

Figure 3.20 is almost certainly a coccolithophore bloom, because it has such a strong reflective signature across all the MODIS visible wavebands, from 412 nm to 667 nm, although in the Bering Sea resuspension of diatom debris has been shown to produce water color mimicking the appearance of *E. huxleyi* blooms (Broerse *et al.*, 2003). The 531 nm wavelength data were chosen because in this case they are brightest, but the eddy patterns are almost identical for all the bands, implying an almost white reflective signature. This contrasts with the normal chlorophyll signature that depends on absorption in the blue to turn the sea green as a tracer of other species of phytoplankton. The brightness of coccoliths means they can be detected by less sensitive, visible waveband sensors with broad spectral wavebands, such as the AVHRR on polar meteorological satellites (Ackleson and Holligan, 1989; Balch *et al.*, 1996), which has allowed scientists to be able to search continuously back to 1981 for evidence of blooms (Smyth *et al.*, 2004). It has enabled Landsat with its much finer spatial resolution but poorer radiometric sensitivity to image blooms clearly and allowed finer scale eddy structures to be observed. Finally, as Figure 3.20 confirms, it allows eddies to be traced at high latitudes where solar illumination is often too weak to provide strong color signatures. Unfortunately clouds are ubiquitous at high latitudes and cloud-free glimpses as wide as the example shown occur infrequently.

In contrast with coccoliths' broadband signature, the more usual phytoplankton blooms reveal the best signatures of mesoscale eddies only after they have been analyzed to retrieve the chlorophyll concentration. Figure 3.21 from the Gulf of Aden represents that kind of image. In this case the coloring of the image represents chlorophyll concentration. This helps to visualize the direction of the current flow which is expected to be along streaks of decreasing concentration. There remains a remarkable similarity between the mesoscale patterns in Figures 3.20 and 3.21, confirming that it is the dynamical structure and not the inherent characteristics of the tracer which is being revealed. The Gulf of Aden image shows quite clearly the importance of patches or regions of high chlorophyll concentration which are needed to seed the flow visualization process. In this case primary production occurs mainly along the southern coast.

The interesting, square eddy shape, presumably constrained by the width of the Gulf, is highlighted as jets of flow across the Gulf entrain plumes of high chlorophyll

concentrations which gradually decay away from the source, but do not finally disappear before they have illuminated the way the gyres split up into smaller eddies. This image also provides graphic confirmation of the usefulness of satellite data for helping to interpret local measurements. Ship-based surveys of the Gulf of Aden, with no reference to satellite data, had previously shown that the currents in these eddies are up to 0.5 m/s at the surface and 0.2 m/s down to the sea floor between 1,000 m and 2,000 m (Bower *et al.*, 2002). Eddies are important for influencing the spreading rates of intermediate-depth Red Sea water. A subsequent study by the same team (Fratantoni *et al.*, 2006), which now makes full use of ocean color and other satellite data in addition to conventional *in situ* observations, is able to present a much fuller understanding of eddy behavior and processes. Comparison of these two papers offers clear evidence of how the perspective gained from satellite data illuminates the challenging task of interpreting relatively sparse *in situ* samples along transects through dynamically complex parts of the ocean.

Another study of SSHA in the Gulf of Aden (Al Saafani *et al.*, 2007) also confirms how altimetry and ocean color data are complementary, since altimetry cannot detect the fine-scale stirring and mixing revealed by visible wavelength radiometry in Figure 3.21. The same paper also provides an interesting comparison, at a resolution scale suitable for altimetry, between simultaneous fields of an 8-day composite of chlorophyll from SeaWiFS and SSHA from DUACS over the Arabian Sea. It is reproduced here in Figure 3.22. It completes our brief review of

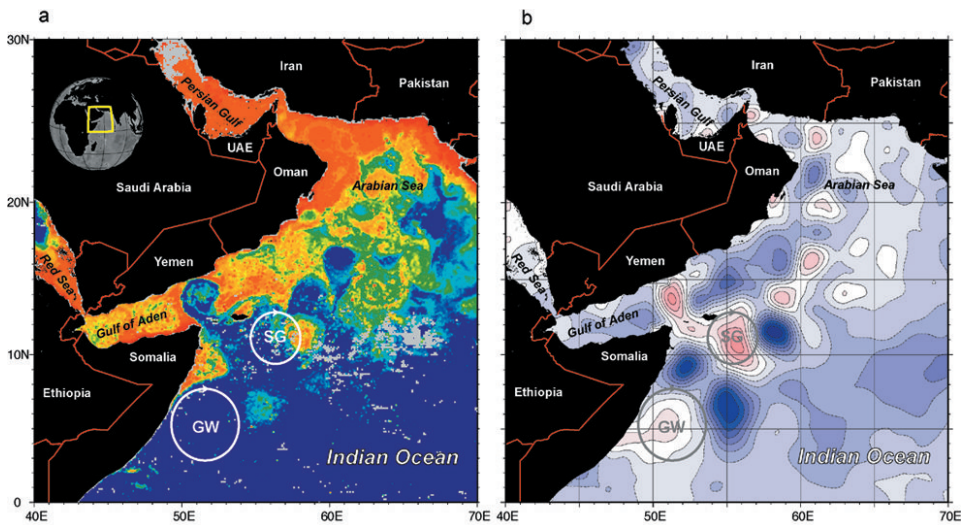


Figure 3.22. (a) SeaWiFS chlorophyll *a* composite image (November 7–15, 1999). (b) Sea level anomaly (November 10, 1999) illustrating the wide range of mesoscale variability evident in the Arabian Sea and the western tropical Indian Ocean. Approximate positions of the anticyclonic Great Whirl (GW) and Socotra Gyre (SG) are schematically indicated, pink shading representing high and blue low SSHA. These images correspond to a time period immediately following the onset of the northeastern (boreal winter) monsoon (this is a copy of figure 1 from Al Saafani *et al.*, 2007.)

eddies observed in ocean color by showing that the patches of high or low chlorophyll that might be interpreted to represent the cores of rings are in fact not precisely co-located with eddies defined by SSHA. While most anticyclonic eddies (positive-SSHA and warm-core) correspond quite well to regions of low chlorophyll, the patches of higher chlorophyll do not precisely overlay cyclonic gyres (low-SSHA, cool-core). This is presumably because phytoplankton populations have a previous history and respond to surface currents advecting them rather than to dynamical geostrophic forces that link temperature and SSHA more closely together. This example should curb a tendency to interpret any circular-looking patch of high chlorophyll as evidence of the core of an eddy, in the absence of any further information.

Despite the interesting studies cited above, and the evidence of some occasional spectacular images, it must be accepted that the use of ocean color imagery for measuring the dynamical characteristics of mesoscale turbulence is in practice quite limited. Because it requires not only cloud-free conditions but also a source of a contrasting water color, this cannot be considered as a reliable method for systematically visualizing ocean eddies. Rather it is a technique to be used opportunistically when the conditions are right, and within that constraint it can provide some unique insights into mesoscale turbulent stirring in the ocean. However, this rather cautious judgment refers only to applications in ocean mesoscale dynamics, since the primary benefits of ocean color sensors are for marine biology and biogeochemistry.

3.7 SURFACE ROUGHNESS SIGNATURES OF EDDIES

It may come as a surprise that mesoscale eddies can be detected in images from synthetic aperture radars (SARs). For example Figure 3.23 shows a very clear example of the depiction of eddies off Norway (Johannessen *et al.*, 1996) . It is placed side by side with a matching SST image for the same day, confirming that the eddy-like patterns on the SAR image do indeed correspond to mesoscale turbulent structures in the SST field.

3.7.1 Hydrodynamic modulation patterns of eddies

The capacity of SAR to reveal mesoscale dynamical processes in this way is consistent with the approach to SAR ocean image interpretation presented in section 10.7 of *MTOFS*, in particular the hydrodynamic modulation mechanism. Wherever there are patterns in ocean surface currents that result in local convergence or divergence at the surface, these tend to modulate the energy of surface wind waves generating zones of steeper waves along convergences and lower amplitude waves at divergences. It is these heterogeneities in the surface wave field that SAR detects as variations in surface roughness, recorded in the pixel-by-pixel estimates of the radar backscatter cross-section, σ_0. The features that stand out most strongly in the SAR image are local fronts where convergence or divergence is strongest, leading to bright

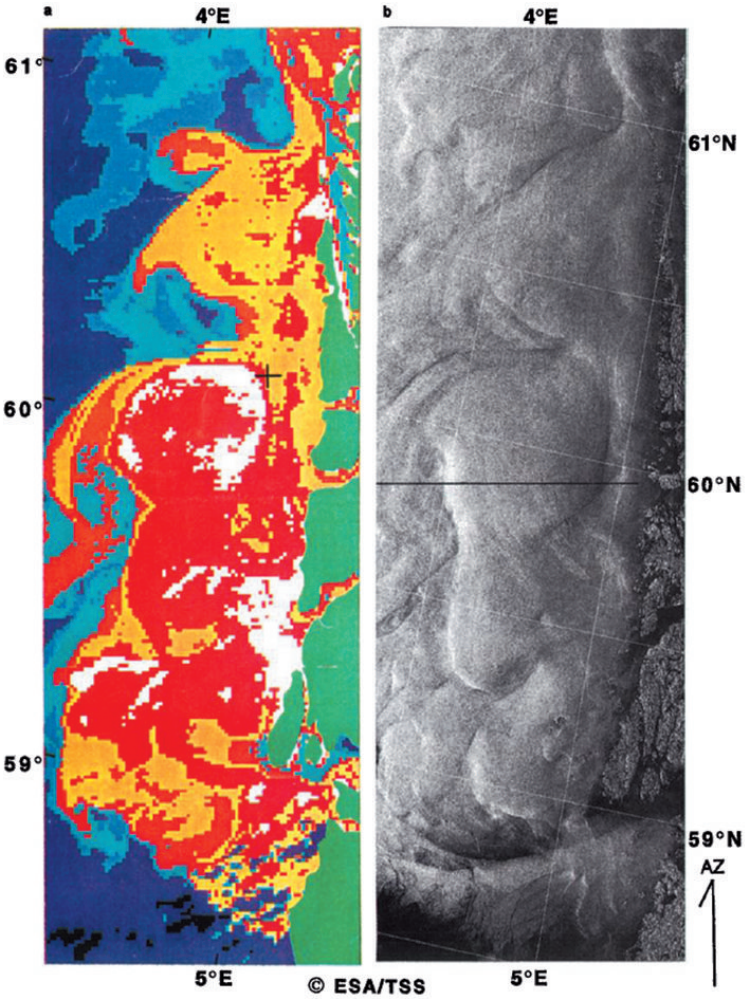


Figure 3.23. Comparison between (*left*) a 1 km resolution NOAA 11 AVHRR IR image acquired at 14:20 UTC on October 3, 1992 (white is 14°C and purple is 12°C; land is masked in green and clouds in black) and (*right*) 100 m resolution ERS-1 SAR image acquired at 21:35 UTC on October 3, 1992. Both images cover the same 100 km by 300 km region off the west coast of Norway between 59°N and 62°N (this is a copy of plate 3 in Johannessen *et al.*, 1996).

or dark lines. Thus the SAR image highlights the parts of the turbulence field where nonlinear processes are locally forcing the creation of fronts (frontogenesis) as the eddies evolve. In the example of Figure 3.23 there appears to be strong interaction between individual eddies as they respond to the vorticity field of their neighbors.

For SARs to be able to detect mesoscale eddies by hydrodynamic modulation requires not only that a suitable amount of surface convergence is present in the velocity field, but also that wind is of moderate magnitude. It needs to be strong enough to create surface waves whose modulation can be detected by SARs, but not so strong that the SAR backscatter measurement is saturated and therefore becomes spatially homogeneous. Furthermore the character of the SAR signatures of local fronts within the eddy system depends on the relative orientations of the convergence field, the wind (and hence wind-driven surface wave direction), and the azimuth

pointing direction of the radar. These need to be optimal for the clearest imaging, and a curved front may change its character from bright to dark along its length. Careful analysis of these processes (Johannessen *et al.*, 1996) suggested that they are capable of being modeled mathematically in order to predict the SAR image response to a given set of surface flow, wind, and wave conditions. However, the complexity of inverting such models (Romeiser and Alpers, 1997; Romeiser *et al.*, 1997, 2001) is very challenging. It is unlikely that systematic retrieval of surface currents in eddies will become an operational possibility by this means in the near future. On the other hand, as operational ocean remote sensing develops in partnership with improved ocean circulation–forecasting models (see Chapter 14), this should provide the required background knowledge of synoptic conditions to enable an inverse modeling approach to quantify surface currents associated with the clearest SAR current signatures. An improved approach along these lines (Johannessen *et al.*, 2005), has been proposed, based on a refined radar-imaging model (Kudryavtsev *et al.*, 2005). It should also be noted that a rather different approach to measuring surface currents directly from Doppler shift measurements retrieved from SAR processing (Chapron *et al.*, 2005) is showing promise. It is mentioned further in Chapter 4.

Figure 3.24 shows another example of hydrodynamic modulation revealing mesoscale turbulent structure, this time from the Kuroshio Current to the east of

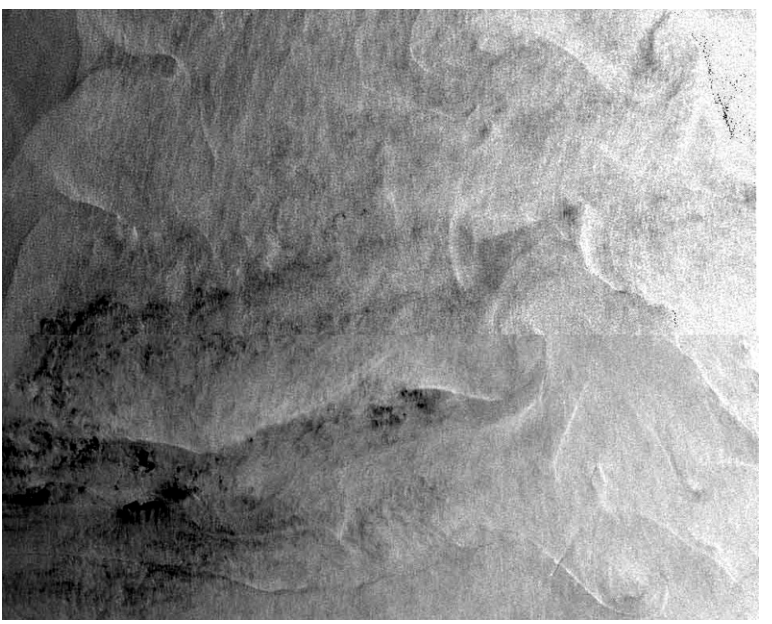


Figure 3.24. ERS-1 SAR image over the Kuroshio Current in the northwest Pacific Ocean, acquired on December 23, 1994 showing hydrodynamic modulation of surface roughness by eddies, meanders, and filaments in the current. It is located at about latitude 40°45′N , longitude 144°E. The width of this image is 100 km. It is oriented with the vertical axis aligned about 10° east of north (downloaded from Alpers *et al.*, 1999).

Taiwan. In this case there appear to be several plume-like features protruding from left to right. As well as the mesoscale structure being defined by dark or light frontal lines there appears to be a difference in tone between the plume regions and elsewhere. These could be intrusions of colder water, since the slightly darker image tone indicative of weaker backscatter could be a consequence of the colder sea producing a more stable atmospheric boundary layer (ABL) leading to a smoother sea surface (see figure 10.28c of *MTOFS*). This is a different kind of imaging mechanism which complements hydrodynamic modulation. With no additional information available, this interpretation can only be an informed guess, although similarity with Figure 3.23, where the thermal image supports this interpretation, gives some encouragement that this speculation is on the right lines.

3.7.2 Slick-modulated signatures of eddies

There is another means by which mesoscale eddies are revealed in SAR images. This is based on the slick modulation mechanism described in section 10.7.2 of *MTOFS*. In very low winds, patches of the sea surface where a surfactant film has aggregated remain completely flat, reflect no radar backscatter, and so appear black on a SAR image. These are what appear to the human eye as slicks on the sea surface. As long as the wind is above about 2 m/s the patches where there is no film are roughened enough to cause a small amount of backscatter, creating a contrast on the image between slick and nonslick regions. If the wind is less than 1.5 m/s the whole surface is flat and produces a featureless dark image. When the wind rises above about 5 m/s the slick-covered regions start to roughen and the contrast is reduced or destroyed altogether. Slicks are important for the detection of oil (see Chapter 14) or biogenic surfactants associated with primary production. However, as Figure 3.25 shows, in the right circumstances they have an evident capacity to reveal mesoscale turbulence patterns.

The patterns in this image look so much like our expectation of mesoscale eddies that it is very easy to jump to the conclusion that somehow we are seeing flow streamlines, comparable with some color tracers (as discussed in Section 3.6). But before we adopt this interpretation a mechanism is needed to make a causative connection between surface slicks and mesoscale turbulence. How are these patterns formed? There are two important points to consider about surface films as tracers. First, they always remain at the surface, unlike colored material in suspension which will be drawn into a convergence zone and then may be carried by downwelling below the depth at which they are detectable. In contrast, surfactant material is drawn by surface convergence towards a converging frontal zone, but then remains trapped at the surface at the point of convergence. Second, the physical properties of surfactant material promote aggregation into a coherent film, and then resist its dispersal, meaning that what is drawn together by convergence is not drawn apart by divergence. This provides an explanation for the linear nature of most slick signatures on the image.

Lines of surfactant material may be produced at a frontal line. More likely they have been produced during a period of moderate wind that generates linear vortex

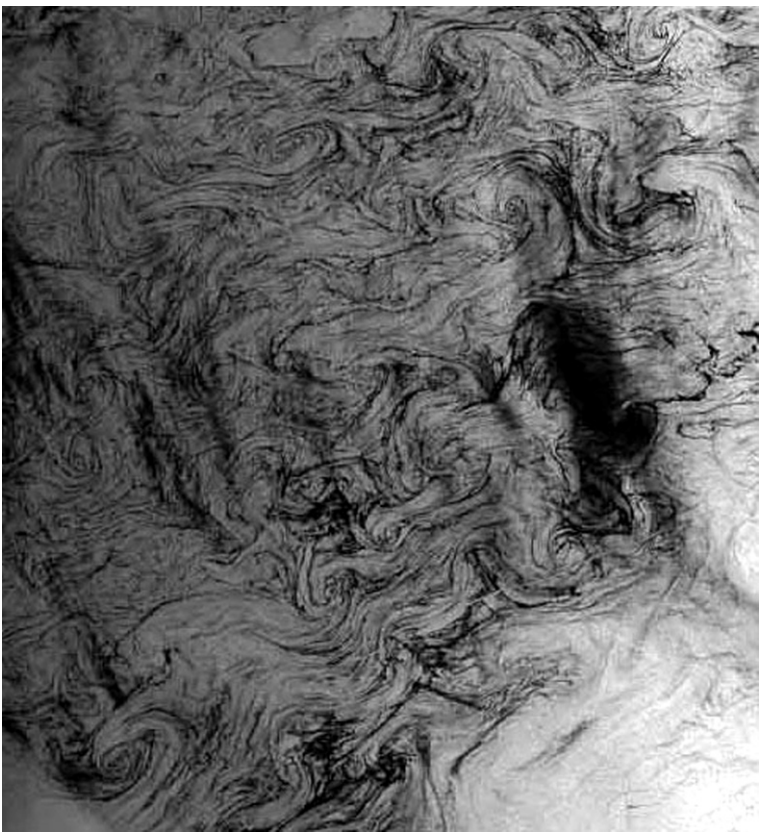


Figure 3.25. Ocean mesoscale eddies in Pacific Ocean east of Japan, revealed in an ERS-1 SAR image by slick modulation of radar backscatter. This image is 100 km wide and is centered at about latitude 42°N, longitude 146°E. It was acquired on September 22, 1995. (ERS SAR data are provided by the European Space Agency). It is oriented with the vertical axis aligned about 10° east of north (downloaded from Alpers *et al.*, 1999).

rolls in either the ocean or the atmosphere, aligned with the wind. These accumulate into parallel slicks along lines of convergence, and can often be seen in the sea where there is a supply of natural or anthropogenic organic material to form the slicks. Once slick lines have been produced, they will remain in existence until a strong enough wind is able to disperse them. In the meantime, if the wind drops they drift in response to surface flow associated with underlying mesoscale motions. In the case of a rotating eddy with slow convergence towards the center, the line is drawn into a spiral, such as can be seen in Figures 3.25 and 3.26. The latter is an example of a SAR image in the Tyrrhenian Sea north of Sicily in the Mediterranean. Meanwhile larger scale flows stretch and squeeze the surface, squashing the circular spirals into ellipses. Interacting eddies generate characteristic hammerhead patterns. Once

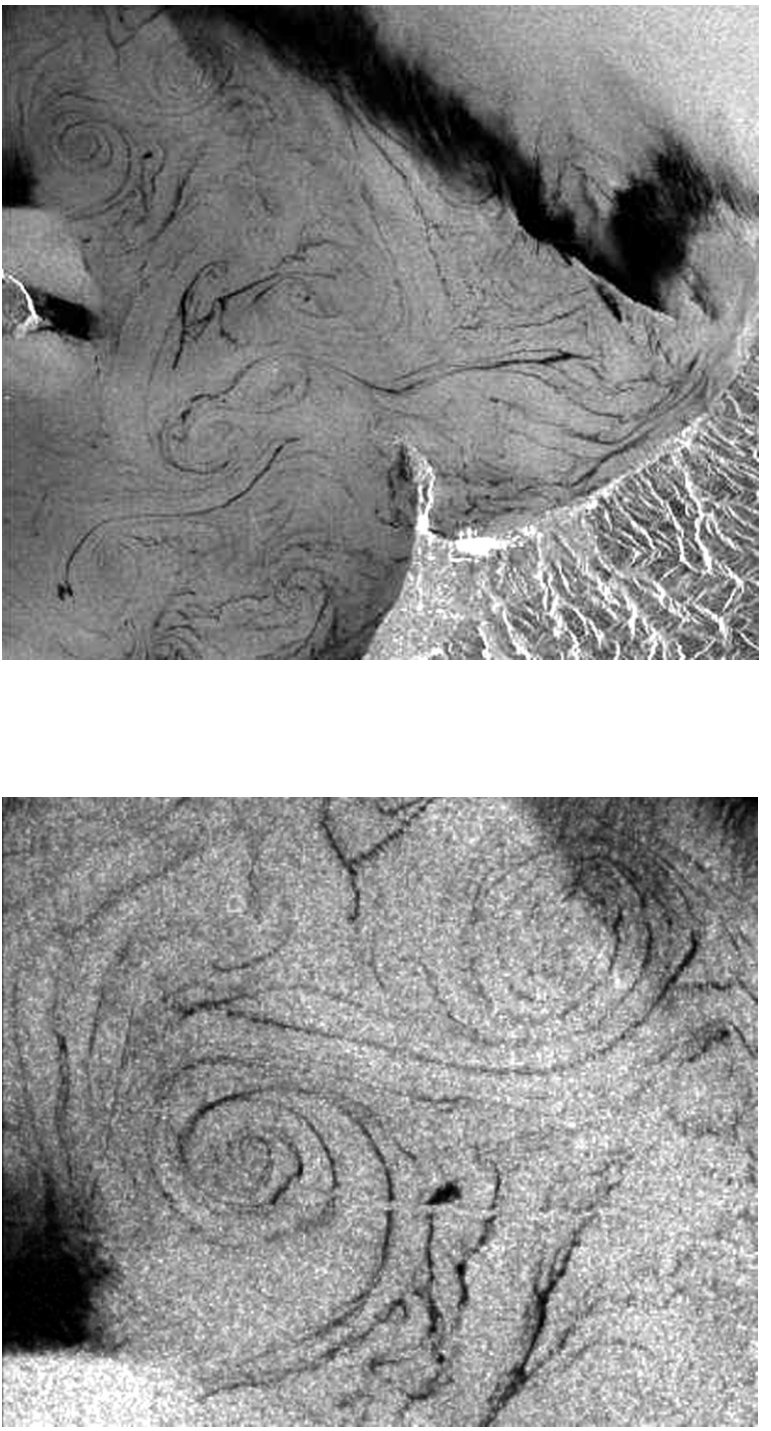


Figure 3.26. ERS-1 SAR image of the Tyrrhenian Sea north of Sicily, acquired on September 19, 1993. (a) 50 km wide overview of the region. (b) A 20 km wide enlargement of the northwest corner of (a) (downloaded from Alpers *et al.*, 1999).

convergence has brought a large patch of slick together it is likely to be "recycled" as a tracer for another eddy structure.

This conceptual mechanism provides a plausible explanation for the patterns that are seen, although little work has yet been done to confirm this by experiment. Yet even if it is valid, it remains difficult to retrieve dynamical information from SAR images. One issue to consider is that slicks are capable of preserving the history of what has happened in the previous hours or even days. For example, a small eddy may wind up a slick spiral, and then decay or move elsewhere. The slick pattern remains on the surface as a relict of the previous motion, until something else happens to deform it again. Thus the very "busy" images of slicks seen in the figures do not necessarily mean that the whole region is full of eddy activity all the time. This presents an interesting challenge for further research.

3.7.3 Sun glitter photography

This section would be incomplete without mentioning another source of information about the surface roughness patterns of mesoscale turbulence. This is the set of photographs obtained by astronaut-oceanographer Paul Scully-Power during a flight of the U.S. Space Shuttle (Scully-Power, 1986). An example is shown in Figure 3.27. These images rely on Sun glint to provide an imaging mechanism for the patterns of surface roughness. The roughness detected is comparable with that observed by a SAR, essentially related to the mean square slope of the sea surface. However, Sun glitter images are more dependent on the viewing geometry than are SAR images. The glitter pattern spreads solar reflection out from the point where the Sun's image would be reflected from a completely flat calm surface. Within this zone any parts of the sea whose mean slope is greater than its surroundings shows up darker, because fewer facets reflect the Sun. Over a region outside the glitter zone, but not too distant from it, rougher patches of sea will appear brighter, because some of the facets in the enhanced roughness now reflect the Sun. Too far away from the glitter zone the roughness imaging capacity is lost. This makes Sun glitter patterns difficult to interpret quantitatively, but these data have nonetheless inspired further analysis of what causes characteristic spiral eddies (Munk *et al.*, 2000).

The size of the area covered by the shuttle photographs is consdierably smaller than a standard 100×100 km SAR image, and the detected features are correspondingly smaller, but it seems that they each observe similar features, while interaction with mesoscale flows is partly explained by the slick modulation mechanism discussed above. When conditions are just right for Sun glitter photography (facilitated by a manned space flight) the technique seems to be more sensitive than SAR for detecting surface eddy signatures. The ubiquity of the spirals observed from the shuttle flight suggests that there remains more to be discovered from SAR data and that it is worth persevering in developing methods which routinely extract information about mesoscale dynamical variability from SAR images.

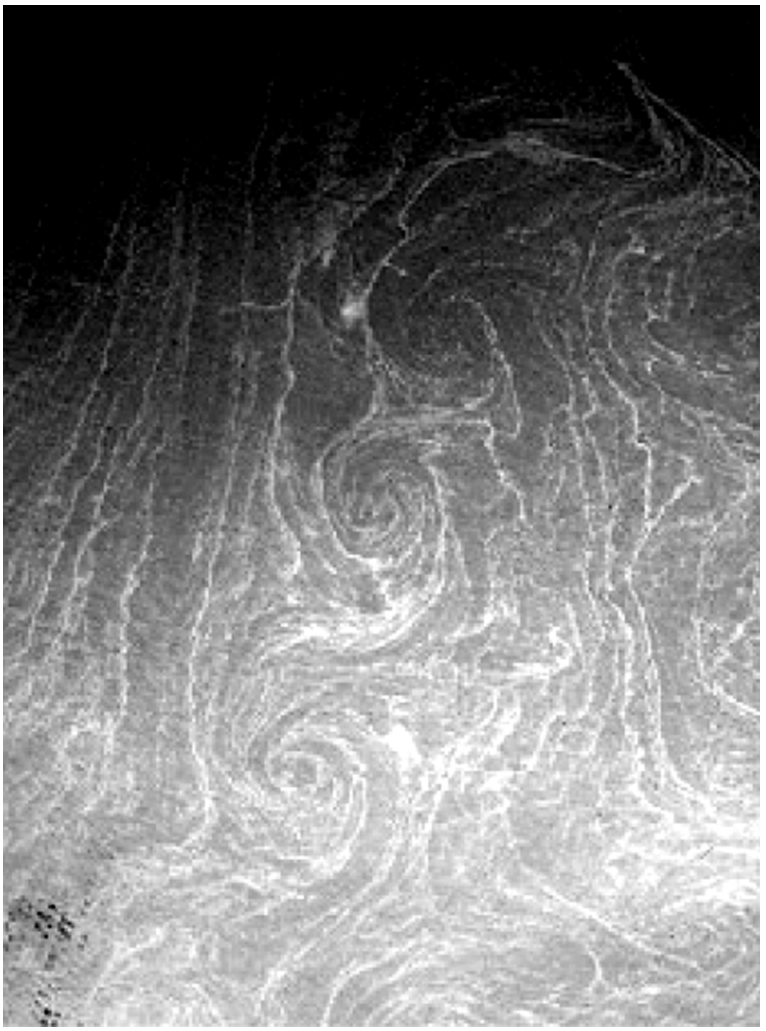


Figure 3.27. Example of spiral eddies in the Mediterranean Sea off Egypt (32°N, 26°E) on October 7, 1984, revealed in Sun glitter photography by Scully-Power (1986) from the space shuttle. The field of view in this image is about 25 km across. North is at the top. This image is discussed further by Munk *et al.* (2000) where it relates to their figure 2.

3.7.4 Can imaging radar become a reliable tool for observing turbulent eddies?

It can hardly be claimed that SAR today offers a systematic and reliable way of measuring ocean mesoscale eddy activity. Despite some spectacular examples, such as those shown here, no more than a few percent (probably less than 3–5%) of all ocean SAR images reveal any significant mesoscale dynamical features, and most show far less detail than the figures presented here. While theory-based analysis of SAR images containing hydrodynamic modulation signatures of eddies has demonstrated consistency in relation to SST images, it is a complex and uncertain procedure to unpack estimates of surface current fields from SAR ocean surface roughness data. The slick signatures of eddies show us detailed patterns of eddy-

like structures when the combination of low winds and surface active material is optimal, but at present no techniques have been developed to retrieve quantitative measures of the mesoscale turbulence they appear to represent. Perhaps in both cases SAR's potential rests in a qualitative mapping role, identifying when and where mesoscale turbulence is occurring, and providing some indication of the variability of lengthscales.

At the turn of the century the available geographic coverage of SAR was fairly poor over the ocean. ERS SAR images covered only 100×100 km and were infrequently acquired over the open sea where the scatterometer took preference. The sparseness of SAR acquisitions and the cost and effort of processing raw SAR data into digital images further discouraged the use of imaging radar for monitoring ocean eddy activity. But the deployment of ASAR on Envisat changed that outlook.

The availability of wide-swath SAR images from Envisat, covering a swath of nearly 500 km, has improved global coverage and achieved a frequency of acquisition (every 3 days) at mid to high latitudes. Because the scale of the eddies is large relative to the 30 m resolution of standard SAR images, wide-swath mode's poorer resolution with 100 m pixels is not a problem. Examples of wide-swath data suggest that an eddy-monitoring role may be feasible one day and is certainly worth further investigation, especially if new methods could be developed to automatically filter all acquired data for evidence of mesoscale features before operator analysis of the selected scenes. Moreover, there are also promising new techniques being developed to measure currents directly from Doppler shift, retrieved as part of standard SAR processing (see Section 4.2.5).

SAR therefore may yet turn out to become a valuable tool for oceanographers observing mesoscale eddies. Nonetheless, at the time of writing it remains the case that the combination of satellite measurements of SST and altimetry offers the most immediate benefits to oceanographic research about mesoscale eddies. This can be supported by ocean color images to elucidate the impact of mesoscale turbulence on primary production. The same is true for meeting today's requirement for operational ocean monitoring; the assimilation of altimetry and SST data holds the key for constraining ocean circulation models to represent mesoscale turbulent structures occurring in the actual ocean (see Chapter 14). The next two chapters about other mesoscale processes tell a similar story.

3.8 REFERENCES

Ackleson, S. G., and P. M. Holligan (1989), AVHRR observations of a Gulf of Maine coccolithophorid bloom. *Photogramm. Eng. Remote Sens.*, **55**(4), 473–474.

Adamec, D. (2000), Eddy flow characteristics and mean flow interactions in the North Pacific. *J. Geophys. Res.*, **105**(C5), 11373–11383.

Alpers, W., L. Mitnik, L. Hock, and K. S. Chen (1999), *The Tropical and Subtropical Ocean Viewed by ERS SAR*. ESA ESRIN, available at *http://www.ifm.uni-hamburg.de/ers-sar/* (last accessed April 25, 2008).

Al Saafani, M. A., S. S. C. Shenoi, D. Shankar, M. Aparna, J. Kurian, F. Durand, and P. N. Vinayachandran (2007), Westward movement of eddies into the Gulf of Aden from the Arabian Sea. *J. Geophys. Res.*, **112**(C11004), doi: 10.1029/2006JC004020.

Balch, W. M., K. A. Kilpatrick, and C. C. Trees (1996), The 1991 coccolithophore bloom in the central North Atlantic: 1, Optical properties and factors affecting their distribution. *Limnology and Oceanography*, **41**(8), 1669–1683.

Bower, A. S., L. Armi, and I. Ambar (1997), Lagrangian observations of meddy formation during a Mediterranean undercurrent seeding experiment. *J. Phys. Oceanogr.,* **27**(12), 2545–2575.

Bower, A. S., D. M. Fratantoni, W. E. Johns, and H. Peters (2002), Gulf of Aden eddies and their impact on Red Sea water, *Geophys. Res. Letters*, **29**(21), 2025, doi: 10.1029/2002GL015342.

Bracco, A., J. LaCasce, C. Pasquero, and A. Provenzale (2000), The velocity distribution of barotropic turbulence. *Phys. Fluids*, **12**, 2478–2488.

Brandt, P., F. A. Schott, A. Funk, and C. S. Martins (2004), Seasonal to interannual variability of the eddy field in the Labrador Sea from satellite altimetry. *J Geophys. Res.*, **109**(C02028), doi: 10.1029/2002JC001551.

Broerse, A. T. C., T. Tyrrell, J. R. Young, A. J. Poulton, A. Merico, W. M. Balch, and P. I. Miller (2003), The cause of bright waters in the Bering Sea in winter. *Continental Shelf Res.*, **23**(16), 1579–1596.

Brown, C. W., and J. A. Yoder (1994), Coccolithophorid blooms in the global ocean. *J. Geophys. Res.*, **99**(C4), 7467–7482.

Caniaux, G., L. Prieur, H. Giordani, F. Hernandez, and L. Eymard (2001), Observation of the circulation in the Newfoundland Basin in winter 1997. *J. Phys. Oceanogr.*, **31**(3), 689–710.

Chapron, B., F. Collard, and F. Ardhuin (2005), Direct measurements of ocean surface velocity from space: Interpretation and validation. *J. Geophys. Res.*, **110**(C07008), doi: 10.1029/2004JC002809.

Chelton, D. B., R. A. De Szoeke, M. G. Schlax, K. El Naggar, and N. Siwertz (1998), Geographical variability of the first-baroclinic Rossby radius of deformation. *J. Phys. Oceanogr.*, **28**, 433–460.

Chelton, D. B., M. G. Schlax, R. M. Samelson, and R. De Szoeke (2007), Global observations of large oceanic eddies. *Geophys. Res. Letters*, **34**(L15606), doi: 10.1029/2007GL030812.

Ducet, N., P. Y. Le Traon, and G. Reverdin (2000), Global high-resolution mapping of ocean circulation from the combination of T/P and ERS-1/2. *J. Geophys. Res.*, **105**(C8), 19477–19498.

Fratantoni, D. M., A. S. Bower, W. E. Johns, and H. Peters (2006), Somali current rings in the eastern Gulf of Aden, *J. Geophys. Res.*, **111**(C09039), doi: 1029/2005JC003338.

Fu, L.-L. (2007) Interaction of mesoscale variability with large-scale waves in the Argentine Basin. *J. Phys. Oceanogr.*, **37**(3), 787–793.

Fu, L.-L., and A. Cazenave (Eds.) (2001), *Satellite Altimetry and Earth Sciences* (463 pp.). Academic Press, San Diego, CA.

Gopalan, A. K. S., V. V. G. Krishna, M. M. Ali, and R. Sharma (2000), Detection of Bay of Bengal eddies from TOPEX and in situ observations. *J. Marine Res.*, **58**(5), 721–734.

Henson, S. A., and A. C. Thomas (2007), A census of oceanic anticyclonic eddies in the Gulf of Alaska. *Deep-Sea Res I*, **55**(2), 163–176.

Holligan, P. M., and S. B. Groom (1986), Phytoplankton distributions along the shelf break. *Proc. R. Soc. Edinburgh, Sect. B, Biol. Sci.*, **88**, 239–263.

Holligan, P. M., M. Viollier, D. S. Harbour, P. Camus, and M. Champagne-Philippe (1983), Satellite and ship studies of coccolithophore production along a continental shelf edge. *Nature*, **304**, 339–342.

Isern-Fontanet, J., E. García-Ladona, and J. Font (2003), Identification of marine eddies from altimetry. *J. Atmos. Oceanic Tech.*, **20**, 772–778.

Isern-Fontanet, J., J. Font, E. Garcia-Ladona, M. Emilianov, C. Millot, and I. Taupier-Letage (2004), Spatial structure of anticyclonic eddies in the Algerian basin (Mediterranean Sea) analyzed using the Okubo–Weiss parameter. *Deep-Sea Res. II*, **51**(25/26), 3009-3028.

Isern-Fontanet, J., E. Garcia-Ladona, and J. Font (2006), Vortices of the Mediterranean Sea: An altimetric perspective. *J. Phys. Oceanogr.*, **36**(1), 87–103.

Johannessen, J. A., R. Shuchman, G. Digranes, D. Lyzenga, C. Wackerman, O. Johannessen, and P. Vachon (1996), Coastal ocean fronts and eddies imaged with ERS 1 synthetic aperture radar. *J. Geophys. Res.*, **101**(C3), 6651–6667.

Johannessen, J. A., V. Kudryavtsev, D. Akimov, T. Eldevik, N. Winther, and B. Chapron (2005), On radar imaging of current features: 2, Mesoscale eddy and current front detection, *J. Geophys. Res.*, **110**(C07017), doi: 10.1029/2004JC002802.

Kudryavtsev, V., D. Akimov, J. A. Johannessen, and B. Chapron (2005), On radar imaging of current features: 1, Model and comparison with observations. *J. Geophys. Res.*, **110**(C07016), doi: 10.1029/2004JC002505.

Le Traon, P.-Y., and G. Dibarboure (1999), Mesoscale mapping capabilities of multiple-satellite altimeter missions. *J. Atm. Ocean. Tech.*, **16**, 1208–1223.

Le Traon, P.-Y., F. Nadal, and N. Ducet (1998), An improved mapping method of multi-satellite altimeter data. *J. Atmos. Oceanic Tech.*, **15**, 522–534.

McDowell, S. E., and H. T. Rossby (1978), Mediterranean water: An intense mesoscale eddy off the Bahamas. *Science*, **202**, 1085–1087.

Mouriño, B., E. Fernández, J. Escánez, D. de Armas, S. Giraud, B. Sinha, and R. Pingree (2002), A Subtropical Oceanic Ring of Magnitude (STORM) in the Eastern North Atlantic: Physical, chemical and biological properties. *Deep-Sea Res.*, **49**(19), 4003–4021.

Munk, W., L. Armi, K. Fischer, and F. Zachariasen (2000), Spirals on the sea, *Proc. R. Soc. Lond. A*, **456**(1997), 1217–1280.

Okkonen, S. R., G. A. Jacobs, E. J. Metzger, H. E. Hurlburt, and J. F. Shriver (2001), Mesoscale variability in the boundary currents of the Alaska Gyre. *Continental Shelf Res.*, **21**(11/12), 1219–1236.

Okubo, A. (1970) Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences. *Deep-Sea Res. I*, **17**, 445–454.

Palastanga, V., P. J. van Leeuwen, M. W. Schouten, and W. P. M. de Ruijter (2007), Flow structure and variability in the subtropical Indian Ocean: Instability of the South Indian Ocean Countercurrent. *J. Geophys. Res.*, **112**(C01001), doi: 10.1029/2005JC003395.

Pasquero, C., A. Provenzale, and A. Babiano (2001), Parametrization of dispersion in two-dimensional turbulence. *J. Fluid Mech.*, **439**, 279–303.

Romeiser, R., and W. Alpers (1997) An improved composite surface model for the radar backscattering cross section of the ocean surface: 2, Model response to surface roughness variations and the radar inmaging of underwater bottom topography. *J. Geophys. Res.*, **102**, 25251–25267.

Romeiser, R., W. Alpers, and V. Wismann (1997), An improved composite surface model for the radar backscattering cross section of the ocean surface: 1, Theory of the model and optimization/validation by scatterometer data. *J. Geophys. Res.*, **102**, 25237–25250.

Romeiser, R., S. Ufermann, and W. Alpers (2001), Remote sensing of oceanic current features by synthetic aperture radar—achievements and perspectives. *Ann. Télécommun.*, **56**(11/12), 661–671.

Saraceno, M., C. Provost, A. R. Piola, J. Bava, and A. Gagliardini (2004), Brazil Malvinas Frontal System as seen from 9 years of advanced very high resolution radiometer data. *J. Geophys. Res.*, **109**(C05027), doi: 10.1029/2003JC002127.

Saraceno, M., C. Provost, and A. R. Piola (2005), On the relationship between satellite-retrieved surface temperature fronts and chlorophyll a in the western South Atlantic. *J. Geophys. Res.*, **110**(C11016), doi: 10.1029/2004JC002736.

Schonten, M. W., W. P. M. de Ruijter, P. J. van Leeuwen, and J. R. E. Lutjeharms (2000), Translation, decay and splitting of Agulhas rings in the southeastern Atlantic Ocean. *J. Geophys. Res.*, **105**(C9), 21913–21925.

Scully-Power, P. (1986), *Navy Oceanographer Shuttle Observations, STS 41-G: Mission Report* (Tech. Rep. NUSC TD 7611, 71 pp.). Nav. Underwater Syst. Cent, Newport, RI.

Smyth, T. J., T. Tyrrell, and B. Tarrant (2004), Time series of coccolithophore activity in the Barents Sea, from twenty years of satellite imagery. *Geophys. Res. Letters*, **31**(L11302), doi: 10.1029/2004GL019735.

Stanton, B., and M. Y. Morris (2004) Direct velocity measurements in the Subantarctic Front and over Campbell Plateau, southeast of New Zealand. *J. Geophys. Res.*, **109**(C01028), doi: 10.1029/2002JC001339.

Stewart, R. H. (2008) *Introduction to Physical Oceanography* (e-book). Texas A & M University, available at *http://oceanworld.tamu.edu/home/course_book.htm*

Vallis, G. K. (2006), *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation* (745 pp.). Cambridge University Press, Cambridge, U.K.

van Leeuwen, P. J., W. P. M. de Ruijter, and J. R. E. Lutjeharms (2000), Natal pulses and the formation of Agulhas rings. *J. Geophys. Res.*, **105**(C3), 6425–6436.

Waugh, D. W., E. R. Abraham, and M. M. Bowen (2006), Spatial variations of stirring in the surface ocean: A case study of the Tasman Sea. *J. Phys. Oceanogr.*, **36**(3), 526–542.

Weiss, J. (1991), The dynamics of enstrophy transfer in two-dimensional hydrodynamics. *Physica*, **D48**, 273–294.

4

Mesoscale ocean features: Fronts

This is the second of three chapters about the remote sensing of mesoscale ocean features, and is best read in conjunction with Chapters 3 and 5. The former provides a general introduction to mesoscale processes and then focuses on observing ocean eddies from space. The latter shows how upwelling is monitored by satellites and considers further applications of remote sensing for studying various mid-scale ocean-dynamical features. The subject of this chapter is the remote sensing of ocean fronts. It starts in Section 4.1 with a descriptive explanation of how fronts are maintained in dynamic equilibrium and in Section 4.2 considers the characteristics of fronts that allow them to be observed from space, reviewing various remote-sensing methods and image enhancement techniques. Section 4.3 describes an analytical method that automatically tracks frontal location and characteristics. The rest of the chapter shows examples of how our present understanding of fronts has benefited from the knowledge and insights provided by satellite data, looking at the global distribution and climatology of major ocean fronts (Section 4.4), at mesoscale frontal variability (Section 4.5), and in Section 4.6 the role of fronts in promoting biological primary production.

4.1 BOUNDARIES IN THE OCEAN

Fronts in the ocean are sharp horizontal boundaries between water masses of different density. Their occurrence is notable because they defy a tendency in the ocean for water properties to approach uniformity on horizontal surfaces. Although very strong vertical gradients can be found in the stratified ocean which is statically stable as density increases with depth, even the smallest horizontal density gradients create horizontal pressure gradients which cannot be balanced by gravity. Consequently the presence side by side of two water types of different density will induce motion. In general the resulting currents tend to advect the different water

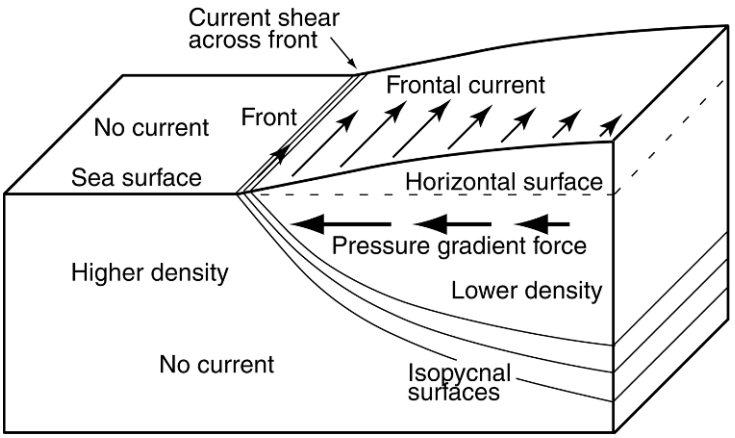


Figure 4.1. Section through an ocean front (northern hemisphere). Note that in this schematic the slope of the sea surface is greatly exaggerated relative to that of isopycnals.

masses so that horizontal density gradients are reduced (e.g., in the absence of other forces the denser fluid will flow underneath the lighter fluid). The currents also promote mixing between the two water masses which can ultimately homogenize their densities. Therefore if frontal boundaries are to be maintained without erosion another force is required. In the ocean this is the Coriolis force which, as we have already noted in Section 3.2, can oppose pressure forces induced by the density gradient while allowing currents to flow along contours of density. The presence of currents is essential in a frontal system, but in a geostrophically balanced front they flow parallel to the density interfaces and so do not directly deform or destroy them.

Figure 4.1 illustrates schematically a typical ocean front that is in dynamic equilibrium. A wedge of less dense water is seen overlying a uniform denser body of water in which there is no motion. In order that there are no horizontal pressure gradients in the deeper water below this wedge, the sea surface above it must slope up away from the front. Consequently the pressure increases from left to right in the upper water body, producing a pressure force from right to left. This pressure force does not occur in the lower fluid, being eliminated by isopycnals sloping in the opposite direction. In the upper fluid the pressure force can be balanced by the Coriolis force associated with currents flowing parallel to the front on the less dense side. Current direction drawn in Figure 4.1 is that for the northern hemisphere; it would reverse for fronts occurring south of the Equator. The whole of the wedge of lighter fluid must be in motion parallel to the front. The highest speed is found close to the front where isopycnals slope down most steeply and the surface slope is also at its maximum. The current gradually decreases to the right and decays away to zero where the slopes of the isopycnals and the sea surface tend to zero (i.e., they are horizontal).

This basic frontal structure applies to all scales of ocean fronts as long as they are sustained long enough to reach geostrophic balance. The largest scale fronts extend for thousands of kilometers and their frontal flows form the major ocean

currents such as the Gulf Stream in the Atlantic Ocean, the Kuroshio in the Pacific and the Antarctic Circumpolar Current around the Southern Ocean. More ephemeral fronts on smaller lengthscales come and go within the general mesoscale turbulence of the ocean. For example, as mesoscale eddies are deformed the isopycnals will be squeezed closer in some places and fronts will evolve dynamically. Certain fronts are the consequence of sources of lower or higher density water. For example, cold, upwelled water may create a sharp interface with its surroundings in the form of a front, or warm water discharging from a large river or estuary may establish a front where it meets the cooler water of the open ocean. Because they are readily observed by satellites, our knowledge of the types and distribution of fronts has benefited greatly from remote sensing, as the rest of this chapter will show.

Some other aspects of fronts are also worth mentioning because they are relevant to frontal detection from space. The schematic in Figure 4.1 is simplified in showing currents only on one side of the front. A steady flow can be added to the whole figure without changing the basic mechanism of frontal equilibrium, or the main current could be located on the denser side of the front. Furthermore, the basic mechanism outlined above represents the first-order, steady-state, dynamical description, whereas in reality there are other, second-order processes which may cause the front to gradually change. Some of these are illustrated in Figure 4.2. For example, a front that is free of topographic constraints may start to curve and develop large lateral oscillations; in extreme cases the resulting meanders can grow until they pinch off to form free eddies (e.g., Figure 3.16 shows Gulf Stream rings being created). At shorter lengthscales the velocity shear across the front can develop local instabilities which promote mixing across the front, tending to reduce density gradients. Localized redistribution of the density perturbs the pressure field causing departures from the basic geostrophic balance and promoting secondary flows. These may converge or diverge at the front. Surface convergence is required for

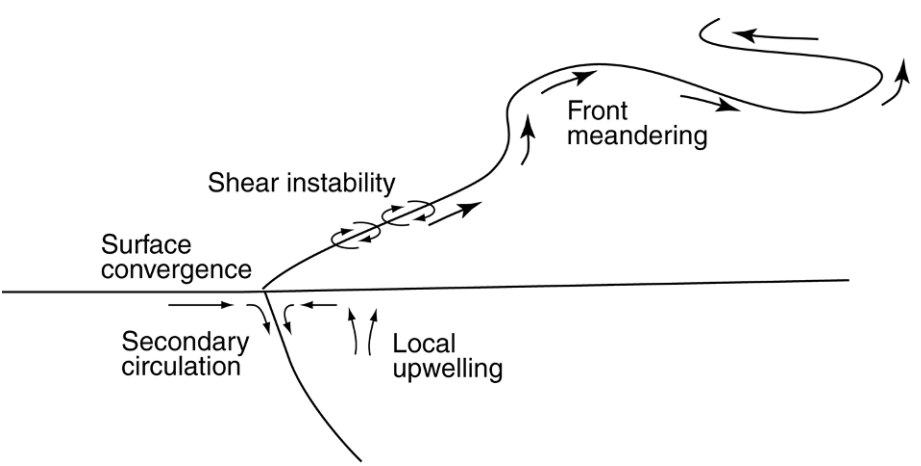


Figure 4.2. Secondary dynamical processes that may occur at ocean fronts.

frontogenesis because it has the effect of steepening horizontal density gradients. It may also modify surface roughness at the front line. Divergence would result in upwelling at the front, with the possibility of promoting primary production.

4.2 THE REMOTE-SENSING SIGNATURES OF OCEAN FRONTS

For fronts to be detected from space, some aspect of their structure must affect one of the four primary observable quantities measured by ocean remote sensing. In fact the structure and processes of fronts can, in the right circumstances, affect each of these quantities; surface temperature, color, surface roughness, and sea surface slope. Signatures of fronts can therefore be found in the data from the several different sensor types described in this section.

4.2.1 Sea surface temperature signatures of fronts

The appearance of fronts on infrared images

It is in SST images that fronts have been most readily observed by satellite remote sensing. The two properties which control the density of seawater and define the different water masses which abut one another at fronts are temperature and salinity. Many, but not all, ocean fronts are characterized by a strong temperature gradient which is immediately detectable in infrared images. Figure 3.16 in the previous chapter shows the Gulf Stream front, between North Atlantic Gyre water and U.S. Shelf water, as a temperature difference of around 10 K at the point where it detaches from the North American coast. It gradually weakens as the flow develops large meanders as it progresses across the Atlantic.

Another example of a very strong thermal front is the Agulhas Current where it detaches from the coast as it flows south-westward from the Indian Ocean and starts to turn west round the southern tip of Africa. The MODIS (Aqua) SST image in Figure 4.3 shows a rare cloud-free view of this region. Evidently a spatial resolution of about 1 km is needed to show the detachment point so precisely and to define the extremely steep front which spans about 4 K to 6 K in as many kilometers. It also reveals most precisely the gradual and then sudden growth of instabilities that grow on the front after detachment, and which eventually lead to the spawning of Agulhas rings (van Leeuwen *et al.*, 2000).

It is not only major ocean fronts that can be detected by infrared remote sensing. Some medium-resolution infrared radiometers operating today have a thermal resolution approaching 0.1 K. Although after atmospheric correction the absolute accuracy of some satellite SST systems may be much poorer than this, brightness temperature images from a single channel within the 10.5 μm to 12.5 μm spectral window should be capable of delineating frontal structures when the temperature contrast across the front is as small as 0.2 K.

The typical spatial resolution of infrared radiometers is between 1 km and 2 km, depending on whether the pixel is below the satellite at the center of a swath or is

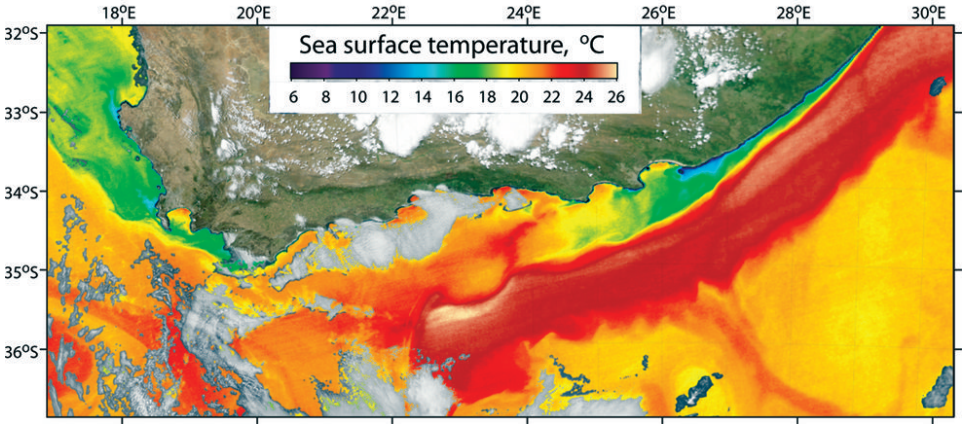


Figure 4.3. SST image of the front where the warm Agulhas Current detaches from the east African coast. The SST field was derived from the infrared wavebands of the MODIS sensor on NASA's EOS Aqua satellite, for an overpass at 12:40 UT on December 30, 2007. It corresponds to the chlorophyll image shown in Figure 4.8.

being viewed obliquely towards the swath extremity. This is very much smaller than the typical along-front length of mesoscale fronts in the open ocean, and comparable with their cross-frontal dimension, typically a few hundred meters to a few kilometers depending on the scale of the front. Therefore infrared images are able to detect small-scale fronts that form within mesoscale turbulent eddy fields, as well as intermediate fronts that are sometimes found to subdivide large-scale ocean fronts. Figure 4.4, a single-channel brightness temperature image from ATSR, shows these features. A major front is evident at A, but in places it is not a single front but a series of steps (B). There is also in this case a pool of cool water along the front (C). The main temperature difference is between water at about 19°C (white tone on the image) and 15°C (gray). Between them the temperature falls to 12°C (black). This is most likely to be a result of upwelling induced by the front. Elsewhere there are more examples of interleaved frontal structures, where there are alternate bands of cooler and warmer water (D), and small-scale fronts that have developed on the flank of smaller scale eddies (E).

However, temperature data displayed in this type of gray-tone contour image is not the best way to visualize fronts. In a display such as Figure 4.4, whose contrast stretch is quite weak (it varies linearly from 11.5 to 19.5°C), the eye can discern only the stronger fronts. Careful inspection will reveal a number of weaker fronts that do not stand out clearly, and it is likely that some even weaker fronts are not visible at all in this image. If color-contoured images are used, such as Figures 3.16 or 4.3, the eye is drawn to strong color contrasts but still misses the subtle color steps across weaker fronts. Therefore different image enhancement is needed to identify all the fronts that may be present.

What distinguishes a front from a gradual change is the occurrence of a near-discontinuity in temperature, a thermal gradient that is much stronger than else-

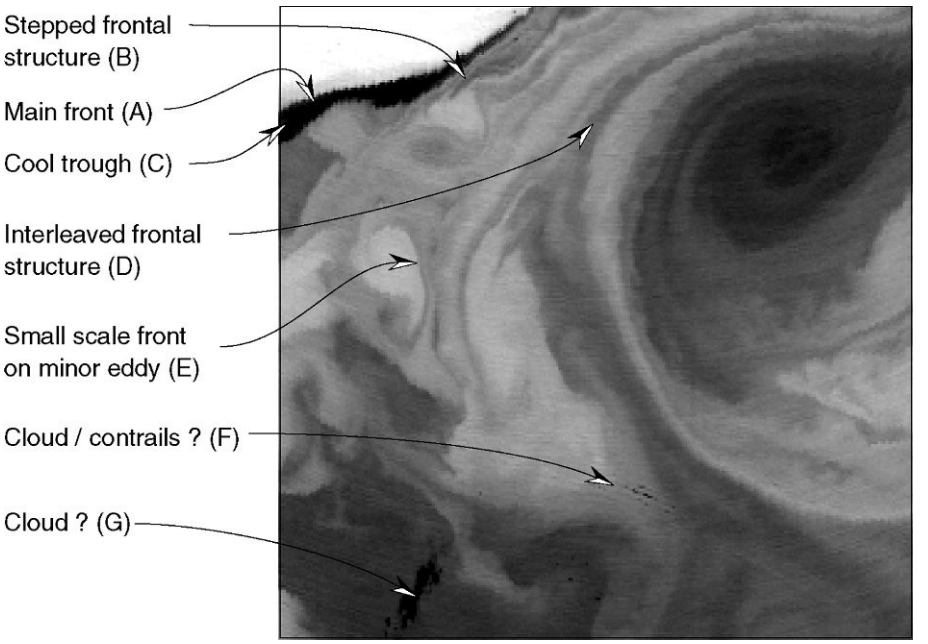


Figure 4.4. Part of a brightness temperature image from the ATSR 10.3 μm to 11.3 μm channel. This image shows a 250 km square region of the South Atlantic Ocean off Argentina in the confluence zone between the Brazil and Falkland currents. Lighter tones indicate warmer water. (The original ATSR data were provided by the European Space Agency.)

where in the surrounding part of the image. Such a discontinuity in temperature implies the presence of a current shear zone, which has oceanographic importance in terms of secondary circulation, upwelling, and mixing between the two water masses. Thus in Figure 4.4 where the front is stepped, the current must be flowing in several parallel bands within which the flow is uniform, but with shear layers between them. In the case of interleaved bands, the layers are probably alternately faster and slower. Fronts such as (E) that form on a small eddy may not have a very large temperature contrast, but their presence is a prelude to further instability that will spawn even smaller scale eddies—part of the process by which large water masses eventually mix towards homogeneity. That is why it is important to be able to identify fronts at all scales, since their existence is an indicator of how this process is proceeding.

Visualizing fronts using high-pass filters

Identifying and highlighting fronts in a temperature image requires application of a suitable high-pass filter. Figure 4.5 shows examples of four such filters applied to the image in Figure 4.4 (see section 5.3.6 of *MTOFS* for more information about the

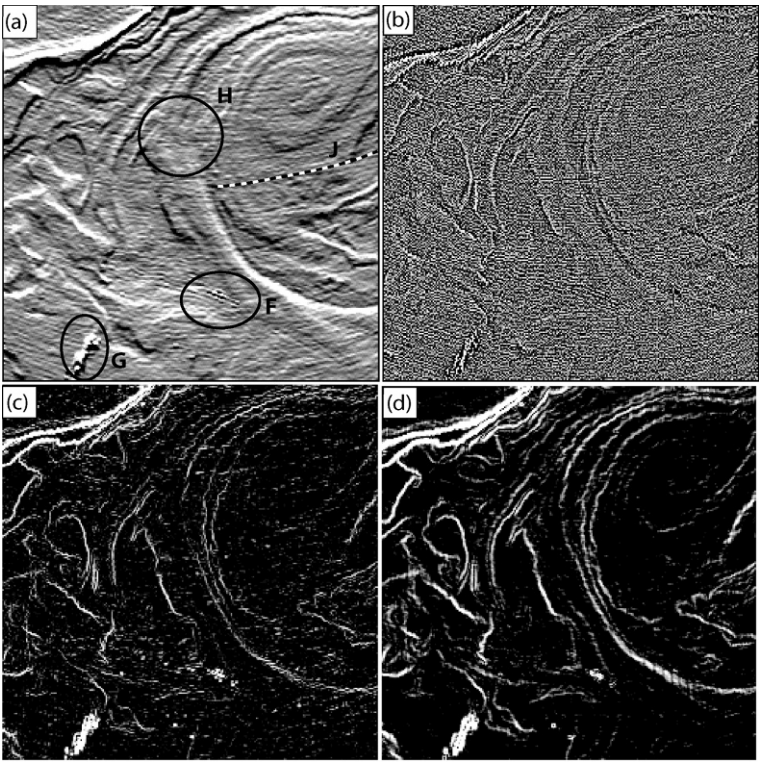


Figure 4.5. Application of high-pass digital filters to the image in Figure 4.4, to enhance visualization of ocean fronts. (a) 3×3 gradient filter, responding to gradients with a component up and down the image. (b) 3×3 Laplacian operator. (c) Roberts (2×2) gradient filter. (d) Sobel filter.

four filters used). Each of these filters, in its own way, emphasizes discontinuities and short-lengthscale variability in the temperature field.

The vertical gradient filter (a) consists of signed (positive and negative) values which represent the direction of the gradient, with bright zones where the temperature increases up the image and dark when it increases downwards. This gives the impression of viewing a three-dimensional surface. It is very effective in discriminating between the positive and negative slopes where the front is interleaved in features B and D of Figure 4.4. It also places emphasis on small-scale fronts as much as large ones, showing that they are often just as steep although they are narrower. For example, feature E is mostly well displayed except where the front lies exactly parallel with the image vertical co-ordinate. Here the filter detects nothing, and this is the serious weakness with this filter. The region labeled H in Figure 4.5a shows this problem at its worst. The signature of the circular frontal system surrounding the large eddy is lost at this point and changes in phase, so that a misleading impression is given.

Figure 4.5b is a Laplacian filter which effectively shows the spatial second differential, responding most strongly to isolated peaks or troughs. It also is signed. It provides a lot of fine detail but fails to emphasize the largest fronts where there is a sustained gradient of temperature over a width of several pixels.

Both this and the one-dimensional gradient filter show up regular circular arcs, parallel to the dashed line labeled J in Figure 4.5a. These correspond to the scan geometry of the ATSR sensor and represent instrument noise appearing in this single-band image. This noise occurs at a level of ± 0.1 K, showing how sensitive the filters are. However, this does not present a problem for the detection of fronts, unless they happen by chance to coincide with the scan geometry. What allows the fronts to be detected even when their signature is comparable in magnitude with sensor noise is that they are characterized by a coherent curvilinear form. This is readily differentiated from sensor noise which is either random or structured in a regular geometrical form.

What may be more of a problem in infrared images is the effect of undetected cloud or subpixel cloud. In fact no cloud detection was applied to this image and so there is a danger that cloud-related artifacts may be present, although the image is evidently almost cloud-free. The parts labeled F and G present features that are not consistent with the rest of the image. At both there is a dark area on the original temperature image and the structure also appears on filtered images. The first two filters both show that F is part of a wider feature consisting of four or five parallel lines stretching nearly 150 km, more than halfway across the image. There is no obvious oceanographic explanation, and it is reasonable to treat this is as an atmospheric phenomenon, with subpixel cloud or enhanced atmospheric water vapor influencing brightness temperature for a few pixels. One possible explanation for the regularity is that this is a commercial air route and these are the accumulated condensation trails from many aircraft. Whatever the reason, this illustrates the need to be vigilant when interpreting SST imagery. The feature at G is more likely to be a region of clouds since it shows no obvious oceanographic structure at all.

The two gradient filters in Figures 4.5c, d are similar to each other in that they represent an absolute value of the two-dimensional spatial gradient. Thus they cannot differentiate between the direction of the gradient, but they both result in a pattern of bright lines in a dark background. Comparison with the other filters and the original image confirm that the bright lines correspond to fronts. The Sobel filter in Figure 4.5d seems to discriminate the fronts most successfully.

What is remarkable is how coherent front lines remain over hundreds of kilometers, even as they spiral in towards the center of the large eddy. Several fronts remain parallel for long distances although it also appears that in some places the spacing between adjacent fronts starts to oscillate, as though some sort of instability is developing. There is potential for more detailed study of image data like this, in conjunction with fluid-dynamical analysis of the processes represented. Concerning which filter to choose for visualizing the fronts, it is probably best to use more than one and combine the insights from several, as demonstrated here. Note that Section 4.3 will present a different approach to the automatic detection and quantitative characterization of fronts.

Problems with cloud and other limitations

As with all applications of thermal infrared data, the view of ocean fronts is

obscured by the presence of clouds. There are some regions (e.g., the Iceland–Faeroes Front in the North Atlantic) where the cloud cover climate is such that it would be futile to expect infrared data to provide reliable regular information about the precise location of the front, which is disappointing for certain operational applications. Instead the best that infrared remote sensing can offer in such a case is occasional clear overpasses which can at least show typical locations and morphologies that the front achieves. This in itself would serve to complement and test the use of numerical models for forecasting frontal dynamics.

There are of course many other parts of the World Ocean where cloud cover is not such a serious issue. However, even when there are no clouds obscuring a front, the user of SST maps derived from infrared sensors must beware of a potential problem caused by the character of the SST field clashing with automatic cloud detection software. One of the standard tests for cloud cover (see section 7.2.4 of *MTOFS*) is to examine the spatial coherency of a block of pixels. If spatial coherency is low (i.e., the temperature variance is high), it is assumed that there is patchy cloud and all the pixels in the block are rejected. The threshold variance for this test is generally set quite low, on the assumption that sea temperature varies smoothly and a cloud-free scene would have low variance. Fronts are an exception to this, and are likely to trigger cloud detection unless steps are taken to suppress it in a region where fronts are expected. Otherwise, the unwary user will find that clouds always seem to congregate where fronts are expected, even when the rest of the sea is completely cloud-free! Of course the cost of relaxing the spatial coherence test by setting a higher threshold is that the SST field tends to be corrupted more often by undetected cloudy pixels. To illustrate the problem, Figure 4.6 shows the result of applying a spatial variance filter to the temperature image used in Figure 4.4. Using a typical spatial coherency threshold the white zones would all be flagged as cloud, which is obviously

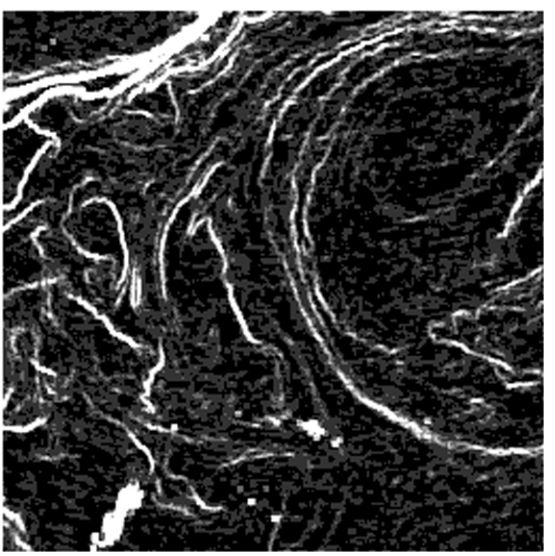


Figure 4.6. Application of a spatial variance filter to the image shown in Figure 4.4.

not the case in this image. On the other hand, relaxing the threshold might prevent the cloud contamination in regions F and G from being detected.

Another reason why smaller and weaker fronts may not always show up as clearly as might be expected is that they may be temporarily covered by a surface mixed layer or diurnal thermocline. Although underlying water masses will eventually influence the properties of the overlying surface layer, in periods of solar heating and low wind, when the surface layer is shallow, buoyant, and disconnected from the water below, the patterns painted in satellite SST images may say more about wind mixing than about underlying water masses. In such situations, nighttime images are likely to give a better representation of frontal structures.

Where cloud cover remains a serious problem for infrared data from polar-orbiting sensors, an alternative is to use infrared sensors on geostationary platforms. Because these deliver many images per day of the same region there is more chance of detecting frontal locations under partly cloudy conditions, allowing some interesting frontal studies to be performed by this means (Legeckis *et al.*, 2002). Another alternative is to use microwave radiometry. The drawback with this (see Section 2.4.4 in this book and chapter 8 of *MTOFS*) is the much poorer spatial resolution. Even when oversampled, the typical pixel size of microwave SST images is 25 km. Thus microwaves should be considered useful for observing major ocean fronts, but not for identifying small-scale frontal structures. Comparison between Figures 3.18 and 4.3 provides a graphic confirmation of this. In the case of the Agulhas detachment region, the microwave radiometer is further disadvantaged by the poor quality of its temperature retrieval up to 100 km from the coast, flagged as no data in Figure 3.18. Nonetheless the capacity to update the view every day or two, with very few open-ocean pixels flagged as invalid, allows a detailed time history of mesoscale features to be tracked very well (see Section 3.5.2) and the positions of smaller fronts can be inferred although their characteristics (such as steepness and width) cannot be measured by microwave radiometry.

Examples of the thermal signatures of fronts, and what can be learned oceanographically from them, are considered in more detail in relation to the climatology of major ocean fronts (Section 4.4) and the variability of mesoscale fronts (Section 4.5). Note also that examples of remotely sensed fronts in shelf seas are discussed in Chapter 13.

4.2.2 Can fronts be detected by altimetry?

In principle, since fronts are associated with strong currents and therefore sea surface slopes (see Figure 4.1) altimetry might seem to be an effective way of observing fronts, independently of cloud cover. However, closer consideration suggests that the altimetric signal of most small-scale fronts is poor. First, the altimeter samples only along-track and so to observe a front the altimeter would need to pass precisely over it, and preferably at right angles to the front. Thus many mesoscale fronts will not be crossed at right angles and may be missed completely. Second, the along-track resolution for SSHA is around 10 km. Thus the entire signature of small-mesoscale fronts would be no more than a small height change between two adjacent SSHA

samples. Without recourse to analyzing a high-resolution two-dimensional field, the linear character of the front would not be detectable. Even if it were possible to detect a height change the coarse spatial resolution would not distinguish between a fast, narrow current or a more gentle and wider current. In short, altimetry does not seem to be well suited to the study of the small-scale dynamic detail of mesoscale fronts.

On the other hand, large ocean frontal systems, and the major currents associated with them, do have a readily measurable altimetry signal. Total geostrophic flow through a section drawn through the ocean is proportional to the absolute dynamical height difference between the two ends of the section (see figure 11.14 of *MTOFS*). If this distance is several hundred kilometers then our existing knowledge of the geoid (determined from GRACE and using drifting buoy measurements of steady currents) allows us to estimate total flow and mean velocity over the section to an acceptable accuracy, but not to identify the frontal edge where that flow may be concentrated. In large ocean currents there are often several separate strands or filaments of stronger current in which the main flow takes place (and where sea surface slope is steepest) separated by zones where the surface slopes less and the current is weaker (see Figure 4.7). Even when an improved definition of the geoid can be obtained using anticipated GOCE gravity data, it is not intended to provide spatial precision better than 50 km to 100 km and so it is not expected to be possible to distinguish the signature of these filamentary, subfront

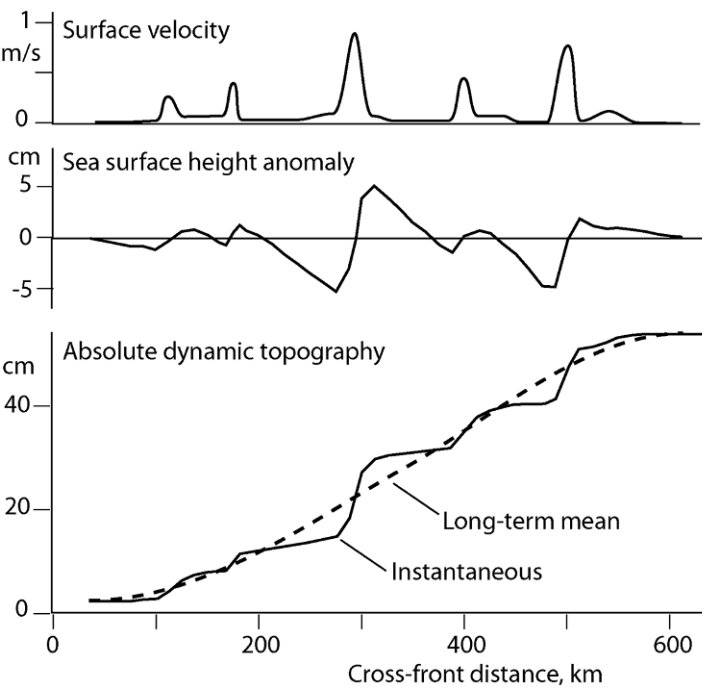


Figure 4.7. Schematic to show the relationship between the filamentary along-front velocity structure in a multiple-core front, associated absolute dynamic topography (ADT), mean dynamic topography (MDT) retrieved from altimetry, a best-fit geoid, and the sea surface height anomaly (SSHA) from altimetry. Slope is scaled to velocity assuming $f = 10^{-4}\,\mathrm{s}^{-1}$.

features in the mean dynamic topography from short-scale fluctuations of the geoid. However, given the inherent turbulent character of large-scale ocean currents, the lateral position of filaments across the frontal zone is likely to shift over months by a few tens of kilometers. Hence their signature will become evident in the SSHA field.

As with the eddies discussed in Chapter 3, so with fronts, the maps of SSHA can be confusing when there is an underlying broad steady current. However, when the absolute dynamic topography (ADT) is constructed by adding the SSHA field to the mean dynamic topography (MDT), the result should provide a satisfactory picture of the absolute current field including the filamentary structure. Figure 4.7 shows some artificial but typical values, scaled for $f = 10^{-4}\,\mathrm{s}^{-1}$ (latitude 43°). Although the MDT cannot be determined to the spatial precision needed to define the subfronts if they were stationary, the fact that finely spaced filaments are continually shifting laterally should cause the true MDT to be rather smooth and therefore quite close to the estimated MDT. This is the basis of recent interesting observations of filamentary structure in the Antarctic Circumpolar Current (ACC) (Sokolov and Rintoul, 2007a), and also in other oceans (Maximenko *et al.*, 2005).

4.2.3 Observing fronts in ocean color images

Some fronts are detectable in ocean color images. Following the discussion in Section 3.6 there are two main reasons for this. The first is that the two different water masses which adjoin each other at a front may have different optical characteristics and therefore different colors, detectable by a multispectral scanner or imaging spectrometer. The second reason is that some fronts provide an opportunity for enrichment of an otherwise nutrient-depleted water body, by upwelling due to secondary circulation in the front or by cross-frontal mixing. Consequently at certain times of the year there may be enhanced primary production along the front which is clearly visible in the chlorophyll data derived from ocean color sensors.

Water masses with inherently different colors

When major ocean fronts have a strong color signature it is generally because there is more productive, greener water on the cool side and less productive, clear, blue water on the warm side. This is not the case for all major fronts, but where it does happen the front is clearly visible in chlorophyll concentration maps derived from visible waveband multispectral sensors. Figure 3.19 shows this clearly in the bottom left corner of this image, where cooler waters by the coast have higher chlorophyll concentrations on the western side of the Gulf Stream front soon after it detaches from the coast. Farther northeast the picture is confused by the development of eddies and the color front no longer closely matches the thermal front.

A similar set of conditions is spectacularly visualized in Figure 4.8 which shows the high chlorophyll concentration inshore of the Agulhas front for the same MODIS overpass whose SST distribution is shown in Figure 4.3. The 1 km resolution visible wavebands are able to show a strong gradient in chlorophyll over a distance which appears to be just slightly wider than the SST front.

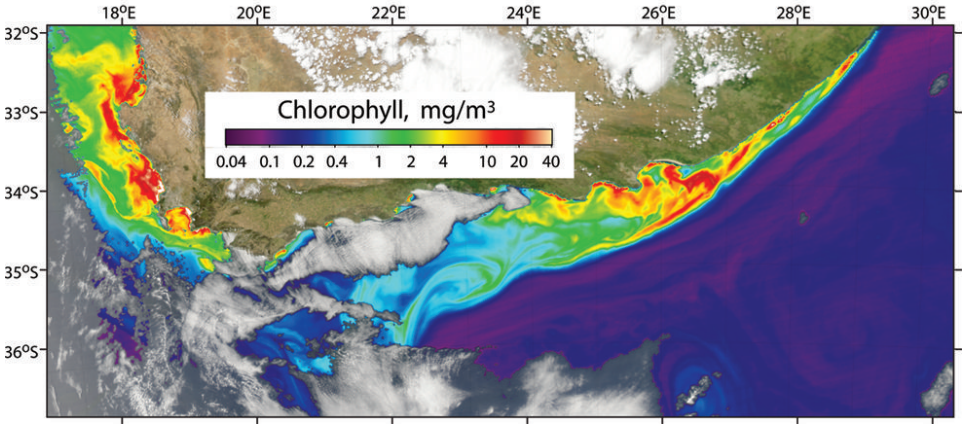


Figure 4.8. An image of chlorophyll concentration revealing the region where the warm Agulhas Current, with low chlorophyll, detaches from the east African coast. The cooler water between the front and the coast is rich in chlorophyll. Chlorophyll distribution was derived from visible wavebands of the MODIS sensor on NASA's EOS Aqua satellite, for an overpass at 12:40 UT on December 30, 2007. It corresponds to the SST image shown in Figure 4.3.

Other typical cases where there is a clearly distinct color difference between the waters on either side of a front are associated with turbid water flowing out of estuaries and rivers into clearer waters of the open sea. River or estuarine water is generally more reflective in the green and red part of the spectrum than clear, blue ocean water, because of the presence of particulate material in suspension, reflecting red and green light that is absorbed in clear ocean water before it can be reflected. In river discharges containing a lot of colored dissolved organic material (CDOM), blue light is preferentially absorbed, making the water appear brown. Fronts associated with river discharge plumes or found at the boundary between shelf sea and open-ocean waters are generally smaller scale than the mesoscale or major ocean fronts considered in this chapter; they will be discussed further in Chapter 13.

However, some general comments can be made here about the front detection capability of color in contrast with thermal satellite imagery. The most important is that whereas most fronts in the ocean have a thermal signature, many do not have a distinctive color signature. Moreover, the color signature is not normally directly related to the density of the water, whereas from knowing the temperature difference the density contrast across a front can be estimated, which has importance for understanding the dynamics of the front. Notwithstanding this general rule, there are exceptional cases (e.g., fronts that are salinity-driven), where there is no thermal gradient, and color contrast signatures are needed if such fronts are to be detected from satellites.

The second point to note about front detection from color rather than temperature is that the location of the color front may not lie on precisely the same line as the surface outcrop of the density interface between the two water masses. This is because the apparent color is an integral of the spectral properties

of the light reflected from the optical depth of the water, which is typically a few tens of meters. If the front slopes at a shallow angle, close to the front on the less dense side (to the right of the front in Figure 4.1), the color observed may still be that of the deeper layer. In the case of the Gulf Stream or Agulhas Current this means that there may be some productive water lying below the surface, but still shallow enough for the ocean color sensor to register higher chlorophyll concentrations, making it appear that high-chlorophyll waters have spread into the warm waters of the main current. In this situation a detailed delineation of the steepest part of the front, as derived from the color signature, may be shifted laterally towards the warm side of the front as determined from the SST signature. Careful comparison of Figures 4.3 and 4.8 shows this to be the case, and can be made very evident if a computer animation is made to switch back and forth between the two images. This is one of those situations where the availability of co-registered simultaneous thermal and visible waveband radiometry from the same sensor, or from sensors on the same platform, is very beneficial.

Enhanced primary production at fronts

The other main type of color signature observed is that of the primary production associated with some fronts. The usefulness of this type of observation is less in relation to identifying where the front is, and more to do with understanding the role of fronts in promoting phytoplankton growth. In this case the combination of thermal and color measurements is powerful—thermal imagery defining where the front outcrops at the surface and the color image revealing where production occurs in relation to it. This is another situation where simultaneous thermal and visible data are needed, not only to assist in co-registration of the data, but also to ensure that if the thermal image is cloud-free then so is the color image, and *vice versa*. A separation of even an hour or two would tend to reduce the frequency of occurrence of cloud-free views for both sensors.

Section 4.6 describes ways in which the study of primary production at fronts has benefited from the unique perspective offered by satellite data. A closely related topic is that of wind-induced upwelling along oceanic coasts or along the Equator. Such upwelling generates its own fronts while providing the location for enhanced biological production., and so could be considered in this chapter. However, the upwelling systems of the ocean are the main topic addressed in Chapter 5 and so are not pursued further here.

Other frontal color signatures

It would be misleading to imply that all ocean color images of fronts can be classified readily into one of the types outlined above. A search through the gallery of interesting images displayed on the websites of the agencies responsible for ocean color sensors will reward the reader with a variety of other interesting frontal systems and processes that may be geographically unique. For example, Figure 4.9 shows the Malvinas or Falkland Current in the southwest Atlantic Ocean, skirting the edge of the Patagonian shelf as it heads north. What is interesting about this image is that

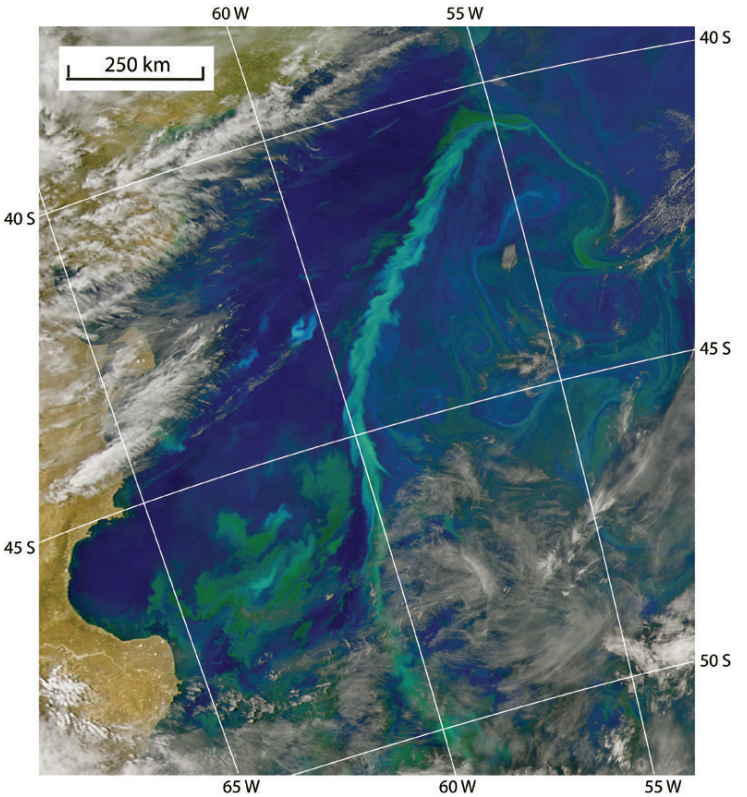


Figure 4.9. The Falklands (Malvinas) current visualized by its ocean color signature. This is a level 1A pseudoreal-color SeaWiFS image acquired over the southwest Atlantic Ocean on December 6, 2004. The geographical grid is only approximate (based on image from NASA Ocean Color website).

whereas major current flows are normally inferred from the frontal boundaries that mark their edge, here the current itself shows up as a different color from its surroundings. This may be associated with enhanced production within the current itself, although the brightness of the reflectance signature in this slightly enhanced real-color image suggests that the current contains a significant quantity of coccolithophores, whose highly reflective liths remain in suspension long after the cells themselves have died (see Section 3.6). In this case, remotely sensed color is serving as a tracer, not of the water masses resident in the region, but of the main current flowing through it. While it is quite common to see streamers of differently colored water marking a localized jet within a complex turbulent flow system such as that shown in Figure 3.21, it is unusual to see the phenomenon persist along a major ocean frontal current for several hundred kilometers, as it does here.

4.2.4 Frontal signatures in radar surface roughness images

It was already noted in Section 3.7 that mesoscale eddies show up in synthetic aperture radar (SAR) images largely because of localized fronts that develop in parts of the eddy, with lengths of less than a few tens of kilometers. Here the

attention focuses on radar roughness signatures of larger scale fronts. Although we now have a fairly good understanding of the various imaging mechanisms by which SAR can visualize certain features in the dynamics of surface currents (see section 10.7 of *MTOFS*) it is never easy to predict from first principles what to expect from SAR images of ocean fronts in any particular situation, as distinct from idealized fronts in model simulations (Ufermann and Romeiser, 1999; Romeiser *et al.*, 2001). With a relatively limited coverage of the world ocean by SAR, we are still at the stage of exploring the data to discover what can be seen (Alpers *et al.*, 1999). Here we shall adopt this empirical approach, presenting some examples of frontal systems that have been successfully revealed by SAR and relating them to established theoretical understanding of SAR imaging mechanisms.

Figure 4.10 provides an example of fronts associated with a major ocean current being revealed in an ERS SAR image. This is in the western North Pacific, off the east coast of Taiwan where we find the northward-flowing Western Boundary

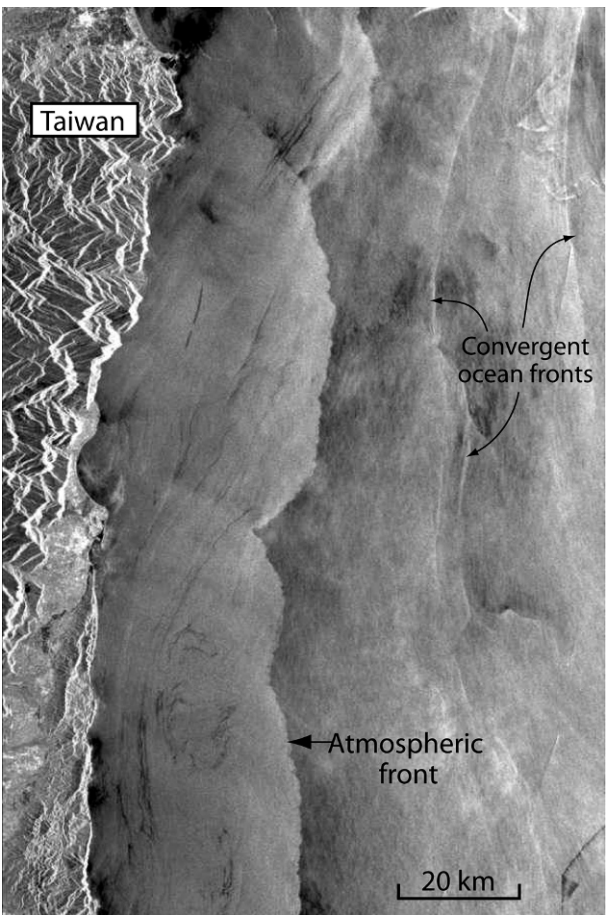


Figure 4.10. ERS SAR image of the region east of Taiwan at about 24°N, showing convergent fronts in the ocean but also a potentially misleading atmospheric front (data provided by ESA).

Current that becomes the Kuroshio Current farther north. There is only a narrow continental shelf here, the sea bed slopes steeply offshore and the major oceanic current system can reach close to the coast. The bright lines parallel to the coast at about 50 km to 80 km offshore are typical of frontal signatures where there is a current shear causing local secondary convergences towards the shear line. The convergence has the effect of enhancing the short wind–wave energy along the shear line, making it stand out in the image. It is reasonable to assume that these convergence lines mark steps in the frontal system, but it is difficult to extract quantitative information with any confidence from an image such as this. Ideally a time sequence of such images is needed before it can be declared that SAR data can monitor the variability of the front and its related flows.

Figure 4.10 also shows a good example of a pitfall when interpreting SAR features. What appears to be the main front, located about 30 km from the coast, does not exhibit any bright lines of convergence. Closer inspection suggests that textural and tonal differences across the boundary are better interpreted as a purely atmospheric effect, associated with the spreading offshore of a land–sea breeze, which have been identified less ambiguously from other SAR images in this region (Alpers *et al.*, 1999).

Figure 4.11 is another ERS SAR image containing evidence of fronts at a somewhat shorter lengthscale than the ocean boundary current in the previous example. In this image each labeled front is identified by its single bright signature extending coherently for up to 50 km but not geometrically straight, in which case an alternative explanation would be sought. Undulations and irregularities in line signatures are characteristic of fronts and increase confidence in interpreting these features as such. The bright signature implies that fronts converge at the surface, assuming that hydrodynamic modulation is the active imaging mechanism. In some cases there is a slight change of radar backscatter from the main bodies of water on either side of a front, which could be associated with different temperatures giving rise to slightly different roughness, as discussed in section 10.7 of *MTOFS*. Note that the regular patterns of groups of parallel bright lines are interpreted here as being internal wave trains (see Chapter 12 for a fuller discussion of measuring internal waves from space). In this image the distinction between lines interpreted as fronts and those assumed to be internal waves is that frontal signatures are sharper, brighter, and narrower than internal waves. They are also found on their own rather than being part of a regular train or group of waves, which is a characteristic of internal waves.

Since rarely are there opportunities for independently verifying an interpretation of dynamical features on SAR images, it is also important to consider whether it is oceanographically reasonable to expect there to be fronts at these locations. The region of sea being viewed is in the northeast Indian Ocean just south of the Indonesian islands of Bali and Lombok. The gap between them, labeled the Lombok Strait, is one of the connections between the Indian Ocean and the Java Sea to the north, and is thus part of the route for water to exchange between the Indian and Pacific Oceans north of Australia. Consequently it is reasonable to expect substantial flows of water through the Lombok Strait. What appears is consistent

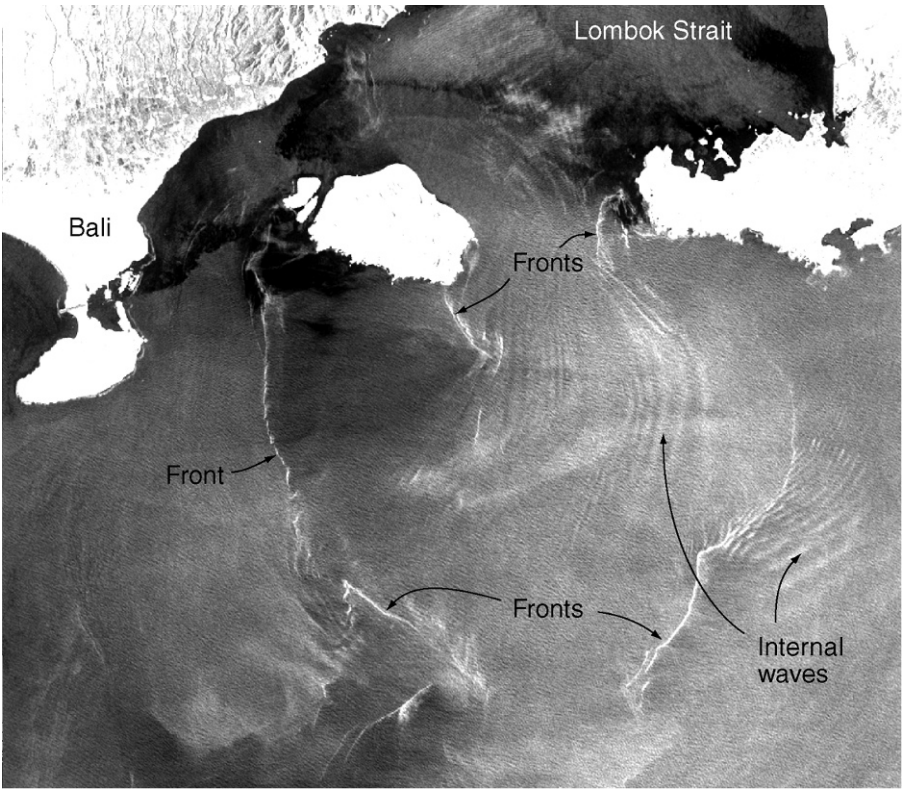


Figure 4.11. ERS SAR image over the Lombok Straits at about 8°S, 116°E, showing a number of local surface convergent fronts. This image is about 100 km wide (produced from digital SAR data provided by ESA).

with water flowing southwards, having a different density to that of the Indian Ocean, and therefore creating fronts at boundaries between outflowing plumes and Indian Ocean waters. The way frontal signatures start close to island headlands is also consistent with this interpretation. Although without further local knowledge these suggestions must be treated as untested hypotheses, they nonetheless illustrate how the detection of fronts on SAR images can raise useful questions and provide insights about the dynamical processes that would support the fronts.

Note also that the apparent relationship between one of the fronts and an internal wave train is quite a common occurrence. It should be recalled that the front is effectively the surface outcrop of a sloping isopycnal. In that situation it appears from examples in many SAR images that flow at the fronts can trigger the generation of internal waves which appear to radiate out from the front.

Finally a third example is shown in Figure 4.12 of a rather different signature of an ocean front. In this case it is a subtle change in texture of the SAR image that marks the front. The rationale for interpreting this as an oceanic front is as follows.

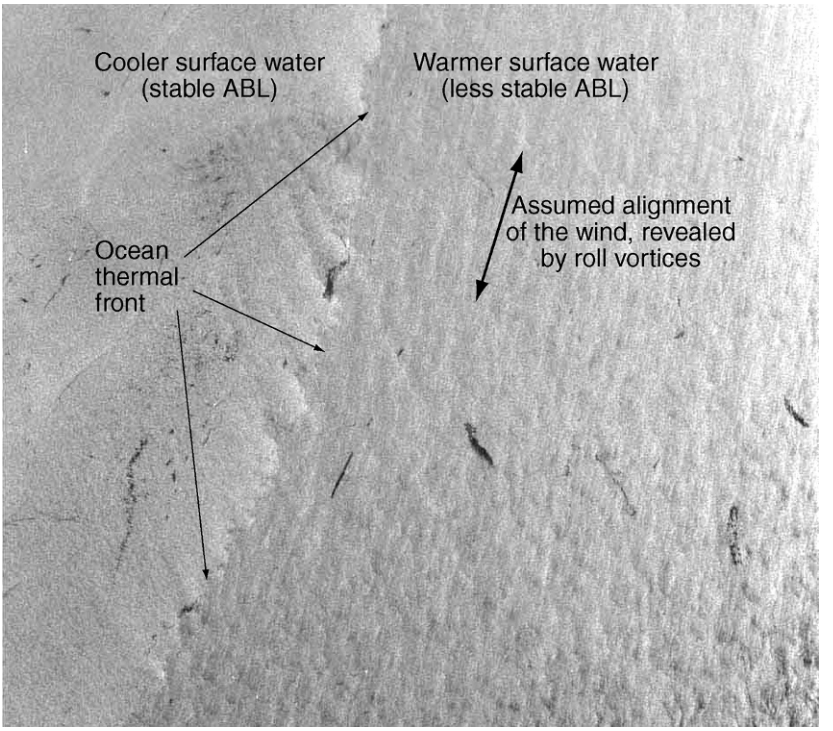


Figure 4.12. This ERS-1 SAR image was acquired on January 7, 1995 over the East China Sea centered at about latitude 28°N, longitude 122.5°E and is 100 km wide. It shows atmospheric roll vortices aligned with the wind in the right half of the field, but not to the left. The line marking the boundary between the two types of image textures is believed to correspond to an ocean front. Warmer water to the east promotes instability in the atmospheric boundary layer which results in roll vortices which give a streaky texture to sea surface roughness (Alpers *et al.*, 1999).

To the left of the front the SST is somewhat cooler than to the right. The atmospheric boundary layer (ABL) is less stable over the warmer water and this allows roll vortices to develop in the ABL, lying parallel with the wind and painting faint striping into the radar backscatter field, aligned with the wind. Over the cooler side of the front this does not occur, changing the image texture and allowing the front to be seen. In this case there is independent knowledge of the SST field (Alpers *et al.*, 1999) which confirms the existence of a thermal ocean front.

However, care must be taken when making an interpretation based on a tone or texture change across the front, because these are more likely to be the characteristics of an atmospheric front. Images like Figure 4.12 do not offer a method capable of systematically detecting ocean fronts since they are so dependent on atmospheric conditions at the time. Note that section 10.8 of *MTOFS* discusses the ways in which atmospheric effects can appear on SAR images of the sea surface.

4.2.5 Direct measurement of currents using Doppler analysis of SAR data

In section 10.4.3 of *MTOFS*, a new method was mentioned for measuring sea surface velocity directly from SARs, based on careful analysis of residual Doppler shifts in the frequency of radar echoes. The principle is simply that the frequency shift between the emitted radar pulse and the echo reflected from the sea surface depends on the relative motion between the satellite and sea surface roughness elements. If the magnitude of this Doppler shift can be measured as the average over all the energy reflected from a given resolution cell of the sea surface (typically a few kilometers in size) during the SAR integration time, relative motion between that patch of sea and the satellite can be determined. Knowledge of satellite velocity and the velocity of the Earth around its axis allows the expected frequency shift to be modeled for each resolution cell on the Earth's stationary surface. Over the ocean any difference between modeled and actual frequency shift is assumed to be caused by the motion of scattering elements on the sea surface. In the case of the sea this is partly due to surface wave orbital velocity which tends to average to zero over the resolution cell apart from some nonlinear effects related to the wave height. What remains after this is taken into account represents U_D, the component of the average velocity of radar-backscattering elements in the direction from the satellite towards the particular resolution cell being examined (as shown in Figure 4.13).

On Envisat, thanks to the very stable satellite orbit and attitude, it has proved possible to estimate precise satellite motion and antenna-pointing contributions to measured frequency shift with sufficient accuracy to distinguish meaningful patterns in U_D. The method has been demonstrated and the practical interpretation and application of the results have now been explored (Chapron *et al.*, 2005), using

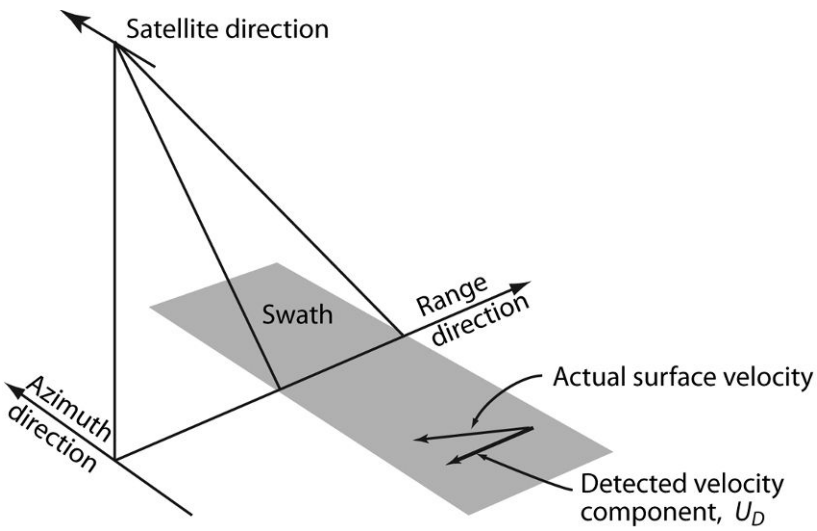


Figure 4.13. Schematic showing the component of apparent surface current, U_D, detected by Doppler centroid analysis of SAR data.

the Envisat ASAR wide-swath mode, for which the σ_0 field has a spatial resolution of 100 m over a swath of about 400 km. Doppler shift is evaluated over cells of 25 km in azimuth and somewhat better resolved (though variable) in range. Chapron *et al.* point out that it should be possible to provide residual Doppler shift as a field within future SAR image products.

The difficulty then lies in interpreting what the retrieved velocity component U_D means. It is the average speed of elements comprising sea surface roughness which reflect microwaves from the radar. Since Bragg waves, a few centimeters in wavelength, scatter most of the echo back to the SAR, it is their speed that controls U_D, although it is not yet fully understood what controls this speed. It obviously depends on the speed of ocean currents in the upper mixed layer of the ocean, which is an oceanographically useful prize to be gained from this method. However, it must also contain the effects of near-surface wind drift which is linked to Stokes drift of the longer waves. U_D is also expected to be modified by effects related to wind–wave current interactions that shape the ocean surface geometry and velocity field at different scales. As a result of Bragg control on relevant ripple wavelengths, U_D will also depend on radar incidence angle and polarization.

For the oceanographer looking for a new way of defining the fine-scale structure of current flows and shear in frontal zones or wanting to determine complex tidal flows in shelf seas, all these uncertainties about how to interpret U_D may seem discouraging. However, the method is still very new and needs to be explored empirically as well as theoretically. For the first published demonstration an Envisat view over the Gulf Stream was selected, since it has a rather large ocean current signature to be detected. The result is shown in Figure 4.14. The measure of U_D has been overlaid as a color-coded field on top of the standard gray-tone SAR image representing the σ_0 field. Some characteristic shear lines are present in the σ_0 image where we expect the Gulf Stream edge to be, where it detaches from the continental slope off Cape Hatteras and heads off into the central Atlantic. A rather coarse view from the altimeter SSHA (not shown here) confirms this. There are strong signatures in the U_D field in the same region, implying that the underlying ocean current is being detected by the Doppler velocity measurement method. However, another σ_0/U_D image (not shown) of the same region when there were strong offshore winds shows patterns of U_D that appear to be responding to the wind field in addition to Gulf Stream currents.

In another recent application of the method shown in Figure 4.15 the sharp frontal edge of the Agulhas Current off South Africa is clearly revealed. Moreover the image seems to show the type of perturbation of the front which can lead to the shedding of Agulhas rings (van Leeuwen *et al.*, 2000). In this image the orientation of the SAR overpass is such that the range direction (in which U_D is measured) lies parallel to the Agulhas Current, providing the strongest possible signal.

This example clearly demonstrates that this method has something to offer the oceanographer, even if only qualitatively at present. Such images can be acquired by Envisat every 2 days on average for any part of the world ocean, and they promise benefits for global ocean monitoring and forecasting systems. Detailed measurements of ocean currents could be used to validate, or even to be assimilated into,

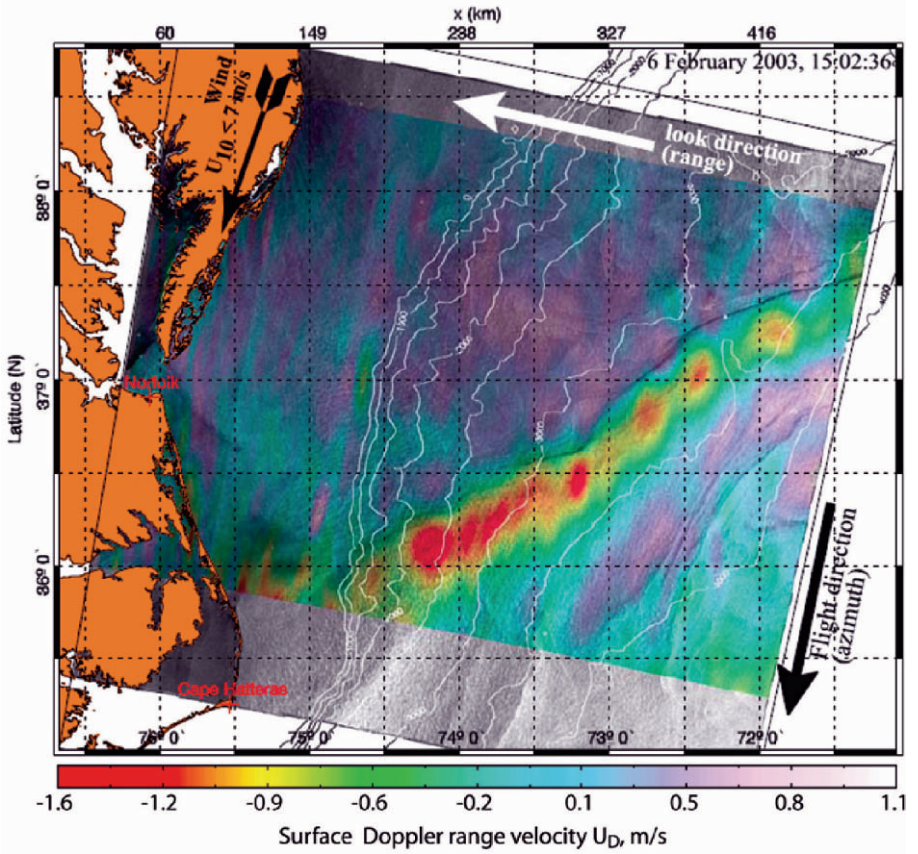


Figure 4.14. Normalized radar cross-section σ_0 (gray shades) and Doppler velocity U_D (colors), analyzed from a wide-swath image obtained by Envisat on February 6, 2003 at 15:12 UTC. Oceanic fronts appear as sharp gradients of σ_0, while surface velocity seen by the radar appears to be related to the Gulf Stream (figure copied from Chapron *et al.*, 2005).

ocean circulation models. However, before the retrieved velocities can be applied quantitatively research effort is required to resolve outstanding questions about how to interpret U_D along with σ_0. It is worth noting finally that this Doppler method used with standard SAR systems is a foretaste of the much higher spatial resolution that could be obtained if along-track interferometric SARs (Romeiser *et al.*, 2005; Romeiser and Runge, 2007) are deployed in future.

4.3 TRACKING FRONTS

4.3.1 Mapping frontal edges

Although the previous section shows that there is a rich variety of ways in which fronts can be observed by remote sensing, the use of thermal infrared remains at

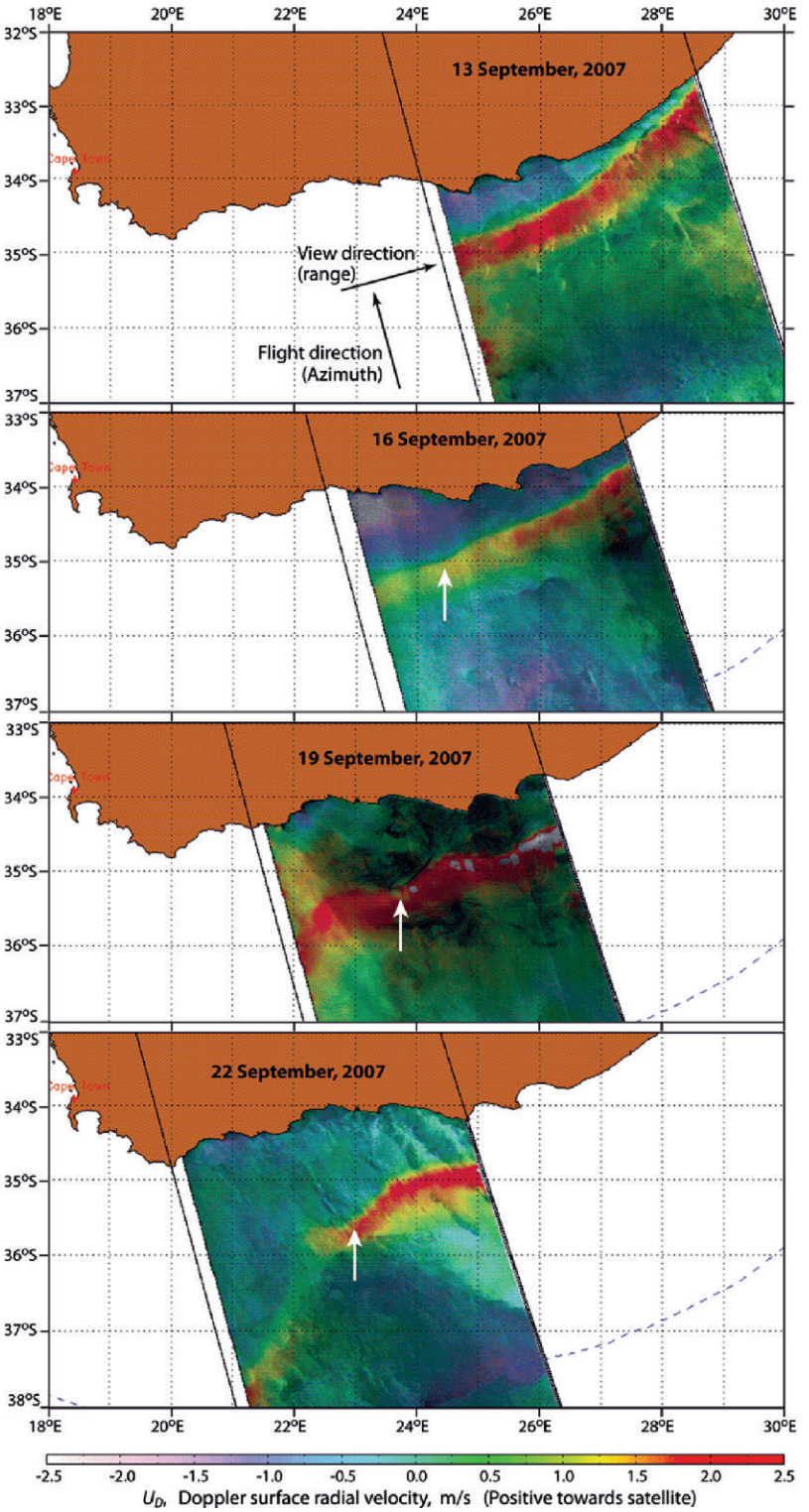


Figure 4.15. Normalized radar cross-section σ_0 (gray shades) and Doppler velocity U_D (colors), analyzed from wide-swath images obtained by Envisat ASAR on four successive overpasses 3 days apart between September 13 and 22, 2007. The arrows denote the movement and growth of lateral perturbations of the front. This is an experimental data product produced for ESA by Fabrice Collard of BOOST Technologies (composed by the author using individual images available at the ESA website).

present the most direct and unambiguous way to detect fronts in support of oceanographic research. Identification of fronts can be performed automatically using edge detection techniques (Canny, 1986). The use of gradient-related operators designed to enhance brightness gradients on thermal images, including the filters discussed in Section 4.2.1, has formed the basis of automated objective methods to identify fronts using image analysis software (Simpson, 1990). In most cases this presents the oceanographer with a mask of pixels or a set of lines on a map, showing where gradients above a given threshold have been detected.

These methods still have the drawback that they may falsely identify image artifacts like undetected cloud as fronts, or fail to record fronts when they are actually present. Thus some operator supervision may be needed for critical operational applications (e.g., knowledge of a front location may be essential for assessing the impact of SST on human survival in a search-and-rescue situation). For research applications based on information about front location and occurrence, their scientific credibility depends on knowing the validated error bounds for the data that are output from automated front detection systems.

With a view to improving the accuracy of front detection, an alternative approach was developed (Cayula and Cornillon, 1992), based on the step-like character of ocean fronts which are sharp boundaries between regions in which the SST is fairly uniform. Instead of searching for strong temperature gradients it analyzes the histogram of SST values (after cloud detection) within windows of 32×32 pixels, searching for evidence of two distinct populations that differ from each other by 0.375°C or more. If so the window is assumed to contain a front, the location of which is mapped to those "edge pixels" at the transition temperature. This is repeated for overlapping windows over the whole image to produce lots of small line segments designated as frontal pixels. An attempt is then made to join the line segments to each other, most importantly requiring the fronts to follow isotherms. Any remaining segments less than 10 pixels long are deleted.

An improvement called the multi-sensor histogram edge detection (MSHED) method (Cayula and Cornillon, 1995) applies single-image analysis to other images of the same region acquired within 2.5 days before and after. This reinforces single-image data and improves coverage when there is patchy cloud. The front segments from each image are then combined, this time using gradients rather than following isotherms to match individual fronts from one image to another, since absolute differences may exist between the different AVHRR overpasses used for this work. Persistent fronts are retained. Then consolidated frontal edges are thinned before being fed back into a final analysis of the single image when once more they must conform to isothermal behavior. Figure 4.16 illustrates some of these stages, being (a) edges (shown in white) detected by single-image analysis overlaid on the original image, (b) multi-image ensemble of edge segments, (c) retaining only the persistent fronts from (b), and (d) the final result after edge thinning and reapplying to isotherm contours on the original image. Ullman and Cornillon (2000) validated MSHED over the Gulf Stream, using *in situ* observations from ship transects operating continuously over 2 years. They confirmed that MSHED produced only 14% false fronts compared with 29% for the gradient edge detector. The number of fronts

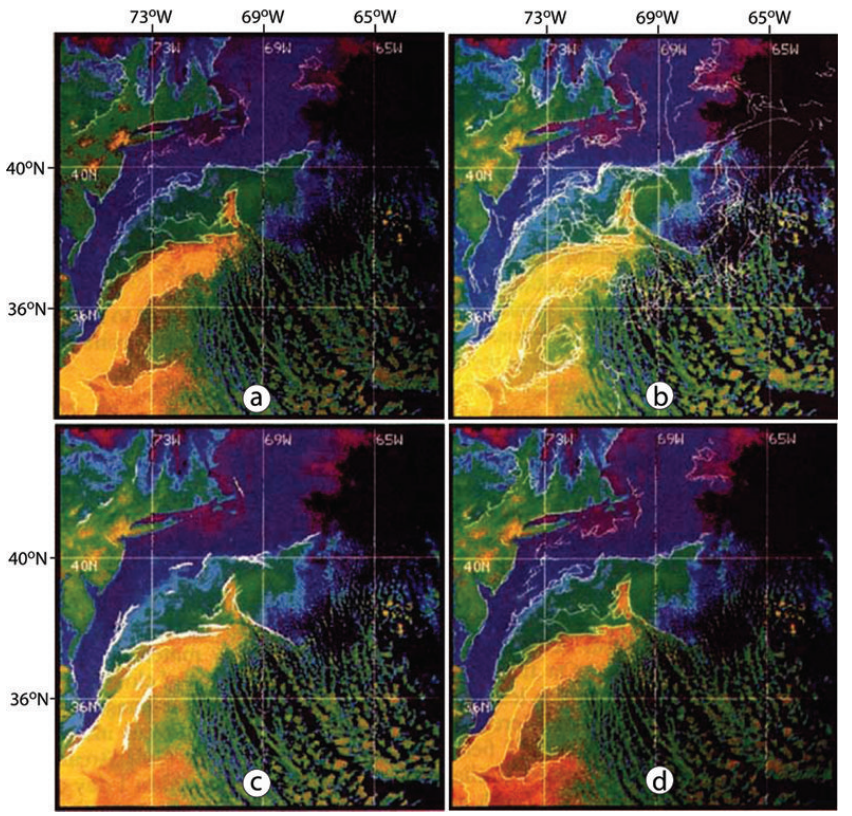


Figure 4.16. Detection of fronts on AVHRR SST image of 18:41 LT on March 26, 1985 (the core image), by the multi-image algorithm. (a) Fronts (shown as white lines) on the core image detected by the single-image edge detection algorithm. (b) Fronts found in all images within 60 h of the core image. (c) Result of editing (b) to retain only persistent fronts. (d) The end product after thinning the fronts in (c) and eliminating any not matching the isotherms of the core image. Note that there are several front segments remaining in (d) which were missed in (a) (adapted from Cayula and Cornillon, 1995).

missed was 5% for both MSHED and the gradient method, as long as only fronts longer than 10 km were considered. However, MSHED tended to perform worse than the gradient method for identifying front segments shorter than 10 km. Overall it was estimated that when MSHED is used for producing climatologies of fronts, errors in frontal probabilities are less than 15%.

The MSHED approach has been proved effective and underpins the global surveys of ocean fronts described in Section 4.4. It would be satisfactory for studies of the distribution of fairly short-lived fronts in a field of mesoscale turbulence, such as the South Atlantic image in Figure 4.4. However, when a permanent front is the focus of a study and satellite data are to be used to observe how it moves and changes strength over time, simple line maps of frontal position tell only part of

the story, and certainly do not contain all the frontal information that is potentially available from thermal image data.

4.3.2 Automatic parameterization of frontal structure

In response to this need, an image data analysis technique has been developed (Shaw and Vennell, 2000) which is able to follow a major front along its length, recording not only the position of its steepest gradient but also its width, the temperature difference across it, and the mid-point temperature between the two extremes. The advantages of such an algorithm, in comparison with the simpler line detection methods, are that it measures the characteristics and position of the front as parameters, and also records how those characteristics change along the front. Each time the algorithm is applied to a different image of the same front, these parameters are re-evaluated to build up a time series which describes more completely the evolving history of the front. This allows these frontal parameters to be analyzed in relation to large-scale forcing factors that may constrain or change the front, and to quantify the effect which changes in the front have on other aspects of local oceanography, such as primary production.

The Shaw and Vennell front-following algorithm was designed to operate on SST datasets derived from AVHRR but in principle the algorithm can be adapted for other sources of SST, preferably provided at the native resolution of the sensor in the original scan line/along-track co-ordinates, and for which each pixel is already geolocated. The procedure operates within a small extraction window defined in (x', y') co-ordinates as shown in Figure 4.17. This is based on a grid of square pixels of linear dimension equal to the mean of the lengths of 1 arcmin of latitude and of longitude, evaluated at the center of the extraction window. The frontal line is defined as the line of inflection where the maximum cross-front gradient of temperature occurs. The extraction window is centered on the estimated position of a point on the frontal line, and is oriented with its y' axis along the estimated line of the front. The extent of the frontal window should be adapted to the scale and character of the particular front being studied. For the case of the Southland Front off New

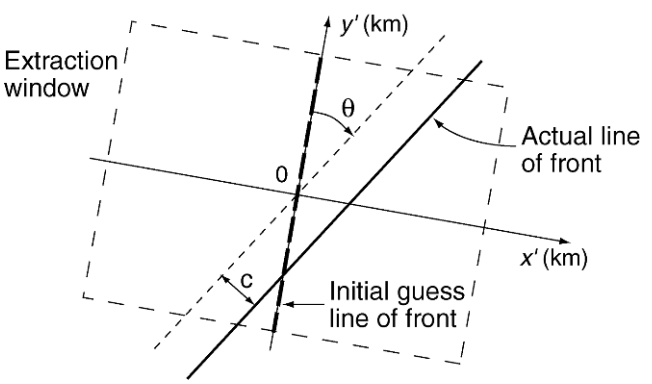


Figure 4.17. Definition sketch of the extraction window used for the front-following algorithm (after Shaw and Vennell, 2000).

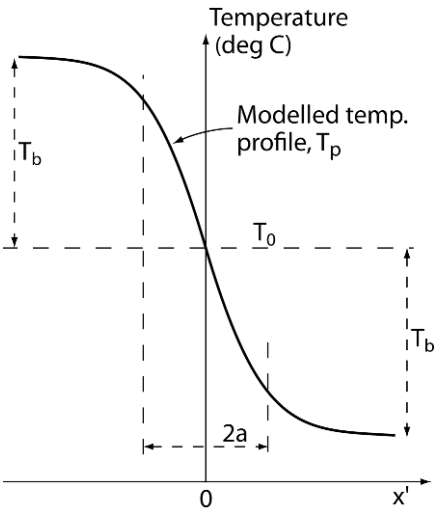


Figure 4.18. Sketch of the hyperbolic functional form used to represent the temperature profile across an isolated front (after Shaw and Vennell, 2000).

Zealand a size of 30 km along x' and 20 km along y' was used (Shaw and Vennell, 2000). At that latitude (about 45°S), grid spacing was 1.6 km so the extraction window contained 19×13 cells. The temperature field is resampled from the satellite grid into the extraction grid, using nearest neighbor substitution.

The shape of the temperature profile, $T_p(x')$ across the front is assumed to be a hyperbolic tangent, which would have the form:

$$T_p = T_0 - T_b \tanh(x'/a) \tag{4.1}$$

as in Figure 4.18, as long as the co-ordinate x' is truly at right angles to the front and that the origin of x' is precisely at the front. T_0 is the temperature at the front line itself, and the parameter T_b is half the temperature difference between the plateau regions assumed to characterize the temperature field on either side of the front. Note that the positive x' direction is from the warm to the cold side of the front. Because of the symmetrical form assumed, T_0 is the same as the mean of temperatures on the warm and cold sides of the front. The parameter a is a lengthscale that characterizes the width of the front. $2a$ is taken to represent the width of the front, and (from Equation 4.1) it is the distance across the steepest part of the front over which the temperature changes by 76.19% of the full temperature step of $2T_b$.

In practice we must expect the line of the actual front to be rotated through angle θ (measured clockwise from the $+y'$ axis) from the estimated orientation, and also that the line of the actual front is shifted a distance c from the estimated origin (measured perpendicular to the front and positive when shifted towards the $+x'$ direction) as shown in Figure 4.17. Then the two-dimensional modeled representation of temperature $T_m(x', y')$ within the extraction window becomes:

$$T_m = T_0 - T_b \tanh\left(\frac{x' \cos\theta - y' \sin\theta - c}{a}\right). \tag{4.2}$$

The algorithm attempts to fit this model to the actual temperature field in the resampled window, using T_0, T_b, a, c, and θ as adjustable frontal parameters. Using a numerical routine which iteratively adjusts the parameters until the square of differences between model and actual temperatures in the extraction window are minimized, the values of frontal parameters are found. Tuning the algorithm so that this process converges efficiently and accurately requires care and attention to the general characteristics of the temperature field in which it is applied (Shaw and

Vennell, 2000). For example, the initial values of parameters in the iteration must be chosen sensibly, limits must be set on the permitted range of each parameter, rules must be set to determine when acceptable convergence of the solution has been achieved, and a maximum permissible number of iterations must be set, beyond which the procedure aborts and returns a null result. A balance must be struck in order to parameterize the front as precisely as possible while still converging on a result as long as there is a well-defined front. This is where knowledge of the region is needed. For example, if the front is stepped rather than smooth, the algorithm must be trained to model the full width of the whole front rather than attempt to follow a single step. Because the solution looks for a broad plateau temperature on either side the algorithm would not converge on a single step. The size of the extraction window is important in this respect. If it is set too narrow in the x' direction then it will not find the plateaux. However, if it is set too wide, detection of the main front may be hindered by the presence near the edges of the window of other temperature features unconnected with the front.

Having parameterized the front at the given starting position, the process then proceeds to follow the front. Using the measured parameters, the position and orientation of the front is predicted at a new location shifted by a given distance in the y' direction. A new extraction window is created for the new location from the original SST image data, and the parameterization algorithm is applied again. The parameters from the previous location (other than c and θ) are used as starting values for iteration at the new location. Shaw and Vennell used a shift of 2 km, and found that the large degree of overlap between the two windows ensured that convergence was almost always reached in a small number of iterations (typically less than 10). If the iteration fails to converge, the window is shifted again and a new attempt is made. The front-following procedure terminates when the parameterization fails for too many successive extraction windows.

4.4 CLIMATOLOGY OF THE MAJOR OCEAN FRONTS

If evidence were needed that the fronts of major currents were one of the very first ocean phenomena whose scientific study benefited from satellite data, it can be found in the first survey of ocean thermal fronts worldwide (Legeckis, 1978). This was based on sea surface temperature observations using VHRR sensors (forerunner of the AVHRR) on the satellites NOAA 3, 4, and 5. With a temperature sensitivity of 0.5 K, limited by electronic noise, and using a single thermal IR waveband of 10.5 μm to 12.5 μm, so that no serious atmospheric correction could be applied, data collected between 1973 and 1977 provided a rich source of new scientific knowledge about the instantaneous spatial structures of fronts as well as their temporal variability. Data are also shown from the early geostationary spin scan radiometers. Thirty years after its publication, this paper is still worth reading although remember that the convention then was to present infrared images with cold as light and warm as dark. The references in that early paper also provide insights to other pioneering studies as space technologists began to contribute to ocean science.

Of course, even then, oceanographers did not have be told by space scientists where to find the major ocean fronts! What was new was to discover the variability and find the evidence that frontal instability led to the creation of large rings and eddies. Today, as SST sensors have improved in both sensitivity and absolute accuracy, a substantial body of knowledge has been accumulated about the probability of where fronts occur and about the character of their variability. This is what we may describe as the climatology of ocean fronts. To be able to achieve statistical reliability based on thousands of samples required the development of the automated front detection methods mentioned in Section 4.3. There is only room in this section to show a few examples and point towards a range of papers produced during the last 10 years.

Many of these have used the multisensor histogram edge detection method (Cayula and Cornillon, 1995). For example, Belkin and Cornillon (2003) use 12 years of AVHRR Pathfinder SST data to present a survey of the thermal fronts of Pacific coastal and marginal seas. Figure 4.19a, b show, respectively, the occurrence of fronts identified in each pixel across the whole Pacific Ocean. This dataset has constant area square pixels with dimension 9.28 km and is generated twice per day. The pixel detection algorithm has been applied to each separate (cloud-flagged) SST field. To evaluate the frequency of occurrence the number of times a front is detected in a particular pixel is counted for all the fields occurring in the specified months over the whole 12 years. That is then divided by the total number of cloud-clear occurrences of the same pixel, counted over the same months. The result is given as a percentage. There is no easy way of knowing whether cloudy pixels introduce a bias into the occurrence.

Figure 4.19 contains a valuable amount of information that should be of wide oceanographic interest, being relevant to all oceanographic phenomena, physical, chemical, and biological, that are associated with fronts. It shows where the fronts are to be found. They can be approximately classified in four groups.

1. Open-ocean fronts lying approximately east–west, and the product of diverse causes: the Antarctic Circumpolar Current (ACC) and the Kuroshio extension are associated with large-scale ocean circulation; the fronts north and south of the Equator are either side of the equatorial upwelling associated with the unique dynamics occuring where the Coriolis parameter is close to zero; the subtropical frontal zone is rather more diffuse.
2. Western boundary fronts along the edge of major currents that supply return flows of basin-wide, wind-driven gyre circulations: the Kuroshio and Oyashio Currents in the North Pacific, the East Australian Current in the South Pacific.
3. Eastern boundary fronts, mainly deriving from wind-induced coastal upwelling off the U.S. west coast, off Central American gulfs (Tehuantepec, Papagayo, and Panama), and off Peru and Chile.
4. Marginal sea fronts, found in several semi-enclosed seas east of Asia. From south to north these are the South China, East China, Yellow, Bohai, Okhotsk, and Bering Seas and they all show high frequency of frontal occurrence during part of the year, although several are seasonal.

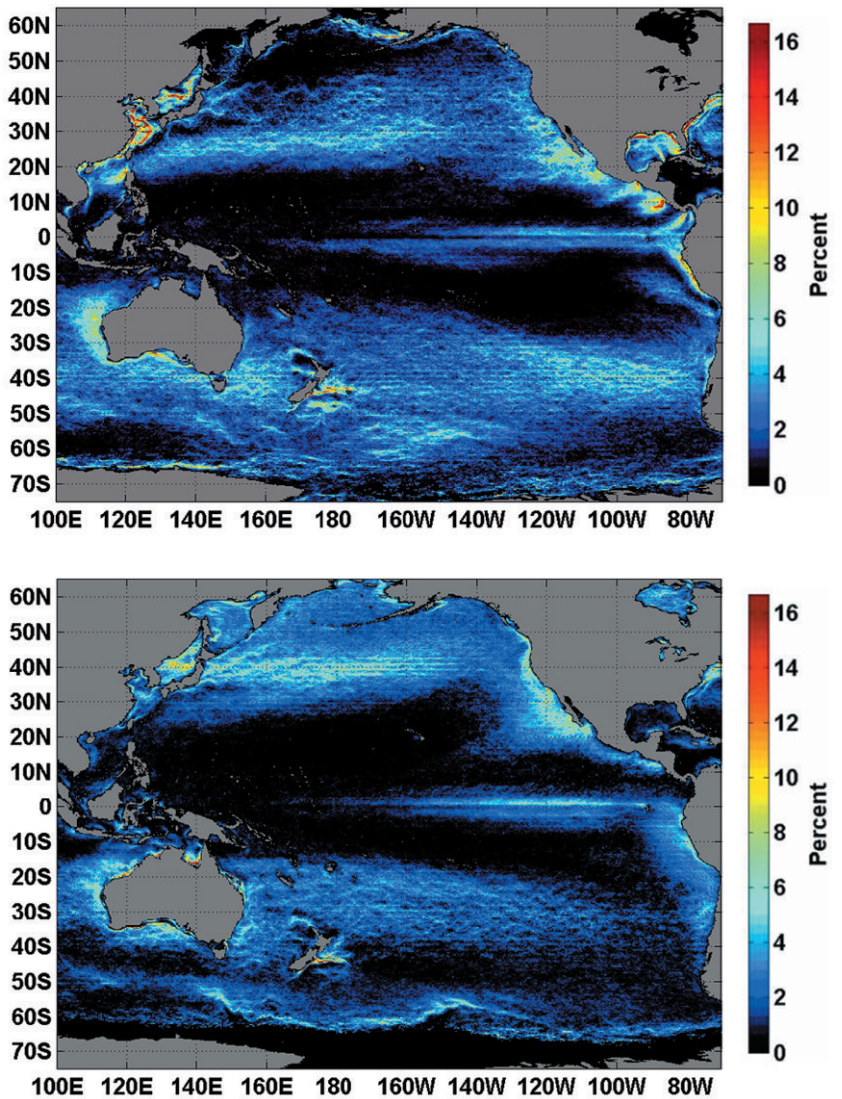


Figure 4.19. Long-term, seasonal-averaged frequency of thermal fronts occurring within each 9.28 km resolution pixel of the Pathfinder SST dataset across the Pacific Ocean, evaluated between 1985 and 1996. (a) Boreal winter (January, February, March). (b) Boreal summer (July, August, September) (from Belkin and Cornillon, 2003).

Figure 4.19 not only shows the location of fronts but also whether they are diffuse over a wide area with a lower frequency, or tightly constrained to a narrow region where their probability is higher. The different distributions between the seasons also contains valuable oceanographic information. Seasonal variability could be more finally resolved by months. Front occurrences could also

be mapped to show interannual variability, or related to major climate indices such as the ENSO index. Now that these front occurrence maps have been produced from satellite data twice per day for 12 years, they represent a useful resource for a variety of applications.

Whereas Figure 4.19 paints the broad ocean-wide picture, the same data have been analyzed on a more localized basis, distinguishing between many distinct local fronts such as those shown in Figure 4.20 for the East China Sea (Hickox *et al.*, 2000). The variety of frontal types found in different marginal seas, such as shelf break fronts, tidal mixing fronts, salinity-driven fronts associated with major river plumes, upwelling fronts associated with local wind patterns, and marginal ice zone fronts in polar seas, are revealed by the satellite data to have different seasonal

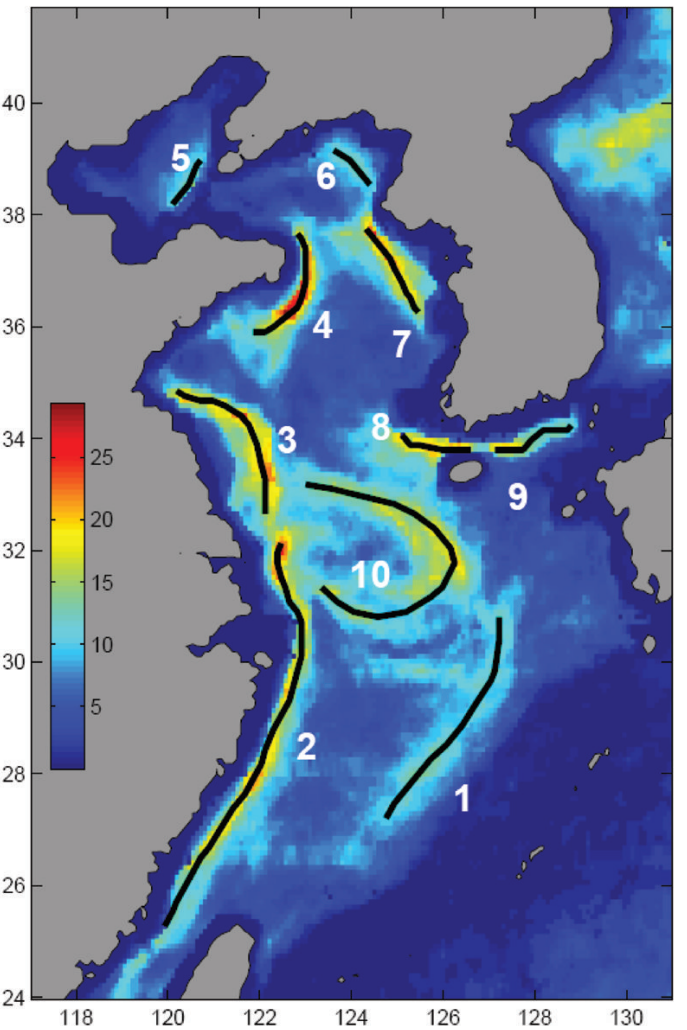


Figure 4.20. Long-term, annual composite frontal probability map in the East China Sea for the 1985–1996 period derived from the AVHRR Pathfinder SST dataset. For each pixel, the percentage of total time that the given pixel contained a front is represented by color (scale represents percentage). Solid black lines mark maximum probability ridges that correspond to most probable locations of fronts. Numbers are: 1, Kuroshio Front; 2, Zhejiang–Fujian Front; 3, Jiangsu Front; 4, Shandong Peninsula Front; 5, Bohai Sea Front; 6, Seohan Bay Front; 7, Kyunggi Bay Front; 8, Western Chejudo Front; 9, Eastern Chejudo Front; 10, Yangtze Bank Ring Front (from Hickox, *et al.*, 2000).

climatologies. Figure 4.21 demonstrates this for the East China Sea by seasonally separating the frontal probabilities shown as annually integrated in Figure 4.20. Similar detailed studies have been made of the Bering Sea (Belkin and Cornillon, 2005), the Sea of Okhotsk (Belkin and Cornillon, 2004), the fronts associated with upwelling off Oregon and the Californian Current system in the northeast Pacific (Castelao *et al*., 2006 who interestingly use SST from the geostationary GOES satellites), and seas off the eastern coast of the U.S.A. inshore of the Gulf Stream (Ullman and Cornillon, 1999, 2001).

Another interesting example of frontal occurrence analysis is that by Saraceno *et al*. (2004) based on 9 years of Pathfinder SST 5-day composite data (1987–1995) over the southwest Atlantic Ocean including the complex frontal structures surrounding the Brazil–Malvinas confluence zone. This uses a gradient method for edge detection, which appears to be effective in this region where cloud contamination of 5-day data is lower than most other parts of the world where ocean fronts are encountered.

4.5 MESOSCALE FRONTAL VARIABILITY

The climatology of fronts explored in the previous section already provides some information about temporal variability in the occurrence of fronts and is capable of further exploitation by oceanographers studying other ocean phenomena that are enhanced or hindered by the presence of fronts. Nonetheless a statistical survey of occurrence does not by itself improve understanding of the processes which cause fronts to vary. For example, if a front ceases to be observed in a particular location we do not know whether it has moved, weakened. or disappeared altogether. Moreover when a front is noted as being present for an extended period of time its strength may change significantly but that is not recorded in frequency maps. We therefore look now at examples of what remote sensing can tell us about the modes and mechanisms of frontal variability.

4.5.1 The Gulf Stream

Quite a lot of work has been done by using the time series of frontal edge maps, such as those used as input to occurrence statistics. For example, using 2-day composite maps of the northeast edge of the Gulf Stream compiled laboriously from AVHRR data between April 1982 and December 1989, the dynamical characteristics of this important current were systematically explored for the first time. Lateral, wavelike perturbations of the front were identified and their evolution followed, leading to a definitive confirmation that at least 40% of all northward meanders having a lengthscale (half the distance between meander crests) longer than 100 km would detach to form warm rings between 75°W and 60°W (Cornillon *et al*., 1994).

It was also observed that the spatially averaged path of the Gulf Stream shifts laterally over an annual cycle (farthest north in November and south in April) in response to shifts in the wind field, while time variability of the amplitude of spatial perturbations is less regular, but has a spectral peak of 9 months and clearly evident

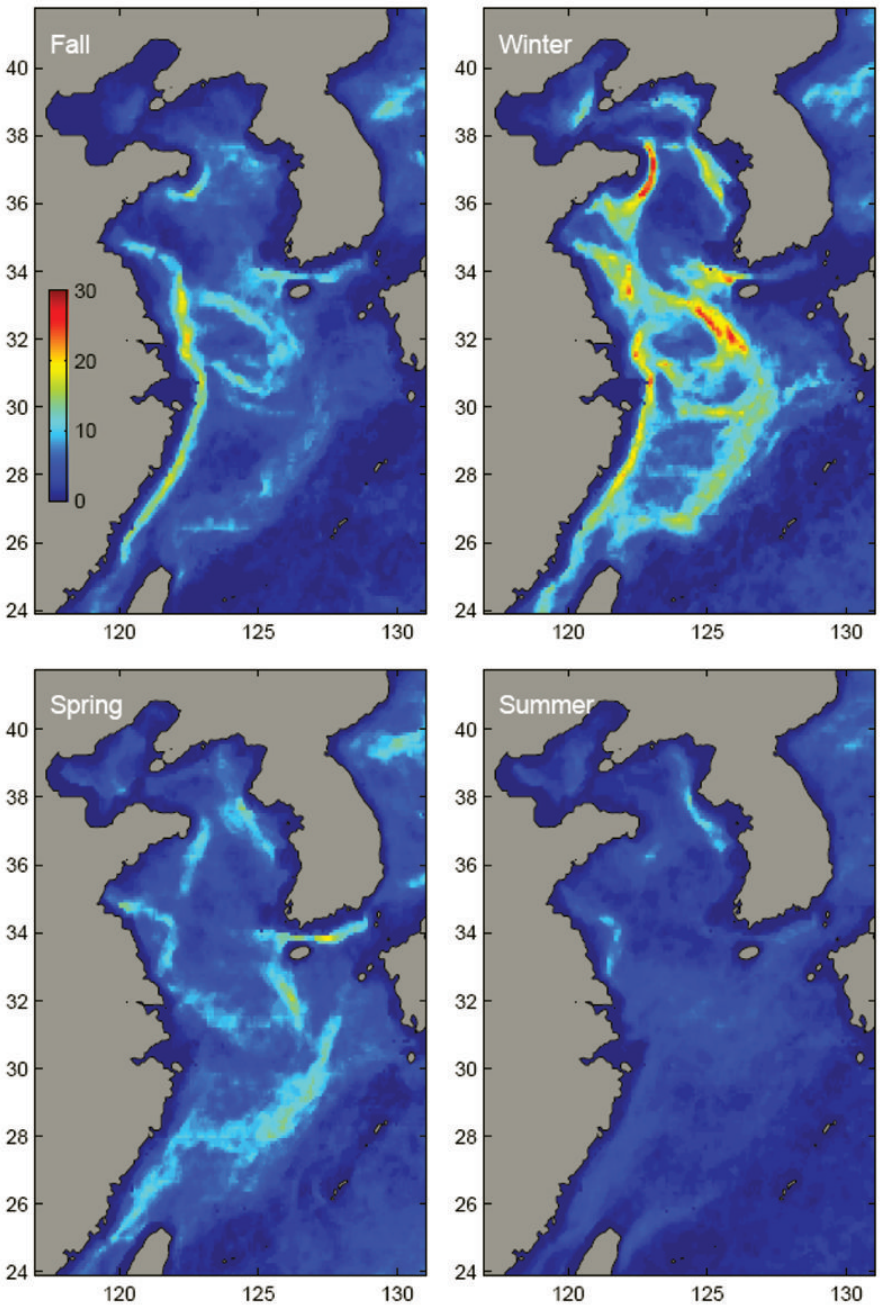


Figure 4.21. Probability maps for the East China Sea showing seasonal breakdown of data presented in Figure 4.20. Each seasonal map accumulates data from 3 months as follows: fall is October to December; Winter is January to March; Spring is April to June; Summer is July to September.

interannual variations (Lee and Cornillon, 1995). This major study of frontal edge evolution was concluded with a spectral analysis in wavenumber–frequency space, and EOF (empirical orthogonal function) analysis in the time and frequency domain, in order to distinguish propagating perturbations from standing waves, and to estimate the speed, growth, and relative energy of different modes. All of these quantitative dynamical measures retrieved from satellite data were related to contemporary *in situ* observations and compared with theoretical underpinning from geophysical fluid dynamics (Lee and Cornillon, 1996a, b).

4.5.2 The Southland Front

Returning now to explicit measurements retrieved from the front-following algorithm described in Section 4.3.2 (Shaw and Vennell, 2000), it is interesting to discover what oceanographic benefits accrued from being able to parameterize characteristics of the Southland Front off the southeast of New Zealand (Shaw and Vennell, 2001). The Southland Front is part of the subtropical front, a global front that extends around the Southern Ocean and is the boundary where Subantarctic Surface Water converges with the Subtropical Surface Water mass. More than 3 years of cloud-clear views of the front were obtained from 277 SST images and the algorithm was applied to every one. The outcome was a time series of frontal parameters which allowed both seasonal and annual variability to be analyzed with statistical rigor.

The mean position of the front was found to be very stable and closely linked to the 500 m isobath, implying that geostrophic topographic steering controls the position of the frontal current. However, the front-following algorithm was able to detect plumes bursting from the front from time to time. It also showed that the steepness and width of the front varies seasonally, being strongest and narrowest in winter. Variability of the frontal position about the mean was also shown to vary interannually. Geographic variation of the steepness of the front could also be explored along its length, relating it to spatial trends in water masses from west to east and showing some relationship to the position of rivers discharging onto the adjacent continental shelf. Over the 3-year period of the study there was a suggestion that the gradient of SST across the front decreased as the Southern Oscillation Index decreased, but a longer timespan of data is needed to confirm this.

4.5.3 Antarctic Circumpolar Fronts

There have been a number of publications over the last 30 years reporting observations of the frontal edges of the Antarctic Circumpolar Current, using mainly thermal infrared sensors (Legeckis, 1977; Moore *et al.*, 1997) or altimetry (Chelton *et al.*, 1990; Gille, 1994). To some extent these found it difficult to make sense of variability in the position of fronts, leading to uncertainty about the number of fronts that are contained in the full current system. They also struggled to find zonal coherence in variability along the flow. In Section 4.2.2 it was noted that techniques have recently been developed for using altimetry to observe the detailed

structures of large fronts that break into several subfronts and where the flow splits into several filaments, co-located with subfronts (as shown schematically in Figure 4.7). Here we discover how this approach has elucidated understanding different strands of the Antarctic Circumpolar Current (ACC), focusing on the sector south of Australasia between 100°E and 180°E but applicable to the complete circumnavigation.

Until recently the growing volume of hydrographic evidence from the Southern Ocean had led to the general picture of a frontal zone spanning 10° to 20° of latitude with three main cores of the ACC located at the three main stages of the front (Orsi *et al.*, 1995). Supposed major frontal steps were, from north to south, the Sub-Antarctic Front (SAF), the Polar Front (PF), and the southern ACC front (sACCf), with the southern boundary—ACC (Bdy)—designated as another feature. Distinctive water masses separated by these fronts could be traced continuously around the ACC. However, as pointed out by Hughes and Ash (2001), the simplicity of this stable continuous structure was not supported by maps of temperature gradients acquired from the ATSR (as shown in Figure 4.22), or by maps of gradients of SSH from multisensor altimetry. Fine-resolution models developed in the 1990s also predicted narrow filaments of flow in short, apparently disconnected segments. Each of these sources implied a more complex structure of at least six to eight or more filaments of narrow jets, seemingly noncontinuous, splitting, and combining apparently randomly in space and time.

It is expected from geophysical fluid dynamics theory that wide zonal streams will naturally organize themselves to concentrate the flow into narrow jets constrained by the β-effect (latitudinal gradients of the Coriolis parameter, f). However, it seems that in the ACC these are more finely split up and narrower than the three main rivers of current originally believed to exist. Furthermore a 3-year time–latitude plot at 130°E of the sea surface height gradient derived from

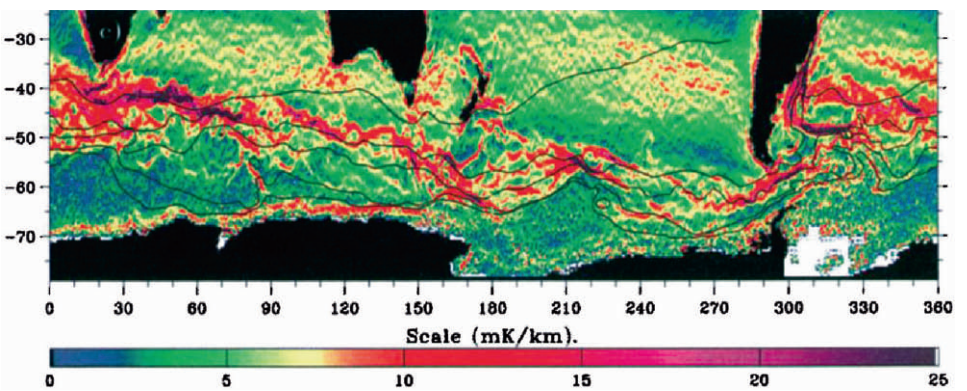


Figure 4.22. Mean gradient of observed sea surface temperature from ATSR. The frontal positions from Orsi *et al.* (1995) are superimposed. They are, from north to south, the Subtropical Front, Sub-Antarctic Front (SAF), Polar Front (PF), south ACC front (sACCf), and southern boundary of the ACC (from Hughes and Ash, 2001).

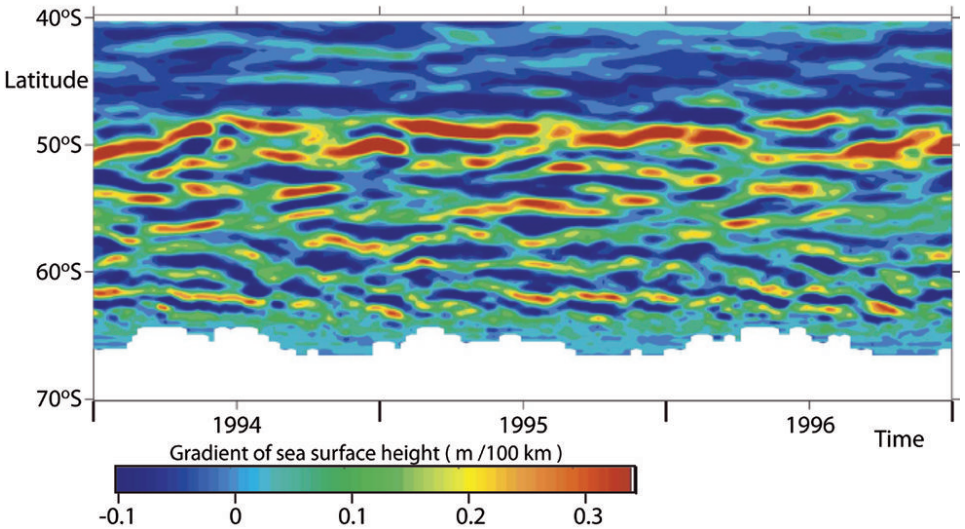


Figure 4.23. Time–latitude plot of meridional gradient of sea surface height (in meters per hundred kilometers) between 1994 and 1997, evaluated at 130°E (from Sokolov and Rintoul, 2007a).

multisensor altimetry (shown in Figure 4.23) confirms that the flow structure is also highly variable in time. This apparently ephemeral nature of filaments raised questions of inconsistency with the strong stabilizing constraints of flow on a β-plane (a theoretical model framework in which f varies linearly with latitude), as well as with historical hydrographic evidence.

This was one of the issues addressed by Sokolov and Rintoul (2007a). They carefully constructed a 12-year sequence of weekly maps of absolute dynamic topography (ADT) by adding the SSHA to mean surface dynamic height relative to 2,500 dbar, derived from climatology. Their next step was to test the hypothesis that current jets (defined by maxima in SSH gradients and corresponding to frontal edges) were associated with particular values of absolute dynamic height. This proved to be the case (as illustrated in Figure 4.24). It was revealed that wherever and whenever strong frontal gradients emerged within turbulent fluctuations they did so along one of eight distinct contours in the ADT, which they named the north, middle, and south streams of the SAF and of the PF, and north and south streams of sACCf. Although each front is not continuous in an instantaneous snapshot of the gradient of SST or SSH, the segments of a particular front that do appear at any one time can be linked to each other along the specific ADT contour for that front, across the whole 80° longitude sector being studied. Thus the specific ADT contour acts as a guideline for where to find the front.

Since the contours of ADT approximate to streamlines of the flow in an equivalent barotropic flow like the ACC (Killworth and Hughes, 2002), this begins to make sense of the apparently random patterns of frontal edges or filaments

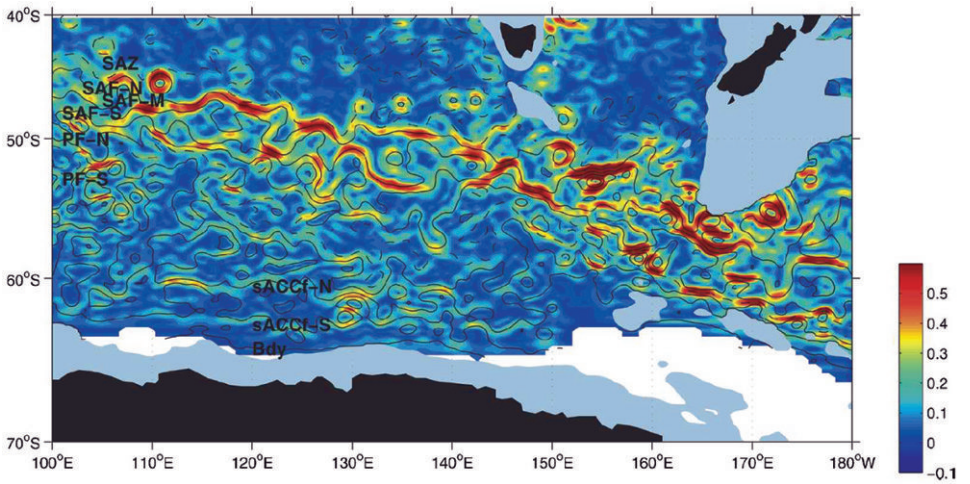


Figure 4.24. A typical SSH gradient field south of Australia and New Zealand (July 3, 2002) overlaid with the mean best-fit SSH contours optimized for the whole period of observations. The contours corresponding to the SAZ feature (the northern limit of the AAC) and the middle branch of the PF are dashed, for clarity. Depths shallower than 2,500 m are shaded (from Sokolov and Rintoul, 2007a).

presented by Figures 4.22 and 4.23. Being able to map the ADT and its turbulent fluctuations is the key. As these streamlines meander north or south in the constrained, but turbulent flow associated with a jet on a β-plane, the template for where to find the fronts also moves, and the strength of each front changes with how close its particular streamline is to its neighbors. That is why isolated snapshots of spatial distribution or continuous monitoring at a single longitude (Figure 4.23), give the appearance of random fluctuations of disconnected strands of front.

Having identified each of the eight fronts by its ADT contour, Sokolov and Rintoul (2007a) could map its time-mean location and the probability of finding it on either side of that position. The spread of probabilities for all eight fronts overlap each other, but of course at any instant they are all kept separate from each other, and ordered in the correct north–south sequence, by being tied to their own ADT contour. Instead of basing our understanding of the ACC structure on vainly attempting to pin down the precise geographical location of each strand of the front, we are given the more helpful picture of each front being carried along with the flow as revealed by the evolving ADT field. Moreover, since unique hydrographic properties are attached to each particular streamline, equivalent to a given contour of the ADT, then in fact each front continues to separate specific water masses.

This largely resolves the discrepancy between the satellite/numerical model view of the ACC and the *in situ* hydrographic view. When both can be interpreted in the context of changing streamline patterns they should tell the same story. This ongoing research offers an excellent example of how remote sensing is starting to give dynamical oceanographers appropriate tools for measuring parameters of the

ocean flow field, in much the same way as they might use flow visualization techniques in laboratory-scale experiments. Improvements to the geoid expected from the GOCE mission should improve the absolute accuracy with which the ADT can be defined. When that happens we can expect this type of research to develop in exciting directions, bringing together satellite oceanographers, hydrographers, numerical modelers, and theoretical fluid dynamicists to provide a more complete view of the ACC, which is such an important phenomenon of the global ocean.

4.6 BIOLOGICAL PRODUCTION ASSOCIATED WITH OCEAN FRONTS

To conclude the chapter we consider how satellite observations enable the primary production associated with ocean fronts to be better understood in the context of physical parameters and processes. Just two examples are shown, linked to frontal zones encountered earlier in the chapter. However, further examples are to be found in several other chapters where the use of ocean color alongside other satellite data provides a multidisciplinary perspective on mesoscale phenomena.

4.6.1 Antarctic Circumpolar Current

The first example comes from a development of work, described in Section 4.5.3, to understand the frontal structure in the ACC. In a companion paper, Sokolov and Rintoul (2007b) describe how they made use of insights gained from their dynamic analysis (Sokolov and Rintoul, 2007a), in order to interpret the distribution of primary production around the whole 360° of the ACC region, as observed previously from satellite ocean color observations of chlorophyll (e.g., Moore and Abbott, 2002) and *in situ* measurements. Figure 4.25 shows a map of mean, summer, chlorophyll concentration for the 100°E to 180°E sector, derived from the level 3, 8-day, 9 km dataset from 5 years of SeaWiFS overpasses (1987–2002). Crucially the mean positions of the eight distinct fronts and southern boundary identified in the earlier work have been overlaid. These can be used to chart the flow of different bands of surface water from west to east and facilitate interpretation of chlorophyll maps within those distinct bands.

Ignoring regions of high concentration which catch the reader's eye off Tasmania, New Zealand, and the Antarctic continent, it can be seen that the chlorophyll concentration within the ACC region is low but not uniform. Water within the SAF region (between the north and south streams of the SAF) is generally lowest in chlorophyll, which increases within the PF region and is greatest within the sACCf, but there is also significant variability along these frontal regions which can be related very clearly to the underlying topography. Chlorophyll is least concentrated where frontal currents traverse the deepest waters and *vice versa*. This is shown even more clearly when chlorophyll is mapped along streamlines taking into account time variation of frontal positions and matching each instantaneous map of ADT to the contemporaneous chlorophyll map. Note that because Figure 4.25 has been integrated on a fixed, geographical grid over 5 years of summer months

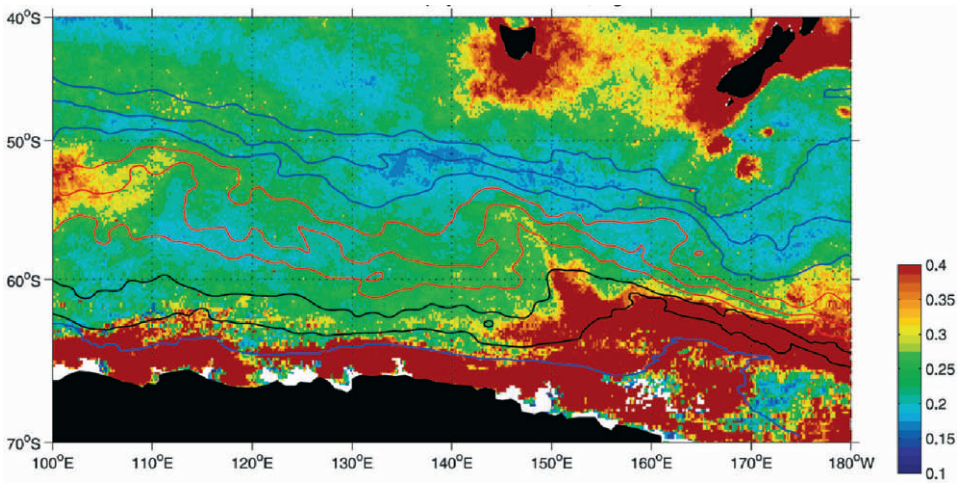


Figure 4.25. Map of mean summer chlorophyll concentrations south of Australia and New Zealand (color scale is in milligrams per cubic meter). Colored lines represent mean summer positions of ACC fronts as determined by Sokolov and Rintoul (2007a). From north to south they are: SAF-N, SAF-M, SAF-S (blue lines); PF-N, PF-M, PF-S (red); sACCf-N, sACCf-S (black); and the southern boundary (blue) (from Sokolov and Rintoul, 2007b).

it must have smoothed out some cross-frontal differences that exist in instantaneous maps. In practice the availability of regularly updated ADT maps provides a more Lagrangian approach to interpretation, allowing the longitudinal and seasonal distribution of chlorophyll to be analyzed separately within each band of the ACC flow.

This led Sokolov and Rintoul to confirm that the distribution of chlorophyll in the Southern Ocean is concentrated in a number of persistent blooms, occurring downstream of islands and bathymetric features. Their location is consistent with a model of nutrients being upwelled into a particular strand of the ACC when it flows over shallow topography and then being advected eastwards to support production in that strand of the current until the nutrients are depleted. Little evidence is found for fronts, by themselves, being associated with enhanced production. They conclude that the upwelling needed to support production in the ACC is driven by bottom pressure torques on the large-scale flow over topography, rather than by small-scale turbulent instabilities in jets.

4.6.2 Fronts in the southwest Atlantic

The Brazil–Malvinas confluence zone is a very complex region, not only dynamically but also in terms of its capacity for primary production in relation to nutrient supply, temperature, and optical properties of water. In order to characterize the biological behavior of a particular location in the ocean, it is customary to classify it as a particular "biogeographical province" (Longhurst, 1998), thus providing some insight into how different species and organisms may perform, based on experience

from other regions having the same or similar classification. Assigning the correct classification is difficult in this complex region where several fronts lie close together and shift seasonally.

In a detailed study of the region, Saraceno *et al.* (2005) used SST, SST gradients, and chlorophyll concentration to identify frontal zones and to characterize the seasonal variability of surface waters. Their primary data sources were level 3 satellite data products from a 6-year period (1988–2003). For the chlorophyll they used the 8-day SeaWiFS dataset, while for SST they used AVHRR, both locally received high-resolution data and coarser global Pathfinder data. In order to cope with cloud cover, worse in some parts of the region than others, they also made use of global, optimally interpolated analysis products and, from 2002, AMSR-E data that can see through clouds although at poorer resolution. Interestingly this is one of the first recorded uses of AMSR-E data in a scientific study. The concept of combining data at high resolution from several sources for this type of application is discussed further in Chapter 14.

With the help of these data, they were able to find much of the information needed to classify the region using histogram analysis based on spatially and temporally detailed measurements. This allowed a more objective approach instead of depending subjectively on more general knowledge that went into the previous assignment of biogeographical provinces in the area. Figure 4.26 illustrates the results, although a single figure cannot do justice to the detailed analysis contained

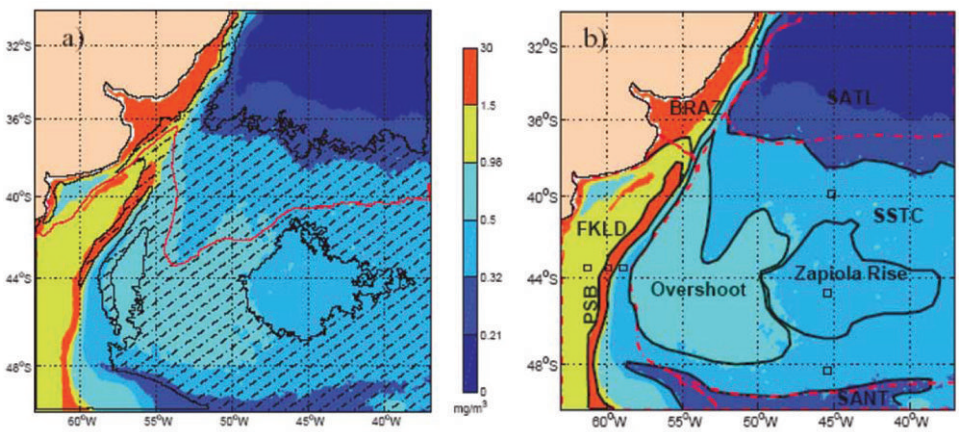


Figure 4.26. (a) Six ranges of chlorophyll-*a* magnitudes, based on respective histograms in regions, represented by background colors. The hatched region indicates areas with SST gradients higher than 0.08°C/km. The solid red line is a temperature threshold deduced from the SST field. (b) Information from the three mean fields (SST, chlorophyll-*a*, and SST gradient) synthesized and compared with Longhurst (1998) mean definition of provinces in the southwest Atlantic Ocean, with background colors as in (a). Regions obtained from histograms are indicated with solid black lines, and the previously defined province boundaries (Longhurst, 1998)—BRAZ, SSTC, SANT, SATL, FKLD—are indicated with dashed red lines (from Saraceno *et al.*, 2005).

in the paper. A total of eight different provinces were classified in the region. Five of the previously defined provinces were confirmed, but with altered boundaries. In certain regions, physical conditions or the primary production response were found to be sufficiently different from surrounding areas to require the creation of three new provinces, called the Patagonian Shelf Break (PSB), the Brazil Current Overshoot, and the Zapiola Rise region (already identified in Figure 3.13).

This chapter has pointed out a variety of ways in which ocean fronts are measured and monitored by satellite remote sensing, but it has only touched the surface of the scientific literature accumulating around this topic. In highlighting some of the more important work it should provide readers with a starting point from which to trace other existing, and future literature which can be expected to refer to those key papers. Fronts will continue to appear throughout the rest of this book, because their strong remote-sensing signatures sometimes give them a special role in revealing other ocean phenomena.

4.7 REFERENCES

Alpers, W., L. Mitnik, L. Hock, and K. S. Chen (1999), *The Tropical and Subtropical Ocean Viewed by ERS SAR*. ESA ESRIN, available at *http://www.ifm.uni-hamburg.de/ers-sar/* (last accessed April 25, 2008).

Belkin, I. M., and P. Cornillon (2003), SST fronts of the Pacific coastal and marginal seas. *Pacific Oceanogr.*, **1**, 90–100.

Belkin, I. M., and P. Cornillon (2004), Surface thermal fronts of the Okhotsk Sea. *Pacific Oceanogr.*, **2**(1/2), 6–19.

Belkin, I. M., and P. Cornillon (2005), Bering Sea thermal fronts from Pathfinder data: Seasonal and inerannual variability. *Pacific Oceanogr.*, **3**(1), 6–20.

Canny, J. (1986), A computational approach to edge-detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **8**, 679–698.

Castelao, R. M., T. P. Mavor, J. A. Barth, and L. C. Breaker (2006), Sea surface temperature fronts in the California Current System from geostationary satellite observations. *J. Geophys. Res.*, **111**(C09026), doi: 10.1029/2006JC003541.

Cayula, J. F., and P. Cornillon (1992), Edge detection algorithm for SST images. *J. Atm. Ocean. Tech.*, **9**, 67–80.

Cayula, J. F., and P. Cornillon (1995), Multi-image edge detection for SST images. *J. Atm. Ocean. Tech.*, **12**, 821–829.

Chapron, B., F. Collard, and F. Ardhuin (2005), Direct measurements of ocean surface velocity from space: Interpretation and validation. *J. Geophys. Res.*, **110**(C07008), doi: 10.1029/2004JC002809.

Chelton, D. B., M. G. Schlax, D. L. Witter, and J. G. Richman (1990), Geosat altimeter observations of the surface circulation of the Southern Ocean. *J. Geophys. Res.*, **95**, 17877–17903.

Cornillon, P., T. Lee, and G. Fall (1994), On the probability that a Gulf Stream meander crest detaches to form a warm core ring. *J. Phys. Oceanogr.*, **24**, 159–171.

Gille, S. T. (1994), Mean sea surface height of the Antarctic Circumpolar Current from Geosat data: Method and application. *J. Geophys. Res.*, **99**, 18255–18273.

Hickox, R., I. M. Belkin, P. Cornillon, and Z. Shan (2000), Climatology and seasonal variability of ocean fronts in the East China, Yellow and Bohai Seas from satellite SST data. *Geophys. Res. Letters*, **27**(18), 2495–2498.

Hughes, C. W., and E. R. Ash (2001), Eddy forcing of the mean flow in the Southern Ocean. *J. Geophys. Res.*, **106**(C2), 2713–2722.

Killworth, P. D., and C. W. Hughes (2002), The Antarctic Circumpolar Current as a free equivalent-barotropic jet. *J. Marine Res.*, **60**, 19–45.

Lee, T., and P. Cornillon (1995), Temporal variation of meandering intensity and domain-wide lateral oscillations of the Gulf Stream. *J. Geophys. Res.*, **100**(C7), 13603–13613.

Lee, T., and P. Cornillon (1996a), Propagation and growth of Gulf Stream meanders between 74°and 70°W. *J. Phys. Oceanogr.*, **26**, 206–224.

Lee, T., and P. Cornillon (1996b), Propagation and growth of Gulf Stream meanders between 75°and 45°W. *J. Phys. Oceanogr.*, **26**, 225–241.

Legeckis, R. (1977), Oceanic Polar Front in the Drake Passage: Satellite observations during 1976. *Deep-Sea Res.*, **24**, 701–704.

Legeckis, R. (1978), A survey of worldwide sea surface temperature fronts detected by environmental satellites. *J. Geophys. Res.*, **83**(C9), 4501–4522.

Legeckis, R., C. W. Brown, and P. S. Chang (2002), Geostationary satellites reveal motions of ocean surface fronts. *J. Mar. Syst.*, **37**, 3–15.

Longhurst, A. (1998), *Ecological Geography of the Sea*. Elsevier, New York.

Maximenko, N. A., B. Bang, and H. Sasaki (2005), Observational evidence of alternating zonal jets in the world ocean. *Geophys. Res. Lett.*, **32**, L12607, doi: 10.1029/2005GL022728.

Moore, J. K., and M. R. Abbott (2002), Surface chlorophyll concentrations in relation to the Antarctic Polar Front: Seasonal and spatial patterns from satellite observations. *J. Mar. Syst.*, **37**, 69–86.

Moore, J. K., M. R. Abbott, and J. G. Richman (1997), Variability in the location of the Antarctic Polar Front (90°–20°W) from satellite sea surface temperature data. *J. Geophys. Res.*, **102**(C13), 27825–27833.

Orsi, A. H., T. W. Whiworth, III, and W. D. Nowlin, Jr. (1995), On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep-Sea Res. I*, **42**, 641–673.

Romeiser, R., and H. Runge (2007), Detailed analysis of ocean current measuring capabilities of TerraSAR-X in several possible along-track InSAR modes on the basis of numerical simulations. *IEEE Trans. Geosc. Remote Sensing.*, **45**, 21–35.

Romeiser, R., S. Ufermann, and W. Alpers (2001), Remote sensing of oceanic current features by synthetic aperture radar: Achievements and perspectives. *Ann. Télécommun.*, **56**(11/12), 661–671.

Romeiser, R., H. Breit, M. Eineder, H. Runge, P. J. Flament, K. de Jong, and J. Vogelzang (2005), Current measurements by SAR along-track interferometry from a space shuttle. *IEEE Trans. Geosc. Remote Sensing.*, **43**, 2315–2324.

Saraceno, M., C. Provost, A. R. Piola, J. Bava, and A. Gagliardini (2004), Brazil Malvinas Frontal System as seen from 9 years of advanced very high resolution radiometer data. *J. Geophys. Res.*, **109**(C05027), doi: 10.1029/2003JC002127.

Saraceno, M., C. Provost, and A. R. Piola (2005), On the relationship between satellite-retrieved surface temperature fronts and chlorophyll a in the western South Atlantic. *J. Geophys. Res.*, **110**(C11016), doi: 10.1029/2004JC002736.

Shaw, A. G. P., and R. Vennell (2000), A front-following algorithm for AVHRR SST imagery. *Remote Sens. Environ.*, **72**, 317–327.

Shaw, A. G. P., and R. Vennell (2001), Measurements of an oceanic front using a front-following algorithm for AVHRR SST imagery. *Remote Sens. Environ.*, **75**, 47–62.

Simpson, J. J. (1990), On the accurate detection and enhancement of ocean features observed in satellite data. *Remote Sens. Environ.*, **33**, 17–33.

Sokolov, S., and S. R. Rintoul (2007a), Multiple jets of the Antarctic Circumpolar Current south of Australia. *J. Phys. Oceanogr.*, **37**, 1394–1412, doi: 10.1175/JPO3111.1.

Sokolov, S., and S. R. Rintoul (2007b), On the relationship between fronts of the Antarctic Circumpolar Current and surface chlorophyll concentrations in the Southern Ocean. *J. Geophys. Res.*, **112**(C07030), doi: 10.1029/2006JC004072.

Ufermann, S., and R. Romeiser (1999), A new interpretation of multifrequency/multipolarisation radar signatures of the Gulf Stream front. *J. Geophys. Res.*, **104**(C11), 25697–25706.

Ullman, D. S., and P. Cornillon (1999), Surface temperature fronts off the East Coast of North America from AVHRR imagery. *J. Geophys. Res.*, **104**(C10), 23459–23478.

Ullman, D. S., and P. Cornillon (2000), Evaluation of front detection methods fo satellite-derived SST data using in situ observations. *J. Atm. Ocean. Tech.*, **17**, 1667–1675.

Ullman, D. S., and P. Cornillon (2001), Continental shelf surface thermal fronts in winter off the northeast US coast. *Continental Shelf Res.*, **21**(11/12), 1139–1156.

van Leeuwen, P. J., W. P. M. de Ruijter, and J. R. E. Lutjeharms (2000), Natal pulses and the formation of Agulhas rings. *J. Geophys. Res.*, **105**(C3), 6425–6436.

5

Ocean mesoscale features: Upwelling and other phenomena

This is the third of the three chapters devoted to remote sensing of mesoscale ocean-dynamical phenomena and processes that are constrained mainly by geostrophy. It starts by showing how the phenomenon of upwelling lends itself to being explored from ocean-observing satellites. It then moves on to somewhat related ocean phenomena, such as those produced when wind blowing off the land is channeled by mountains, or those generated in the wake of isolated islands. The chapter also discusses the satellite signatures of large river plumes discharging into the ocean, and of processes in the marginal ice zone. It concludes with examples of experiments using remote sensing to observe processes that promote primary production in "high nutrient, low chlorophyll" regions of the ocean.

5.1 UPWELLING

5.1.1 The causes and consequences of upwelling

Upwelling is a well-known ocean phenomenon in which the interaction between wind stress over the sea and geostrophic forces produces conditions in which water is pumped to the surface to supply the flow of diverging surface currents, over regions with a size scale of tens to hundreds of kilometers. Upwelling water is cooler and richer in nutrients than typical surface water and the consequence is an enhancement of primary production leading to increased biological production across all trophic levels in the immediate region and beyond. Although geographically restricted, the phenomenon of wind-induced upwelling is therefore a very important process with a global-scale oceanographic impact and of considerable economic importance for those who harvest the ocean's living resources.

To understand what drives upwelling it is important to grasp the concept of Ekman transport. When the wind blows steadily across the sea surface it exerts a

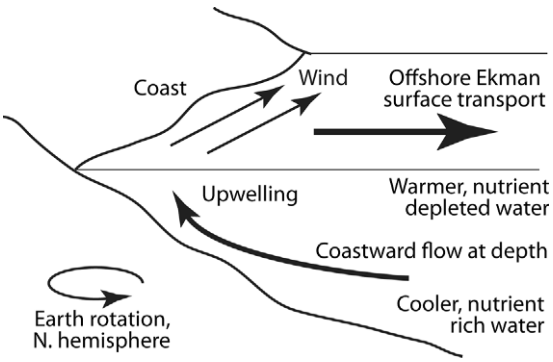


Figure 5.1. Coastal upwelling.

stress in the direction of the wind. The sea responds at first by accelerating in the direction of the wind, but since the Earth is rotating the Coriolis effect starts to swing this current towards the right of its flow direction in the northern hemisphere (to the left in the southern hemisphere). Within a few hours it is moving at 90° to the wind and its velocity adjusts so that the Coriolis force exactly balances wind stress (as in Figure 3.4 but replacing the pressure force with the wind stress force). The flow resulting from the wind is called Ekman transport. This is an essential, but not the only, ingredient necessary for upwelling to occur. The other essential factor depends on which of the two types, *coastal upwelling* or *equatorial upwelling*, we are dealing with.

In coastal upwelling the wind must be blowing parallel to the coast, or at least have a significant component parallel to the coast. Moreover its direction must be such that it flows with the coast to the left of its direction in the northern hemisphere or to the right in the southern hemisphere. This is necessary for Ekman transport to cause water to flow away from the coast (as shown in Figure 5.1). To maintain the steady offshore flow needed to geostrophically balance a steady, along-shore wind stress, it is necessary for water to flow in towards the coast at depth (where it is not subject to wind stress) and be drawn to the surface. Readers wishing to probe the dynamics of this process further should consult a standard physical oceanography text (e.g., Stewart, 2008).

Some studies refer to an upwelling index (UI). This is the term normally used to indicate the degree to which the wind is suitable for forcing upwelling. The UI is normally evaluated as the offshore Ekman volume transport per unit length of coastline averaged over a given period of time, but its precise definition varies according to the study. For example, the Environmental Research Division (ERD) of NOAA publishes 6-hourly, daily, and monthly UIs at 15 positions along the west coast of North America (ERD, 2008). These are derived from synoptic (6-hourly), sea level, atmospheric pressure, gridded fields. The surface wind stress driving Ekman transport is assumed from model analysis of atmospheric pressure. However, when individual studies generate their own UI, the UI typically has a slightly different definition although the same general meaning (Lathuilière *et al.*, 2008).

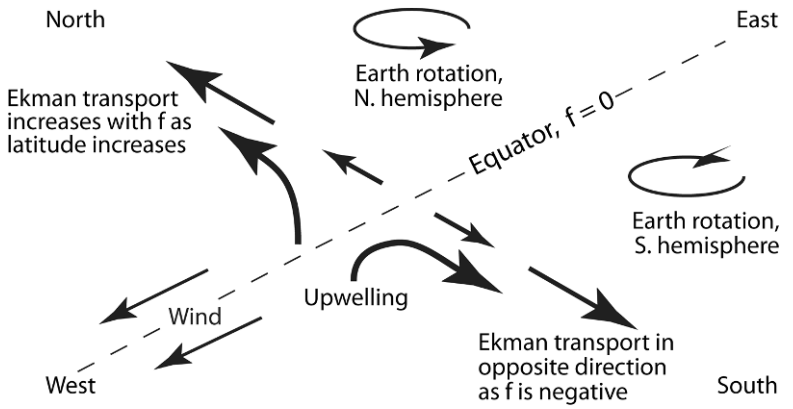


Figure 5.2. Equatorial upwelling.

Equatorial upwelling, as its name implies, occurs at the Equator. Although the Coriolis effect is zero at the Equator itself, and small at very low latitudes, the important factor leading to upwelling is the fact that the Coriolis parameter, f, changes its sign across the Equator. Thus for an easterly wind (blowing towards the west), the Ekman transport north of the Equator is towards the north, whereas to the south of the Equator it is towards the south. Whereas in general Ekman transport occurring in the open ocean away from a coast does not produce upwelling, this is not so at the Equator, where the change of direction of the Ekman flow across the Equator creates a divergence which requires upwelling to maintain the flow (as shown in Figure 5.2).

For both types of upwelling, if the wind is blowing in the opposite direction to that specified for upwelling, Ekman transport is onshore or towards the Equator. This drives coastal or Equatorial downwelling, but this is of far less importance oceanographically. In its undisturbed state, the ocean is stratified with warm surface waters overlaying cooler, denser water. Nutrients in the upper layer soon become depleted by primary production in the sunlit zone, whereas the deeper, dark waters are richer in nutrients. It requires energy to mix or advect the denser, deeper water with its increased nutrient content into the upper layer where it can support primary production. Upwelling is one of the major mechanisms across the world ocean which can achieve this. In major upwelling systems, denser, nutrient-rich water is drawn from depths of 50 m to 100 m or more. In contrast, downwelling does not contribute at all to nutrient enrichment of the surface.

On the shelf seas other processes can enrich nutrients in the photic zone, but for the deep ocean it is the zones of coastal or Equatorial upwelling where the richest sources of primary production are found, and it is these that ultimately support most animal life in the ocean. Consequently the location of upwelling zones, their spatial extent, the strength of nutrient enrichment, their seasonality, and their climatic variability are of fundamental importance for marine biology on the global scale. The detection of when and where upwelling zones occur around the world is a task that is very well suited indeed to the capability of satellite remote sensing.

5.1.2 Aspects of upwelling detected by satellites

In the most basic terms, oceanographers recognize the occurrence of upwelling by three factors: the appearance of low sea surface temperatures in a region which is favorable for upwelling dynamics as discussed above; a recent history in the previous few days of wind strong enough and in the right direction to induce upwelling; and enhancement of the phytoplankton population in the vicinity of the cooler SST. Each of these factors can be observed directly from satellites. SST is measured readily by infrared sensors with a spatial resolution of about 1 km, which is sufficient to monitor the detailed structure of fronts of cold water that bound the upwelling region. Scatterometers can provide daily or twice daily maps of wind speed and direction over the ocean, with a resolution of a few tens of kilometers. Near-surface phytoplankton populations can be detected by chlorophyll concentrations measured by ocean color sensors, at a resolution of 1 km. In other words, upwelling lends itself to being observed from satellites because its characteristic features are among those properties that ocean remote sensing is best at detecting. Figures 5.3, 5.4, and 5.5 illustrate this with examples showing, respectively, the instantaneous distribution of SST, chlorophyll, and wind vectors over the Benguela upwelling region on a particular day in March 2005.

These images demonstrate the benefits from viewing three different aspects of the upwelling phenomenon using nearly simultaneous images from different sensor types. In this case, by using MODIS data from the EOS Aqua platform, SST and chlorophyll maps are acquired simultaneously. SST provides the most direct evidence of immediate upwelling conditions. Although in general, averaged over time, the whole Benguela coastline supports upwelling for a distance of more than 2,000 km, in this particular snapshot the strongest upwelling is occurring between 33°S and 34°S. Water cooler than 14°C appears to have been upwelled recently and drawn offshore up to 50 km from the coast. Farther north the coolest water reaches only about 20 km from the coast and is warmer. From this we may infer that, at this time, upwelling is weaker between 28°S and 32°S than it is farther south.

Reference in Figure 5.5 to the wind distribution observed by QuikScat the evening before the MODIS image was acquired shows that the region off Cape Town (at 34°S) where the upwelling appears most vigorous is also where winds were strongest, about 14 m/s to 15 m/s from the south. Farther north along the coast they are weaker, also consistent with weaker upwelling. In contrast, the chlorophyll map shows that the high concentrations of biomass are found not only where the water is coldest but stretching offshore for 100 km to 200 km beyond the core of the upwelling. In practice the chlorophyll map shows the integrated history of primary production from the previous few days. Supplied by upwelled nutrients the phytoplankton population grows and is then carried away by both offshore Ekman transport and the Equatorward current parallel to the coast. Thus the highest production also stretches several hundred kilometers north of the observed upwelling center. However, this may not necessarily be entirely due to northward advection but may have been supplied by nutrients a few days earlier from an upwelling center farther north, depending on where the strongest southerly winds

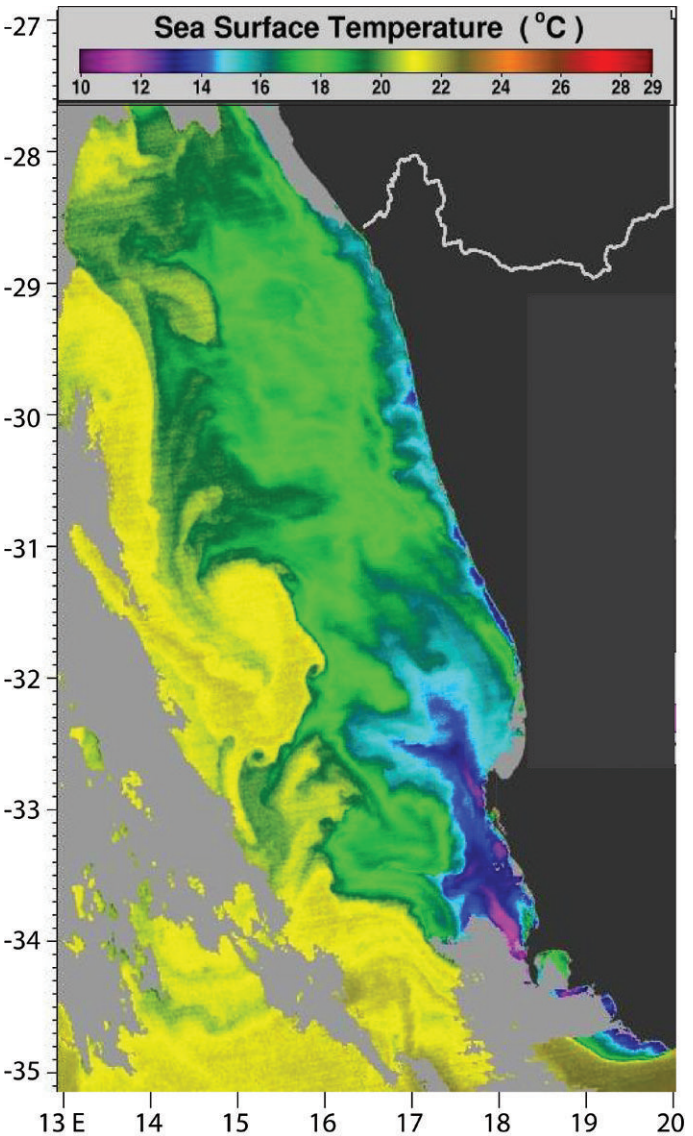


Figure 5.3. SST in the Benguela upwelling region of the southwest Atlantic Ocean, centered on 31°S, 17°W, retrieved from a single MODIS (Aqua) overpass on March 3, 2005, at the same time as Figure 5.4.

were occurring then. Although clouds prevent a clear view of the previous few days, this one example demonstrates the richness of information and insight that can be gained from multiparameter remote sensing of upwelling. Put simply, we may consider that image datasets of wind, SST, and chlorophyll tell us, respectively, about the forcing, the immediate response, and the longer term integrated result of an upwelling event.

There remain problems for detailed monitoring of upwelling processes from space, the foremost of which is the presence of cloud that restricts the use of infrared

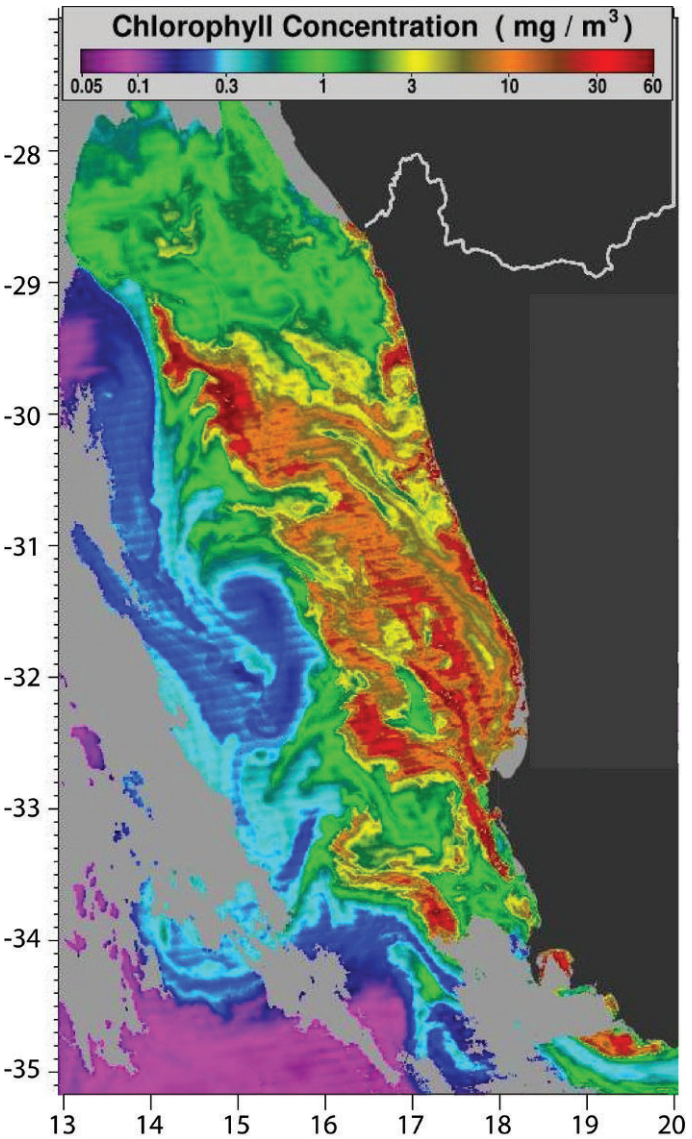


Figure 5.4. Chlorophyll-*a* concentration in the Benguela upwelling region of the southwest Atlantic Ocean, centered on 31°S, 17°W, retrieved from a single MODIS (Aqua) overpass on March 3, 2005, at the same time as Figure 5.3.

and ocean color sensors. In some parts of the world low cloud or sea fog may be associated with the cool SST found in the upwelling zone. One way of ameliorating this is to build up daily SST composite maps from infrared sensors on geostationary satellites, provided cloud does not persist throughout the day (Castelao *et al.*, 2006). This is feasible because most of the important upwelling regions are found at the same latitude as subtropical gyres and within the view of a satellite parked over the

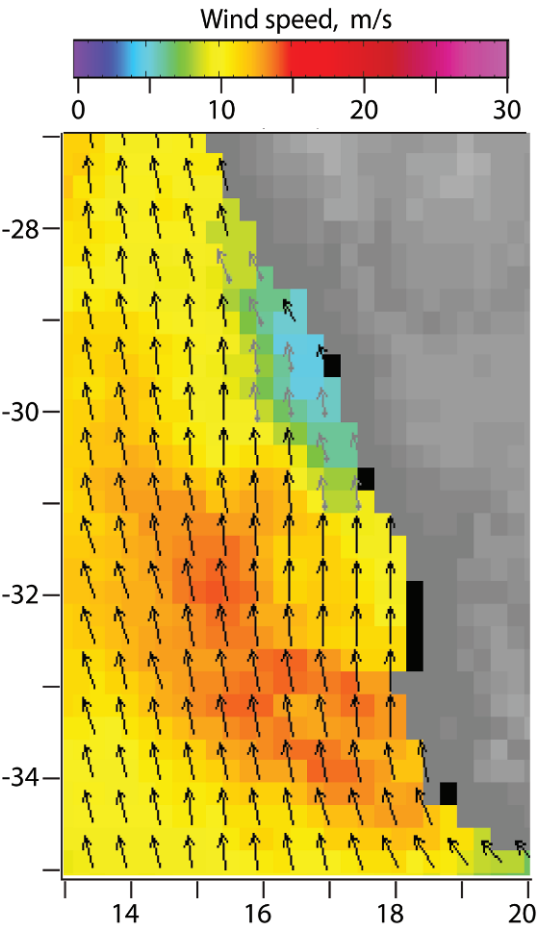


Figure 5.5. Wind vectors from QuikScat over the Benguela upwelling region from the evening overpass of QuikScat on March 2, 2005. QuikScat data are produced by Remote Sensing Systems and sponsored by the NASA Ocean Vector Winds Science Team (adapted from a graphic image map acquired from *www.remss.com*)

Equator. If cloud is very persistent then microwave radiometry can be tried, but only for very large regions of upwelling that extend a long way offshore. Because their spatial resolution is no better than 50 km and their accuracy is compromised within up to 100 km of the coast by stray radiation from the land, microwave radiometers are not able to define the spatial structure of upwelling in detail and would not detect any of the cool water found within 20 km of the shore throughout the whole coastline shown in Figure 5.3.

In principle the sea surface height signature at the frontal edge of major upwelling regions should be measurable by altimetry. When upwelling is established it pushes the warm, surface, mixed layer away from the coast, replacing it with colder, denser water. When this happens the sea level lowers over the upwelling region, sloping down steeply at the front where cold, upwelled water meets warm, offshore, surface water (as shown in Figure 5.6). This surface slope is maintained in

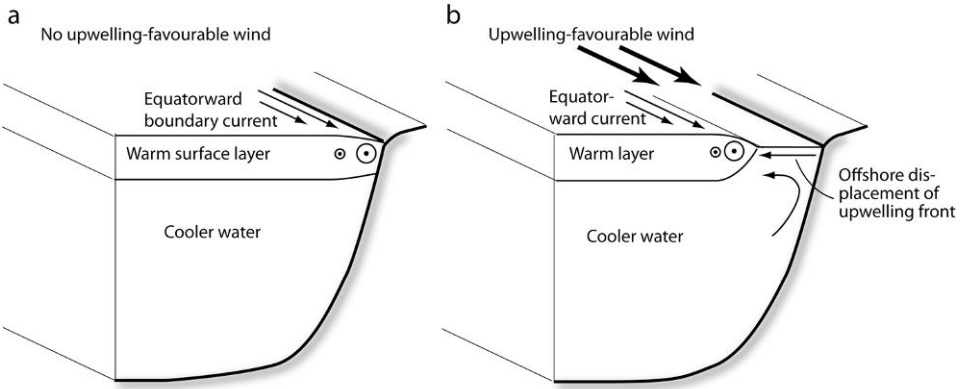


Figure 5.6. Schematic of the cross-section through an eastern ocean margin showing the typical thermal structure, the equatorward boundary current, and associated SSH when (a) there is no upwelling and (b) when there are upwelling-favorable winds.

geostrophic balance by an equatorward jet that is an essential element of the upwelling system. Therefore the upwelling zone does have a sea surface height signature. As shown in the figure, this moves offshore when there are upwelling-favorable winds, and disappears or moves inshore when the upwelling ceases. If altimetry could resolve SSH at short lengthscales with sufficient accuracy then it would play a useful role in monitoring the position of the upwelling frontal edge, but so far this has not been achieved. Current research activity (Madsen *et al.*, 2007; Bouffard *et al.*, 2008) is attempting new analyses of existing altimeter data to improve accuracy close to the coast in semienclosed seas, and these techniques may eventually be applied so that altimetry can monitor upwelling and meandering of the upwelling front. Wide-swath altimeters (see section 11.5.5 of *MTOFS*), if they are ever developed and deployed, would also be valuable for detecting the upwelling frontal edge. Meanwhile, altimetry already provides very good evidence for the changes in sea level associated with upwelling in the eastern equatorial Pacific, and its reduction during an El Niño (as discussed in Section 11.2), although that occurs on a larger spatial scale than most upwelling zones around the world. Altimetry is also applied to studies of eastern boundary currents which, although distinct from the actual upwelling, are closely associated with the fate of the upwelled water. Altimetry identifies the variability of coastal currents out to 500 km from the coast (Strub *et al.*, 1987; Strub and James, 2000) and the occurrence of offshore jets that carry productive water towards the open ocean.

SAR images also have potential for detecting fronts present in upwelling systems, but this approach has not so far given oceanographers any information that they cannot obtain more easily from other sensors. It remains to be seen whether new developments in SAR measurement of currents, using the Doppler centroid method mentioned in Chapter 4, will bring something new to the study of upwelling.

5.1.3 Upwelling regions of the world seen from space

The main upwelling regions of the world are shown in Figure 5.7, mapped onto a 6-year, average annual mean, SST map from MODIS Aqua. Although a long-term average map tends to smooth out the mesoscale detail of instantaneous snapshot images such as Figures 5.3 and 5.4, and also weakens any features that are strongly seasonal, it is used deliberately here because any small-scale features that it does reveal must be those that are strong most of the time and whose location is rather constant. Hence the narrow ribbon of cold water along certain coasts, between 3°C to 6°C cooler than the adjacent ocean, provides strong evidence of persistent upwelling.

In fact these correspond to places where prevailing winds are known to be upwelling-favorable throughout the year. They are found at the eastern margins of the Pacific and Atlantic Oceans at the latitudes of subtropical gyres, in both the northern and southern hemisphere. The four regions of the world where upwelling occurs persistently are the Canary coast off northwest Africa, the Benguela coast off southwest Africa, the Peru coast, and the Oregon coast. Beyond these are regions where upwelling is strong during part of the year (normally the summer months) but not for all of the year, and their signatures in the annual SST map are somewhat more diffuse. Thus, the wider coastline of Oregon and California, the Chile coast, and the Iberian coast are labeled as seasonal.

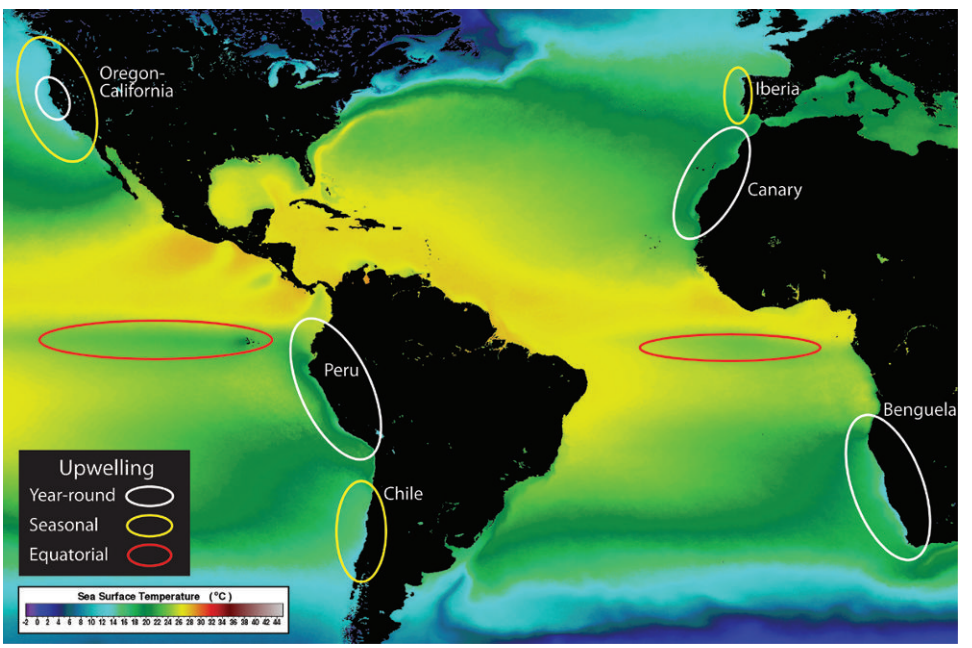


Figure 5.7. The major upwelling zones around the world. The background image is the 6-year average annual mean SST 9 km map produced from MODIS data.

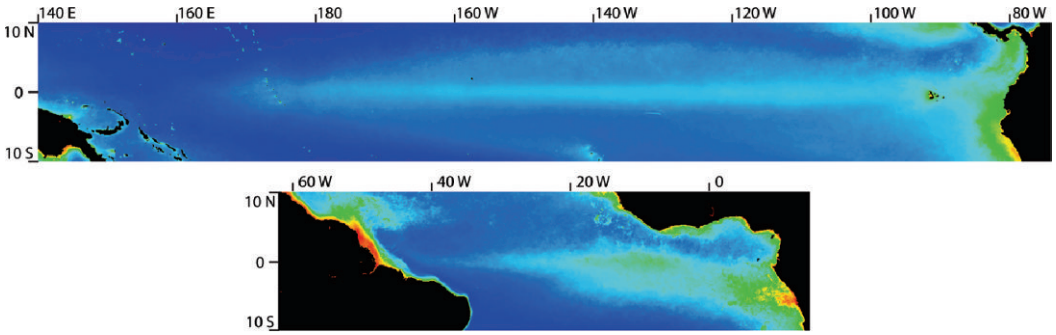


Figure 5.8. Evidence of equatorial upwelling in the 6-year cumulative (January 1, 2002–February 29, 2008) average map of chlorophyll concentration as measured by the MODIS sensor on Aqua, showing the Equatorial Pacific (*above*) and the Equatorial Atlantic (*below*), both between 10°N and 10°S. The color scale for this map is the same as for the chlorophyll-*a* maps of Figures 5.9–5.12 (MODIS data downloaded from the NASA Ocean Color website at *http://oceancolor.gsfc.nasa.gov*).

Elsewhere on the eastern boundaries, as well as on the western boundaries of all the oceans, the prevailing winds most of the year blow in the wrong direction to cause upwelling. Figure 5.7 shows only the East Pacific and East Atlantic Oceans because there are no permanent upwelling regions in the Indian Ocean, the West Pacific, and around Australasia, although that is not to say that upwelling never occurs there. Seasonal upwelling is experienced in particular locations around the world at specific times of the year, when the prevailing wind blows in the correct direction. For example, in the Indian Ocean and South East Asia there are a number of places where monsoon winds produce strong upwelling for a short part of each year, making a considerable impact on regional oceanography. These are mentioned again in Section 11.3.

Equatorial upwelling is found in the Pacific Ocean whenever there are easterly winds (westward blowing), and it is remarkable in global climatologies of chlorophyll (see Figure 5.8) to find a narrow, straight line along the Equator where chlorophyll is enhanced, even though the cool signature of upwelling is rather diffuse in satellite SST climatologies like Figure 5.7. There is also some evidence of equatorial upwelling in the Atlantic, although not as clearly as the Pacific, while the Indian Ocean shows no sign of this phenomenon. During El Niño events, satellite data show how equatorial upwelling in the Pacific is interrupted (as discussed in Chapter 11). We shall consider in Chapter 7 how ocean color remote sensing offers a new perspective on the impact which equatorial upwelling and permanent coastal upwelling centers have on global biological production, Here we explore the four coastal upwelling zones in more individual detail.

In each case, for a particular month when upwelling at that location is expected to be fairly strong, monthly mean SST distribution is presented alongside the monthly mean chlorophyll map. Figure 5.9 shows this in the case of the Benguela region off southwest Africa, in February 2004. Superficially this may seem very similar to snapshots of a single, cloud-free day shown in Figures 5.3 and 5.4.

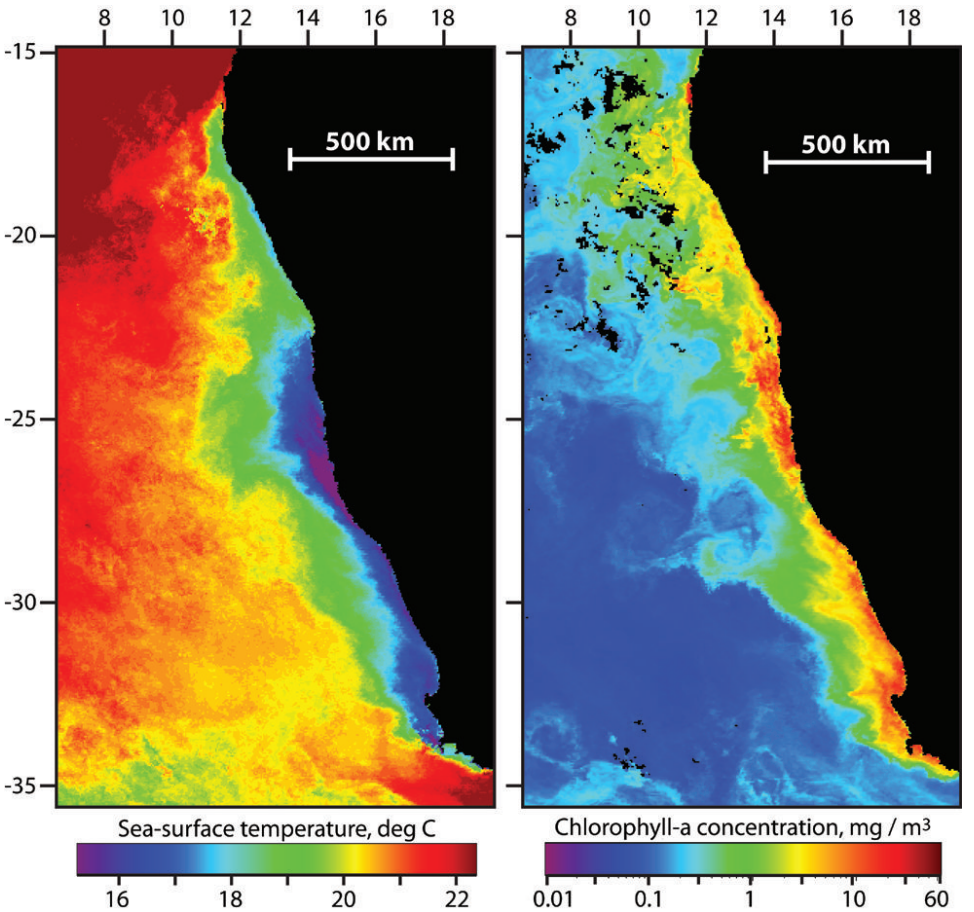


Figure 5.9. Benguela upwelling monthly average for February 2004 of SST (*left*) and chlorophyll (*right*) both from the MODIS datasets (produced using extracts from level 3 mapped images downloaded from the NASA Ocean Color website at *http://oceancolor.gsfc.nasa.gov*).

However, Figure 5.9 is an integrated record of upwelling over a whole month, providing a more complete view of how upwelling is influencing the primary production for the whole region. It smooths out day-to-day variation in the strength of upwelling and in the location of upwelling centers, which depend on the wind and its orientation in relation to different headlands and parts of the coastline. Figure 5.9 shows that some upwelling is experienced all along the coast to as far north as 14°S, although it is strongest between 23°S and 34°S. Note that from a monthly averaged SST map it is not possible to distinguish between a part of the coast where upwelling is persistent but not very strong, and another part where upwelling is intermittent but very strong when it happens. What can be inferred is that the coldest regions on the map indicate locations of the greatest rate of nutrient-rich water being drawn to

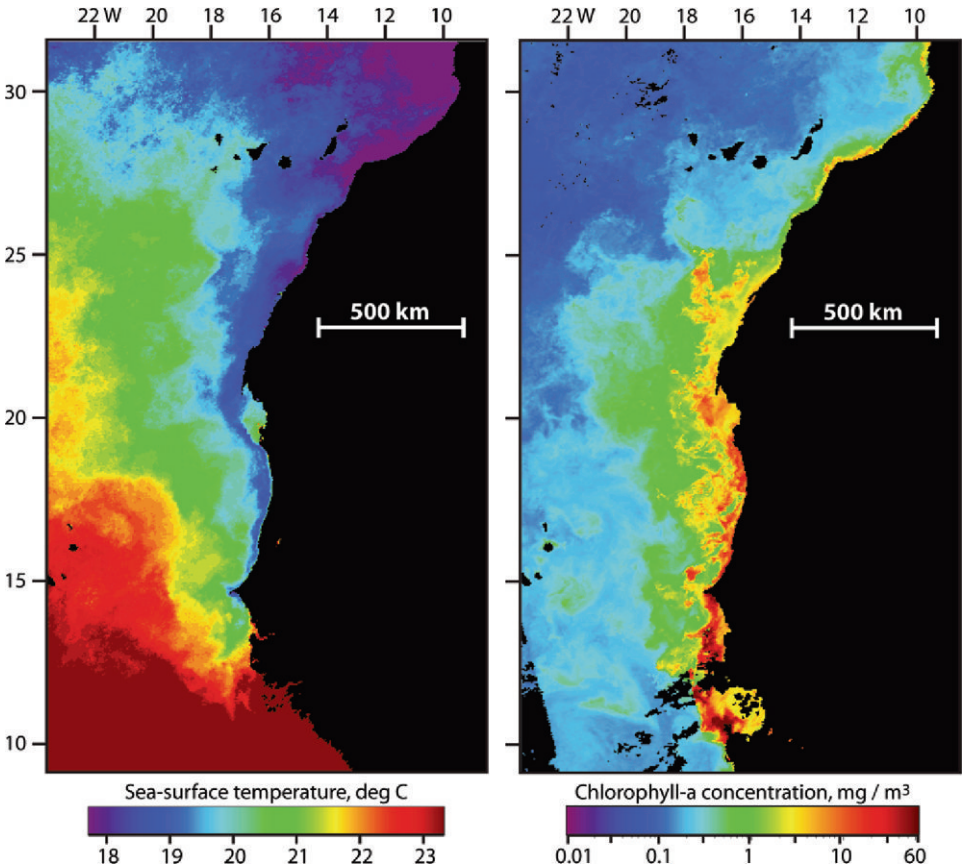


Figure 5.10. The Canary upwelling along the coast of northwest Africa. Monthly average for February 2004 of SST (*left*) and chlorophyll (*right*) both from the MODIS datasets (produced using extracts from level 3 mapped images downloaded from the NASA Ocean Color website at *http://oceancolor.gsfc.nasa.gov*).

the surface where it supports enhanced production. Certainly in this example there is a strong correlation between the distribution of cold water and chlorophyll concentration, Moreover, it appears that where upwelling is strongest its influence is felt farthest offshore. Between 23°S and 27°S, for example, it seems that cooling as a result of upwelled water reaches 500 km offshore, and phytoplankton enhancement is swept even farther beyond this.

Figure 5.10 depicts the upwelling off the northwest African coast, the so-called "Canary coast", also in February 2004. Although the prevailing winds are upwelling-favorable here, the coastline orientation is very variable at a lengthscale of 50 km to 100 km, with a number of prominent headlands. Thus quite small changes in wind direction can shift the upwelling action from one side of a headland to the other, according to which stretch of coastline aligns closest to the actual wind.

In this case there is an absence of upwelling south of the prominent headland at 21°N. Because there is also a fairly steep gradient of SST with latitude, the temperature of upwelled water also changes. Thus between 15°N and 20°N the upwelled water is about 18.5°C (blue on Figure 5.10), whereas north of 24°N it is less than 18°C (purple on the image). Nonetheless it appears from the chlorophyll image that the band of 18.5°C to 19°C water reaching 200 km off the coast is spreading nutrient-rich water offshore implying that upwelled water close to the coast is entrained into the equatorward Canary Current. A number of studies of this upwelling region have made use of ocean remote-sensing data, both SST (Van Camp *et al.*, 1991; Nykjaer and Van Camp, 1994) and chlorophyll from ocean color (Pradhan *et al.*, 2006; Lathuilière *et al.*, 2008).

The potential for upwelling along the Pacific coast of South America spans more than 30° of latitude (as shown in Figure 5.11 for April 2004). The strongest, most persistent upwelling is along the Peruvian coast from 5°S to 18°S, where there is a thin line of upwelled water all along the coast. Chlorophyll concentration does not exactly match the strength of upwelling as indicated by the temperature of coastal upwelled water. Around 12°S to 13°S, where coastal temperature is at its warmest, production appears to be strongest and extends farther offshore. Presumably advection by the along-shore current is influencing where nutrient-enriched, upwelled water is carried, but from only a single month's snapshot no firm conclusions should be made. It is clear between 5°S and 10°S and off Chile between 25°S and 35°S that where enhanced production stretches a long way from the coast it is being carried in offshore jets which also have a cooler SST signature. Satellite SST data have been used for various studies of upwelling in this region (e.g., Thomas *et al.*, 2001; Marin *et al.*, 2003).

A similar picture of streamers of productive water carried offshore by perturbations of the coastal current is evident during July 2004 in the Oregon–California upwelling region (shown in Figure 5.12). The Oregon and California Coastal Current and its relation to upwelling have been extensively studied and many recent papers make use of satellite data alongside, and to provide a wider context for, *in situ* measurements in understanding the impact of upwelling on the region (see, e.g., Barth and Wheeler, 2005 and papers in the associated special section of *J. Geophys. Res.*).

5.1.4 Using satellite data in upwelling research

The brief survey of the world's major upwelling zones given in the previous section makes it clear that we cannot isolate upwelling in its narrowest sense (i.e., wind-driven surface divergence producing an offshore transport of water drawn from depth) from the wider context of coastal currents and offshore jets that determine the longer term fate of upwelled water and the enhanced primary production it can support. Much recent research on upwelling processes is therefore concerned with the dynamics of eastern boundary currents. Increasingly such work draws on satellite data to complement *in situ* measurements and numerical modeling analyses. There have been investigations of offshore jets or filaments that carry upwelled, enriched

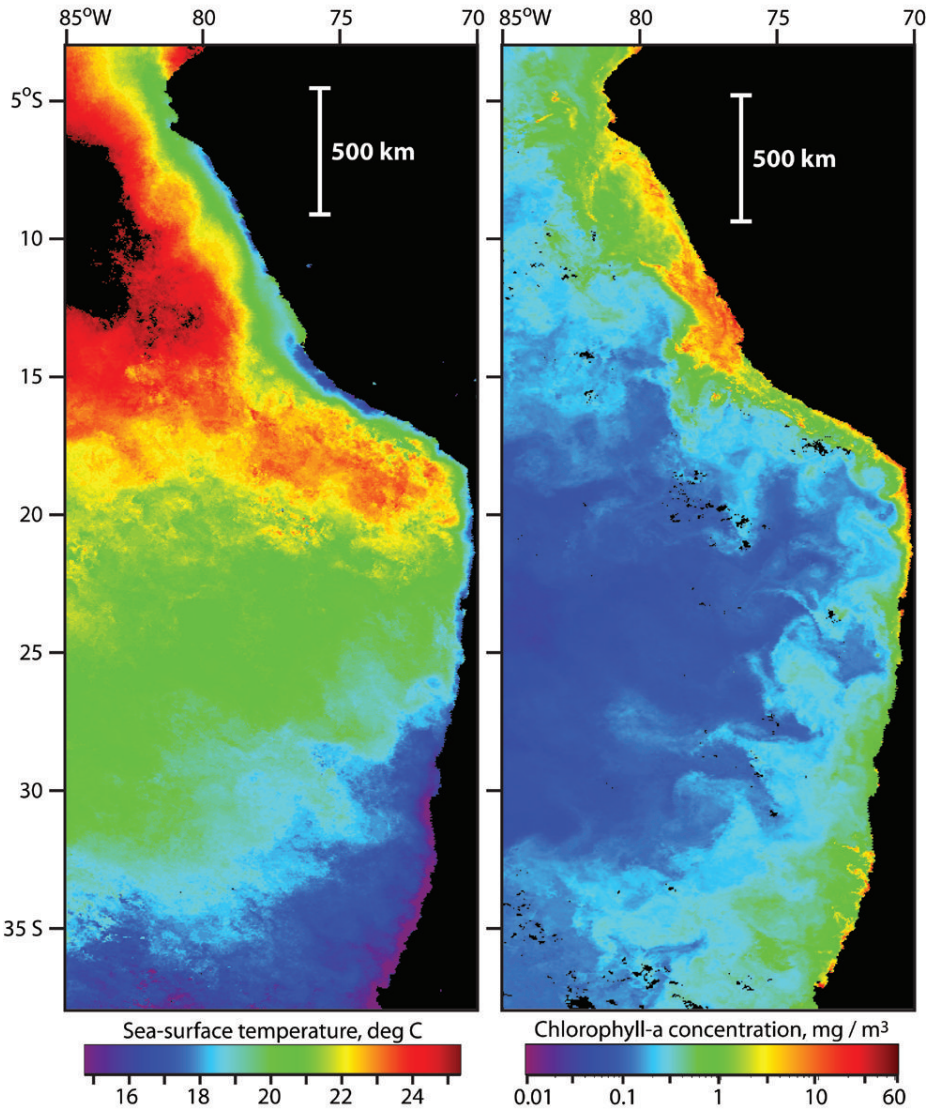


Figure 5.11. Upwelling along the coasts of Peru and Chile. Monthly average for April, 2004 of SST (*left*) and chlorophyll (*right*) both from MODIS datasets (produced using extracts from level 3 mapped images downloaded from the NASA Ocean Color website at *http://oceancolor.gsfc.nasa.gov*).

water well away from the coast (Strub *et al.*, 1991; Barth *et al.*, 2000) and studies concerned with the effect of local topography on this process (Barth *et al.*, 2005). Marin *et al.* (2003) use satellite data to study the occurrence of upwelling shadows, locations where upwelling is inhibited by the relation between wind direction and coastline orientation. Castelao and Barth (2006) explored the role of wind stress

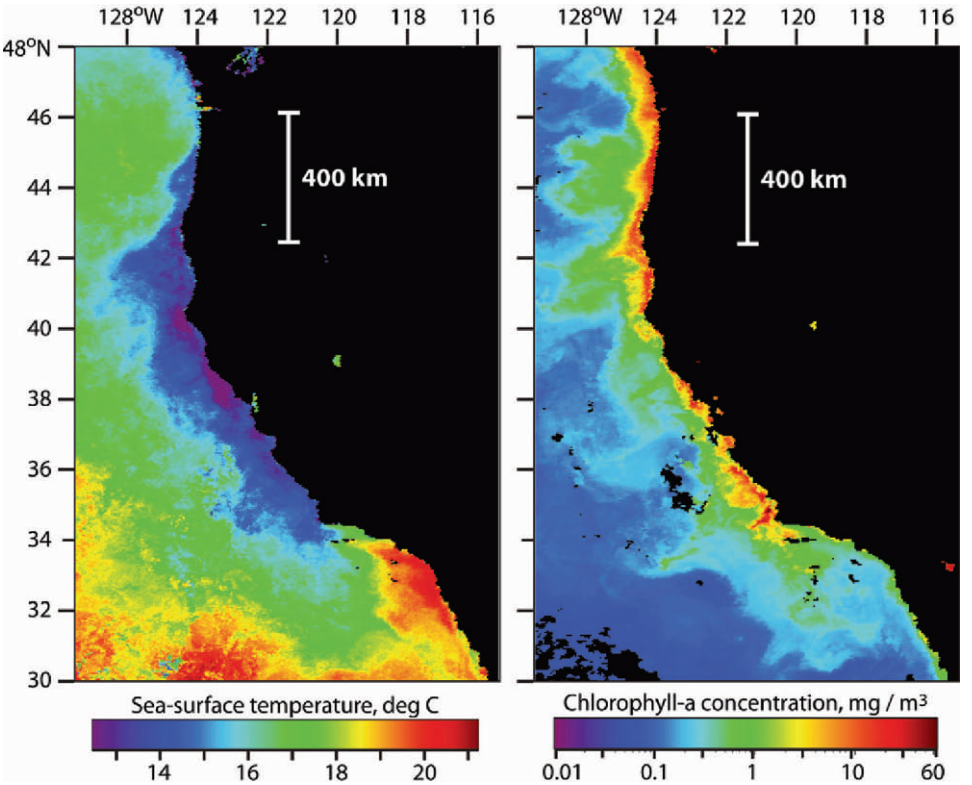


Figure 5.12. Upwelling along the Oregon and California coasts. Monthly average for July 2004 of SST (*left*) and chlorophyll (*right*) both from MODIS datasets (produced using extracts from level 3 mapped images downloaded from the NASA Ocean Color website at *http://oceancolor.gsfc.nasa.gov*).

curl in generating a local upwelling zone in the southwest Atlantic on the Brazilian coast.

Moreover, scientific studies of upwelling are concerned with much more than simply whether or not the phenomenon is occurring, and require detailed information that can only come from *in situ* measurements and sampling. For example, those concerned with the dynamical processes of upwelling onset and evolution, the role of bottom topography, interaction with along-shore currents and the outbreak of offshore plumes of cool surface water need to record changes in temperature and salinity distribution over depth. Direct measurement of currents using ADCP is also needed. Marine chemists trying to understand biogeochemical processes need to measure the concentration of nutrients in upwelling water and the rate at which they are consumed once they reach the photic zone. Marine ecologists need to determine the species of phytoplankton in blooms following an upwelling event, and to observe the structure of zooplankton communities that are supported.

Marine biologists study greatly enhanced fish stocks, major populations of sea mammals, and large populations of sea birds which all depend on enhanced primary production in upwelling regions. None of these factors can presently be measured or observed by remote sensing. Nonetheless, the design of field experiments to obtain such information can be greatly enhanced if satellite data are used in real time to identify when and where best to sample (Barth and Wheeler, 2005).

Finally it is worth pointing out that satellite data are potentially very useful for analyzing changes in the climatology of upwelling because, with long time series available of daily SST, chlorophyll, and winds, it becomes feasible to accumulate time series of upwelling behavior at fine spatial resolution along coasts. There are already examples of such work, studying decadal changes in upwelling climatology (Nykjaer and Van Camp, 1994; Santos *et al*., 2005), and exploring variability in relation to the El Niño index (Thomas *et al*., 2001).

Lathuilière *et al*. (2008) used satellite data extensively, augmented with climatological *in situ* measurements of nitrates, to examine and compare seasonal and intraseasonal variability of chlorophyll concentration between three different segments of the northwest African coast where upwelling is encountered. Three different upwelling indices were defined specifically for this work:

- Cross-shore Ekman transport (CSET) was specified as $(\tau_{al}/\rho f)$; where τ_{al} is the locally alongshore component of wind stress in the direction towards the Equator; ρ is water density; and f is the local Coriolis parameter. Wind stress was obtained from 1/2° QuikScat wind stress products delivered by CERSAT, IFREMER. Wind stress evaluated 50 km off the shelf break were the closest reliable data to the coast.
- The temperature-based upwelling index (TUI) was defined to be the difference between offshore SST (taken as mean SST in a band located within 500–700 km from the coast) and SST on the shelf. SST data were obtained from the AVHRR Pathfinder data product from NOAA.
- The chlorophyll extension index (CEI) was defined as distance from the coast where the 9 km, SeaWiFS, level 3, chlorophyll concentration is equal to three times the mean value in the region 1,200 km to 1,500 km offshore.

By performing climatological analyses on these three indices—which essentially represent, respectively, wind forcing of upwelling, ocean dynamic response, and resulting impact on primary production, using consistent data over a 5-year period (2000–2004)—some useful insights could be gained about the way different parts of the coast responded to atmospheric and oceanographic changes in the wider region. As available satellite datasets are extended in time we may expect to see more research of this character.

5.2 WIND-DRIVEN, OFFSHORE, DYNAMICAL FEATURES

Wherever the wind blows strongly over the ocean the sea responds, but the variability of the wind usually means that the ocean's response is not sustained

long enough to become a characteristic behavior unless there is some factor to cause the same response to repeat in the same location. The phenomenon of coastal upwelling discussed in the previous section was an example of this—the presence of the coast ensures that the upwelling process occurs wherever and whenever the wind is in the appropriate direction to produce offshore Ekman transport. Another situation which we consider here is where the wind is constrained and steered by the land terrain so that it tends to blow over the sea in specific, repeated patterns, producing a recurring characteristic ocean response.

Those handling small boats near the coast are familiar with the way wind blowing off the land may be strongest where valleys open onto the coast and are weaker where the land rises steeply from the shore in cliffs and mountains. In such cases the spatial distribution of wind stress over the coastal sea area has a banded structure that can be seen clearly in SAR images (such as figures 10.36 and 10.37 in *MTOFS*). At the scale of a few kilometers such behavior does not extend far offshore and has little impact on ocean hydrography. There are, however, a few places where the lengthscale of the wind channeling effect is sufficiently large to induce mesoscale changes in ocean hydrography and a geostrophic dynamic response. In the examples to be considered here, revealed as plume-like patterns of cooler, upwelled water stretching for several hundred kilometers from the Central American coast towards the southwest into the Pacific Ocean, the pattern is frequent enough and sufficiently strong to appear in the multiyear mean SST map shown in Figure 5.7.

The cause of this is best discussed in the context of Figure 5.13, a weekly average map of wind speed and direction based on QuikScat observations. This shows three regions on the Pacific coast of Central America in each of which the wind is blowing strongly offshore within a contained region. Each region can be traced back to a gap in the high ground and mountains that elsewhere produce a barrier for low-altitude winds. Each gap corresponds to the lowest point of one of the three isthmuses, Tehuantepec, Papagayo, and Panama. White parallel bars have been marked on the land where those gaps occur. In fact this is a very intermittent wind pattern which occurs for only a few days at a time and usually in boreal winter months. On monthly mean wind maps the pattern is not very strong, and it is unusual to find a weekly average like Figure 5.13 in which strong winds blow through all three gaps at the same time. Nonetheless the meteorological conditions in which flow occurs through one or other of the gaps is well documented (Chelton *et al.*, 2000) from the analysis of scatterometer data. The northwestern of the three gaps—where what is known locally as "Norte" winds blow a strong southward jet over the Gulf of Tehuantepec—is the most active of the three. Despite this intermittency, the fact that the ocean always responds in the same location is sufficient to lower long-term mean temperature in affected areas by a few degrees Celsius relative to the surrounding seas. Oceanographers have been aware for a long time of these local dynamic responses to sudden wind jets, but when infrared images of SST became available they showed more clearly just how far the impact of the cool, upwelled water reached (Legeckis, 1988).

Figure 5.14 provides a finer resolution view during a Norte event. It is clear from the SST map how the wind jet forces the ocean to respond with an offshore jet of

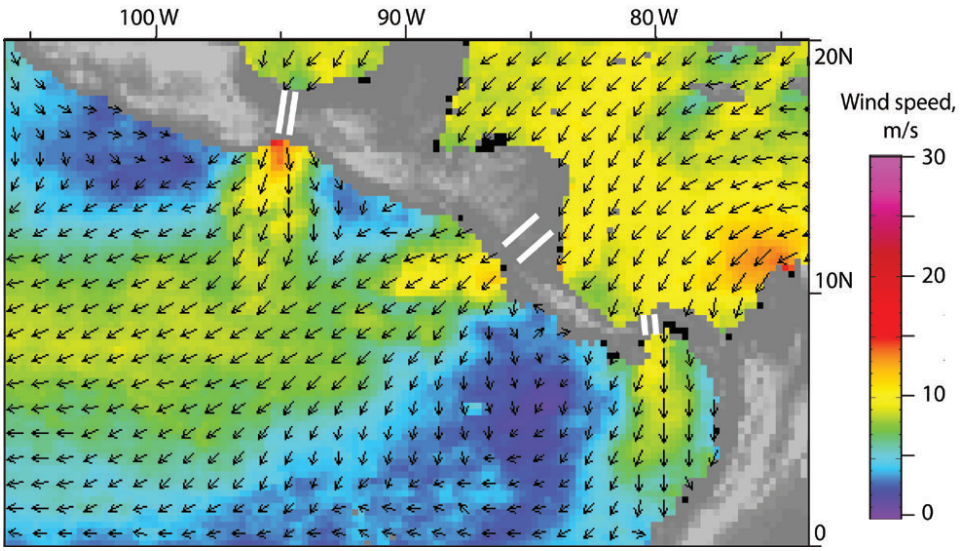


Figure 5.13. Weekly average wind speed and direction from QuikScat over Central America for week ending February 18, 2006. The white parallel bars show the three gaps in the mountain chains through which the wind is steered. QuikScat data are produced by Remote Sensing Systems and sponsored by the NASA Ocean Vector Winds Science Team (adapted from a graphic image map acquired from *www.remss.com*).

cold, upwelled water more than 150 km wide. This extends over 400 km offshore and gradually turns clockwise. In this case there is a 150 km diameter circular region of intermediate surface temperature water which is often found in this location; and its frontal edge has on other occasions been clearly revealed in SAR images (Martinez Diaz de Leon *et al.*, 1999) (such as figure 10.53 in *MTOFS*).

The chlorophyll image in the lower panel of Figure 5.14 demonstrates the considerable impact which upwelling has on the distribution of phytoplankton. Not only does the wind jet draw up nutrients to enhance the primary production of the region, but stirs them into complex patterns which, apart from their interesting dynamics, are indicative that phytoplankton are transported well beyond the localized area where upwelling occurs. This suggests that these phenomena have an impact on the marine biology and fisheries of a much wider region.

They also have an impact on regional mesoscale dynamics through the periodic generation of eddies whose westward propagation has been monitored using altimetry in the case of the Tehuantepec jet (Zamudio *et al.*, 2006), and also Papagayo (Palacios and Bograd, 2005). Studies of the ocean response in Papagayo Bay have also been made using satellite-derived SST data and buoys (Ballastero and Coen, 2004). It seems that we owe quite a lot of what is known about these interesting and influential phenomena to the capacity to view them from space using several different classes of sensor.

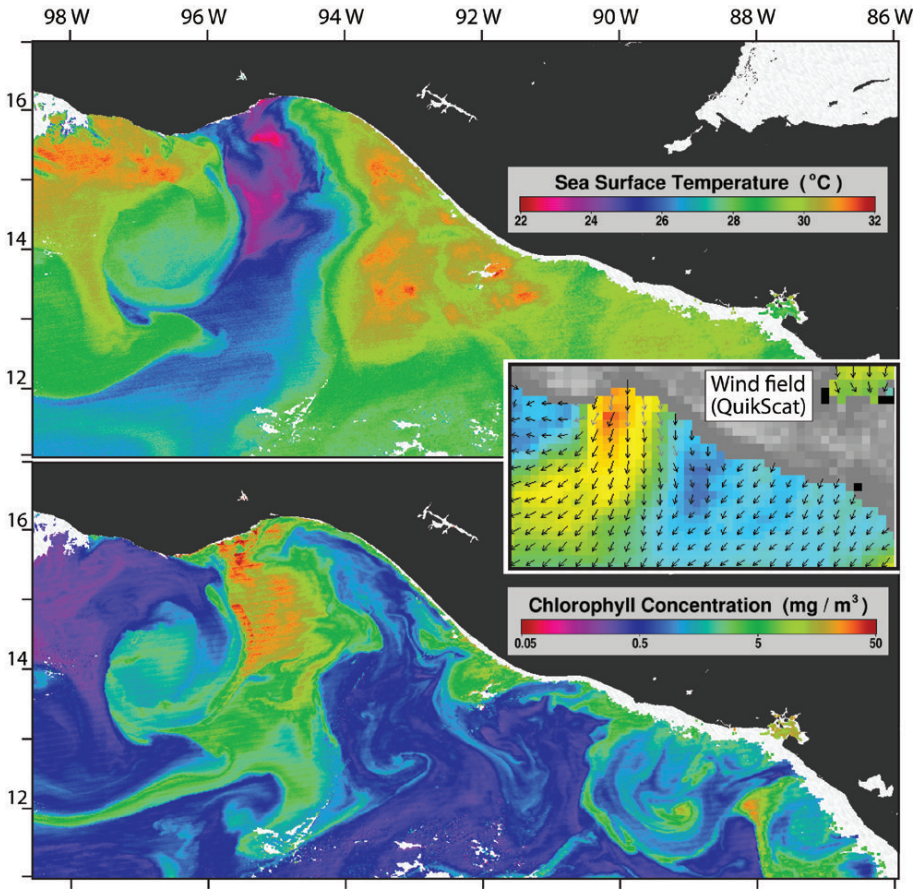


Figure 5.14. Upwelling during a "Norte" event in the Gulf of Tehuantepec off the Pacific coast of Mexico. *Upper panel* is the SST measured by Aqua-MODIS on November 15, 2004. *Lower panel* is the chlorophyll concentration derived from the ocean color channels of the same sensor at the same time. The *inset panel* shows the average wind field for the week ending November 13, 2004, derived from QuikScat. The color scale for wind speed is the same as that defined in Figure 5.14. The MODIS images were obtained from the NASA Ocean Color website (QuikScat data acquired from Remote Sensing Systems).

5.3 LARGE RIVER PLUMES

When rivers discharge into the ocean, their distinctive temperature or color may be detectable at the surface for a distance until they are mixed and diluted with the water of the open sea. This gives river plumes the potential for a signature in remotely sensed images, particularly in SST or ocean color. The extent of such signatures depends on volume flux of the river, contrast with properties of the open sea, and the relative density of river and seawater. If river discharge is more

buoyant it will tend to overlay seawater and spread over a wider area before becoming mixed into the sea and losing its distinctive characteristics. Mostly the extent of river plumes is no more than a few kilometers, and is one of the factors to consider when interpreting color or temperature images of shelf seas (as mentioned in Chapter 14). Here we briefly consider those major world rivers whose volume flow is so great that their influence on water properties measured from space extends beyond the local sea into which they discharge and can be found hundreds of kilometers out into the open ocean, giving them a signature with dimensions comparable with the mesoscale phenomena being considered in this and the previous two chapters.

Figure 5.15 is a chlorophyll map derived from Aqua MODIS over the North Brazil coast in September 2006. The mouth of the Amazon, between 0° and 1°N, is rather obscured by cloud in this image. However, it appears that the rich nutrients of the river are supporting a strong phytoplankton bloom as the discharging water is entrained into the North Brazil Current (NBC) which flows north-eastwards along the Brazilian coast. However, this may not be the only source of nutrients (McGillicuddy *et al.*, 1995), and it has also been noted that the color signature is strongly influenced by river-derived CDOM as well as chlorophyll, so chlorophyll retrievals from the standard Case 1 algorithm should be treated with caution. Nonetheless images such as this have been used to quantify estimates of plume dispersal (Hu *et al.*, 2004).

At the time of the year when this image was acquired the NBC is constrained, in the upper 100 m of the water column, by the opposing Atlantic North Equatorial Counter-Current (NECC) flowing eastwards near the surface at about 6°N (Bischof *et al.*, 2004). At the confluence of the NBC and NECC, retroflection of the NBC may occur and eddies are formed. This is what makes the image in Figure 5.15 appear so spectacular. The NECC is strongest between June and September and diffuse or nonexistent between northern winter and spring, so flow patterns such as this would not be found throughout the year. The NECC transports the productive water across the Atlantic towards the west African coast, ultimately feeding the Guinea Current. Satellite measurements of the plume area in the Atlantic have been used to estimate the size of the associated atmospheric carbon sink (Cooley *et al.*, 2007). However, while the image appears to give the impression that the enriched flow meandering across the Atlantic is almost like an extension of the Amazon River, that would be a misleading way of thinking about it. The typical discharge of the Amazon is of order 1×10^5 m^3/s to 2×10^5 m^3/s (i.e., 0.1–0.2 Sv), whereas the NECC transports about 16 Sv drawn from the NBC. Thus while the color of the Amazon serves as a tracer of the NBC–NECC confluence, in fact only about 1% of the water volume being transported along this path has come from the Amazon.

There are no other river discharges as spectacular as this, but there are still a number of river plumes around the world for which remote sensing has been used to study their influence on the properties of the adjacent ocean. For example, the plume of the Mississippi River, which discharges into the Gulf of Mexico, has been traced by its suspended sediment signature in the visible waveband of AVHRR satellite data (Walker, 1996). Off the coast of Oregon, the Columbia River plume can also

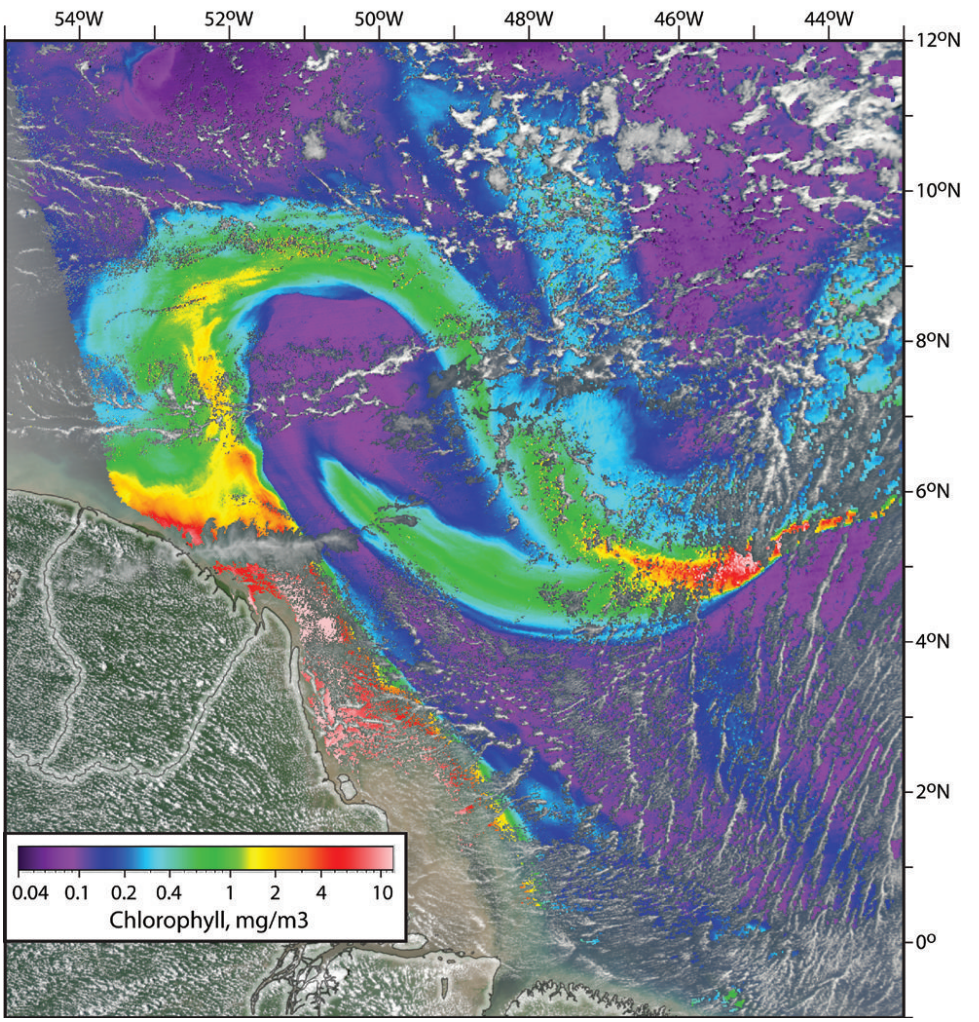


Figure 5.15. Chlorophyll map of west Equatorial Atlantic Ocean derived from Aqua MODIS on September 30, 2006 showing the track of nutrient-rich water from the River Amazon.

sometimes be detected in ocean color imagery. As an example of the unexpected phenomena that can be discovered by careful analysis of satellite data, SAR images have revealed large internal waves created at the edge of the Columbia River plume and radiating out into the North Pacific (Nash and Mourn, 2005).

5.4 ISLAND WAKES

Another example of a remotely sensed phenomenon that is tied to a specific location in the ocean is that of island wakes. An isolated island, or small group of islands, can

sometimes be seen in satellite images to have attached to them a signature over the adjacent ocean, typically streaming away in a dominant direction giving the appearance of a wake. There are three ways in which this may occur. The first is a purely atmospheric effect. Islands that have fairly high mountains, typically of volcanic origin, produce disturbances in the wind flow that have a satellite signature (e.g., as cloud streets trailing downwind in visible wavelength radiometer images, or as wind shadow zones visible in SAR images as low backscatter regions where the wind stress is reduced behind the island—see also Chapter 9). Typically the disturbance extends a few hundred kilometers at most beyond the island in the prevailing wind flow direction, but occasionally a high mountain can influence the wind flow much farther over the ocean. This is the case in Hawaii where the presence of the islands is credited with reducing the westward wind flow of easterly trade winds in a shadow behind the mountains that stretches for 3,000 km, resulting in the formation of a narrow, easterly, ocean surface current (Xie *et al.*, 2001), apparently far removed from the island whose presence caused it.

The second mechanism is where the wind shadow wake behind high mountains on an island creates a response locally in the ocean. This may be superficial, as when a diurnal thermocline develops in the wind shadow area so that warm patches of SST are clearly seen in thermal infrared imagery. However, in the case of steady winds creating a persistent wind shadow region the shear between shadow and nonshadow regions produces Ekman convergence and divergence (as shown in Figure 5.16). This leads to upwelling and downwelling zones, the former appearing cooler on SST images. Upwelling also brings increased nutrients that may lead to enhanced primary production with an ocean color signature. This can account for the patterns of temperature and color that appear on the downwind side of island chains such as the Hawaiian and Canary Islands (Aristegui *et al.*, 1997; Barton, 2001). Barton *et al.* (2001) show how the combination of SAR images revealing wind shadow and wind

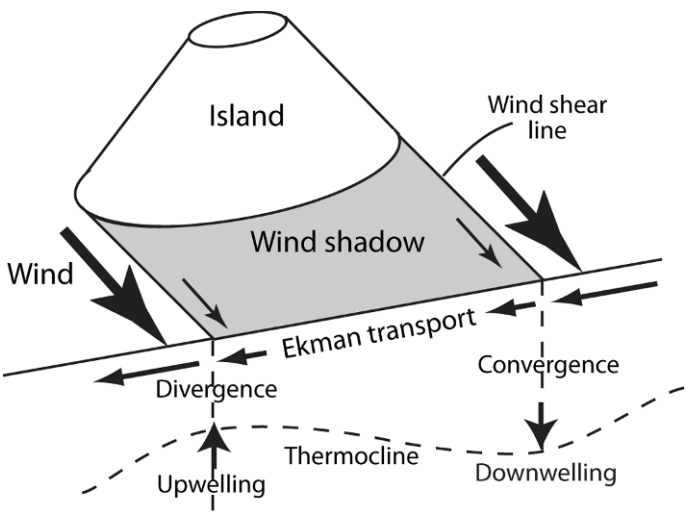


Figure 5.16. Schematic cross-section through the sea showing wind shadow, shear lines, and upwelling driven by Ekman transport downwind of an isolated oceanic island.

shear zones, coupled with SST images, provide considerable insight into this phenomenon.

The third mechanism is the shedding of ocean eddies as a steady ocean current flows past isolated volcanic islands that rise steeply from the ocean floor. This has been modeled as a classic fluid-dynamic problem in which the depth of the ocean, the diameter of the island, and the Coriolis parameter can be shown to determine whether eddies are shed from the island or else a recirculation region is trapped behind the island (Dietrich *et al.*, 1996; Dong *et al.*, 2007). Eddy motion leads to vertical transport of water, and so upwelling occurs in part of the wake region, providing the potential for enhanced production. In shallow seas a similar process of vertical stirring is driven by tidally oscillating flow past an island (as mentioned in Chapter 13). Eddies and wake regions show up in thermal or color imagers. Figure 5.17 shows a characteristic deep-ocean example, the Galapagos Islands, as seen in a chlorophyll map retrieved from SeaWiFS. It is interesting to note differences between the three images spaced a few days apart, illustrating just how variable the dynamical process can be.

Barton (2001) notes that for islands with high mountains where the prevailing wind is parallel to the dominant ocean current it is difficult to distinguish the ocean current effect from the wind-driven mechanism. In the middle of a deep oligotrophic ocean where there might otherwise be very little production, hydrodynamic interaction between the ocean and the island engineers a locally rich source of primary production that may have great significance for the marine biology of a wider area. Such "oases" of production stand out in global satellite maps of chlorophyll concentration. Among many examples that have been reported since ocean color images became available are the enhanced production west of the Maldives as the westward Indian Ocean North Equatorial Current flows past (Sasamal, 2007), and the "island mass" effect of Madeira (Caldeira *et al.*, 2002).

5.5 ICE EDGE PHYTOPLANKTON BLOOMS

Phytoplankton blooms that occur at the melting ice edge in high-latitude seas are important phenomena, the study of which increasingly makes use of remote sensing. Although strictly not mesoscale phenomena they have elements in common with coastal upwelling and river plumes, and so it is convenient to mention them in this chapter. Apart from the short section on sea ice in Chapter 11.4, the scope of this book does not extend to polar remote sensing for which there are several comprehensive texts available (e.g., Lubin and Massom, 2006; Rees, 2006). What concerns us here is primary production that occurs in the region adjacent to the ice edge, during the period each year when the ice edge is receding as a consequence of the annual melt.

Until the mid-1970s it was assumed by biological oceanographers that primary production needed to support the relatively large biomasses of the higher trophic levels in the Southern Ocean was the result of summer sunlight and nutrients supplied by large-scale upwelling at a circumpolar divergence. Further observations

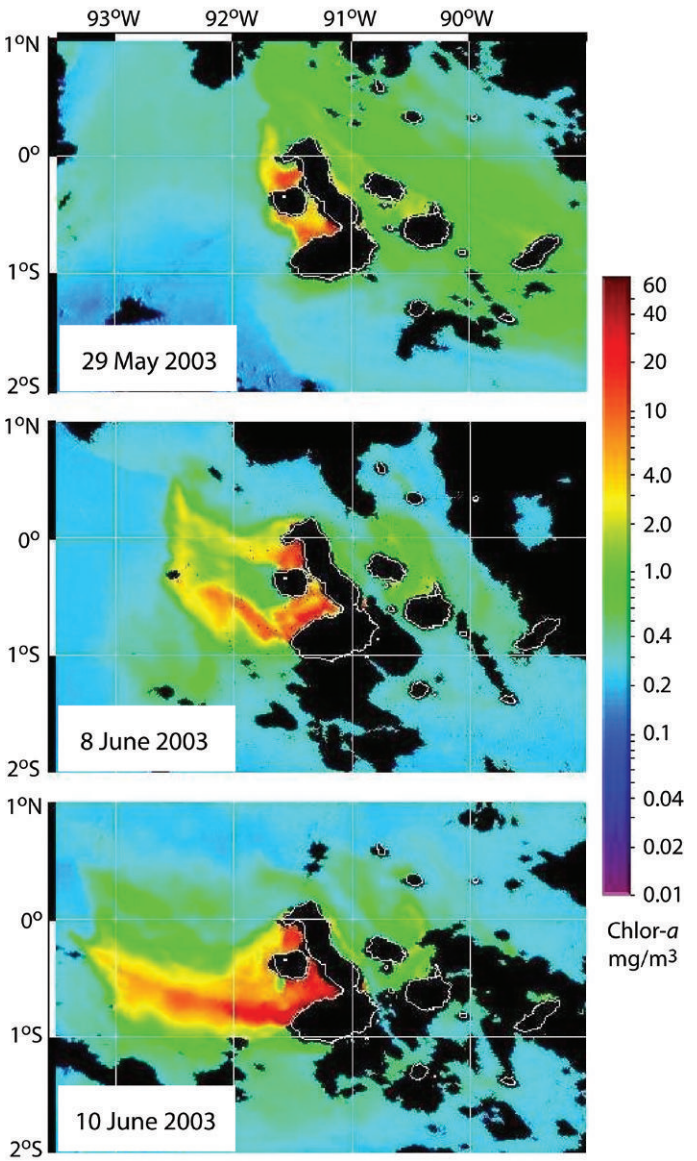


Figure 5.17. Three maps of chlorophyll-*a* concentration over the Galapagos region, derived from SeaWiFS ocean color data, spanning 13 days in May/June 2003, illustrating the "island wake" effect.

revealed that primary production across the Southern Ocean is generally low, implying that there must instead be strong, but very localized, blooms supporting the massive concentrations of krill on which large populations of birds and sea mammals thrive around Antarctica (Smith and Nelson, 1986). A proposed candidate for this type of bloom was the intense production occurring at the melting ice edge in summer. This had previously been largely overlooked because of the difficulty of reaching and monitoring in the marginal ice zone. In the last two decades this

phenomenon has been much more closely studied from special cruises in the Southern Ocean, which follow the ice edge south during the spring melt. Studies have also made use of satellite ocean color observations, initially with historic CZCS data (Comiso *et al.*, 1990). and later with SeaWiFS (Buesseler *et al.*, 2003) to integrate and extrapolate local results to the whole circumpolar region, This was one of the first phenomena to be explored as soon as ocean color data became routinely available from SeaWiFS in the late 1990s (Moore *et al.*, 1999; Moore and Abbott, 2000). Satellite data confirmed that mean chlorophyll concentrations were quite low throughout the Southern Ocean ($<$0.3–0.4 mg m^{-3}) and the only phytoplankton blooms where chlorophyll concentration exceeded 1.0 mg m^{-3} were found in coastal/shelf waters, the vicinity of major Southern Ocean fronts (as discussed in Chapter 4), and areas associated with the seasonal sea ice retreat.

The key mechanism which encourages a strong bloom to occur is the layer of low-salinity melt water left behind in spring as the seasonal ice starts to melt and the ice edge recedes poleward. It contains nutrients, and may even provide a source of micronutrients, such as iron. Most important is the strong stratification at the halocline that is stable to erosion by wind mixing and typically is shallower than, or comparable with, the depth of the photic zone at the time of the bloom's peak (Buesseler *et al.*, 2003). Thus phytoplankton cells in the surface layer are exposed to strong solar illumination and are not mixed down out of the light. This provides the conditions for an intense bloom that lasts a short time but, as the ice retreats further, the zone of intense production follows and the high production rate is maintained. This persists through the summer and early autumn as long as the ice continues to melt. Although on any day the bloom is limited to the zone within a few tens of kilometers of the ice edge (as shown in Figure 5.18), nutrients are effectively utilized from a wide area of previously ice-bound ocean as the melting edge sweeps across it. Given that the typical area in the seasonal ice zone (SIZ) that freezes and melts each year is about 16×10^6 km^2, this process is able to make a large contribution to integrated primary production of the Southern Ocean. Presumably this would be reduced if the area of the SIZ were to decrease in response to climate change.

Following these developments in the Antarctic, a number of other satellite-related studies of primary production in the SIZ have developed the subject further. As Chapter 7 shows, estimation of column-integrated primary production or biomass from satellite chlorophyll data needs to make assumptions about the typical depth distribution of chlorophyll: these have been revised for the special case of the density profile associated with blooms at the ice edge (Engelsen *et al.*, 2004) based on measurements in the Barents Sea. There has been other research in northern polar seas including Labrador and Newfoundland seasonal ice (Wu *et al.*, 2007) and the Greenland coast (Heide-Jørgensen *et al.*, 2007), the latter making use of MODIS data with algorithms adapted for high latitudes. Mustapha and Saitoh (2008) have examined how the strength of blooms in the Okhotsk Sea varies with the timing of sea ice retreat, in the context of the impact on scallop fish-farming. Finally a study of the relationship between chlorophyll blooms and the production of dimethyl sulfide (DMS) in the upper ocean leading to atmospheric sulfate aerosols has shown how the typical climatological distributions of these variables

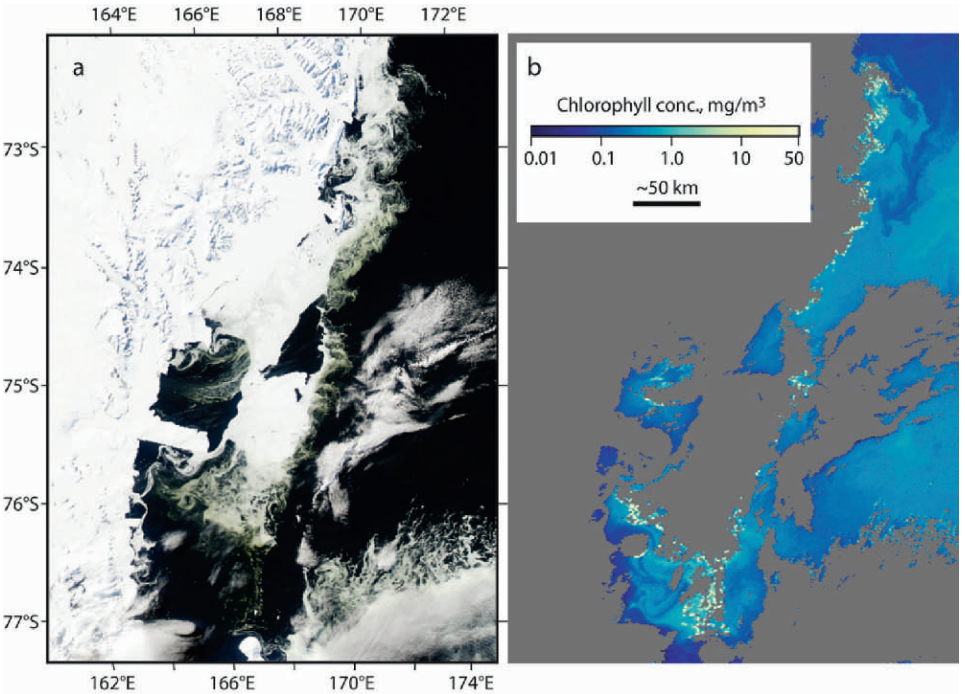


Figure 5.18. Aqua MODIS view of the Ross Ice Shelf on February 24, 2008. (a) Level 1 real-color composite. (b) Level 2 chlorophyll distribution. This shows plankton blooms associated with the melting ice edge (MODIS data from which this image was created were acquired from NASA's Ocean Color website at *http://oceancolor.gsfc.nasa.gov*.).

in the Southern Ocean are affected by the melting behavior of sea ice (Gabric *et al.*, 2005).

5.6 REMOTE SENSING IN IRON LIMITATION STUDIES

One of the major questions of interest to biogeochemical oceanographers at present is to understand why it is that some parts of the ocean have fairly low primary production even though they are not depleted of macronutrients at the end of the growing season and appear to have adequate light available in the summer months (Venables and Moore, 2009). It is suggested that some of the so-called high-nutrient, low-chlorophyll (HNLC) waters, such as those encountered in the Southern Ocean, are limited by a lack of iron. Over the last decade there have been a number of research experiments to study the iron limitation hypothesis. Because of the logistical challenges of measuring phytoplankton populations in the Southern Ocean, against a background of typical mesoscale variability, some of these studies have made explicit use of remote sensing in their experimental design. Here we mention just

two of these experiments, but they serve as pointers towards how, more generally, satellite observations can not only be used to provide background climatological knowledge about mesoscale variability, but also have the capacity to supply near real–time data as an integral part of an experimental program.

One way to test the iron limitation hypothesis has used experiments to artificially enrich a region by depositing iron in a suitable form and monitoring the response of the phytoplankton bloom (Boyd *et al.*, 2007). The scale of the resulting bloom is large enough to be detected in level 2 ocean color data products. In some experiments, satellite data have been used explicitly to monitor the long-term fate of the bloom for several weeks after the experiment when prohibitive costs have not allowed *in situ* sampling to continue beyond the initial stages. For example, in the SOIREE experiment (Abraham *et al.*, 2000), the fertilized patch was detected several times in the few weeks following the iron release on February 9, 1999. Figure 5.19 shows the clearest view acquired on March 23. By this time the bloom, which started as a small circular region, had been drawn into a ribbon shape some 150 km long and then advected by mesoscale turbulent motion into this semicircular shape with a diameter of about 50 km. Maximum chlorophyll concentration is estimated to be 3 mg/m^3. Further analysis of typical SeaWiFS data for the previous 2 years confirmed that such magnitudes had not been detected previously in this region where the mean value was 0.20 ± 0.06 mg/m^3. Although frequent cloud cover prevents a full dynamic description of the growth and evolution of the bloom, occasional clear views like this provide an overall perspective that assists enormously in interpretation of biogeochemical field measurements. Since then similar use has been made of SeaWiFS and MODIS data during the SERIES and SOFeX experiments (Coale *et al.*, 2004).

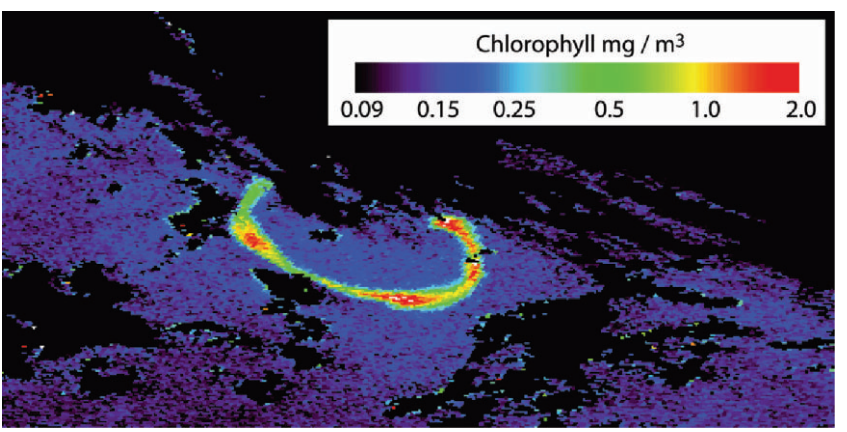


Figure 5.19. SeaWiFS-derived chlorophyll image of the bloom resulting from the SOIREE iron enrichment experiment. This image was acquired on March 23, 1999, several weeks after the initial release. The bloom is centered at about 141°E, 60.5°S. The diameter of the semiring feature is about 50 km (original image obtained from NASA Ocean Color website).

Another way to understand more about iron limitation in the Southern Ocean is to study regions where production is higher than elsewhere, in order to determine by what processes the iron is being supplied. This was the broad rationale of the CROZEX project (Pollard *et al.*, 2007a), at the core of which were two cruises of the RRS *Discovery* between 2004 and 2005, in the region around the Crozet Islands, 2,000 km southeast of South Africa. The project tested the hypothesis that HNLC conditions are lifted in the area by a natural source of iron, most probably from sediment around the islands and on the plateau. Concerning the subject of this chapter, what is especially noteworthy about this large multidisciplinary experiment is the intrinsic use that was made of satellite data through all stages of the project in: setting the concept; planning the program; real-time cruise management; analyzing data from a variety of sources; and placing conclusions in a seasonal and climatological context.

Thus, for example, it was satellite ocean color data that in the first place highlighted the importance of the Crozet Plateau as a region of high production in otherwise HNLC waters. The availability of several years of SeaWiFS archive allowed the interannual variability of the timing and location of blooms to be explored. Where mesoscale turbulence dominates it is useful to have a reasonable awareness of the envelope of variability so that contingency plans can be prepared in case the bloom is earlier or later than expected in relation to the time when the research vessel is available. Once the cruise was under way, near real–time satellite data revealed the evolving regional picture and were used to change the cruise plans so as to encounter important phenomena.

Satellite data, when post-processed to higher accuracy, were also integrated into data analysis and interpretation of dynamical and advective processes, as clearly revealed in two of the key papers coming from the project (Pollard *et al.*, 2007b; Venables *et al.*, 2007). In particular, the use of DUACS absolute dynamic topography (ADT) data from multimission altimetry (as also mentioned in Chapter 4 in relation to studying the fronts in the ACC) was vital for understanding the path of ARGO floats, and also for understanding the spatial distribution of chlorophyll concentration. In Figure 5.20 the streamlines derived from altimetry-based ADT data are superimposed on simultaneous chlorophyll distribution at the time of the field experiment, showing a strong connection between chlorophyll and velocity fields. Although ADT data are not generally suitable for defining flow streamlines until we have improved knowledge of the geoid, there are reasons discussed in Chapter 4 for expecting this approach to be acceptable in the ACC. Here it is evident from the match between patterns of chlorophyll and ADT structure that the ADT is reliably revealing the velocity structure at lengthscales down to less than 200 km.

Figure 5.20 demonstrates how chlorophyll appears to be strongly constrained by the water flow, and helps to make sense of the apparently random variability. The flow follows an S-bend around the shallower topography, and flow over the Plateau itself is weak. This was a strong factor in reaching the conclusion that production over the Crozet Plateau and to the east is entirely consistent with the stirring up of iron into those waters as they pass over the Plateau. Weak circulation means that

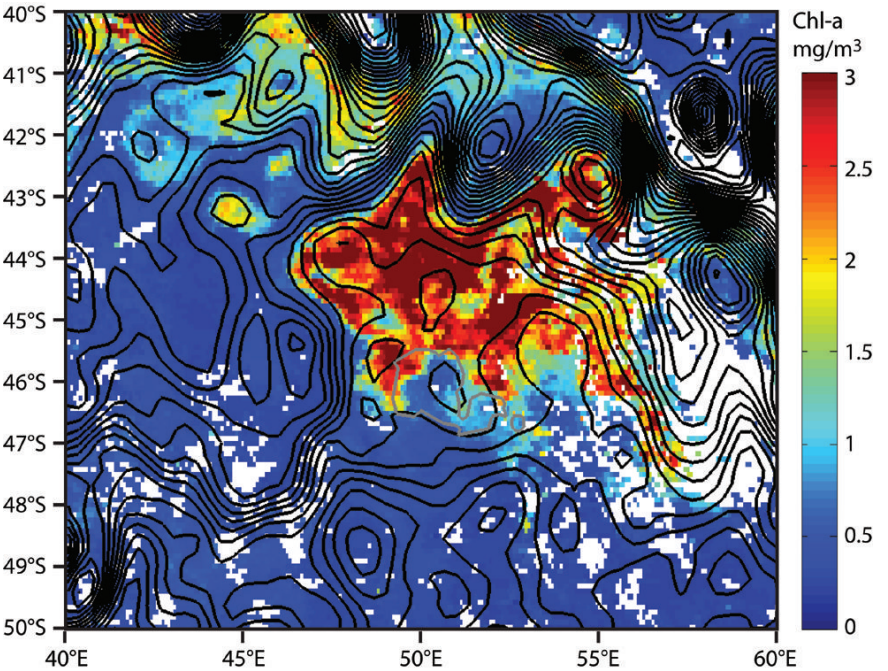


Figure 5.20. Contours of the ADT field, equivalent to streamlines of the surface absolute velocity field, overplotted onto the satellite-derived chlorophyll-*a* field surrounding the Crozet Plateau, for the week of October 23–30, 2004, during the CROZEX experiment (image provided by Hugh Venables, British Antarctic Survey).

iron-fertilized water can build up concentrations in the bloom area during winter, ready to initiate a bloom over a wide area once spring stratification develops.

Another interesting analysis applied to chlorophyll data was to determine the date at which the bloom peaked for each pixel in the area (as shown in Figure 5.21). The relation between these maps of bloom timing and flow patterns point the way to understanding ecosystem behavior in relation to the flow. Finally, given nearly a decade of ocean color and other satellite data, it was possible for the understanding gained in a single year from a specific experiment to be applied and interpreted in relation to interannual variability in the region. Figure 5.22 shows, for example, how differences in main flow streamlines has altered chlorophyll distribution from year to year.

5.7 MAKING THE MOST OF SATELLITE DATA FOR MESOSCALE STUDIES: CONCLUSIONS FROM CHAPTERS 3–5

At the end of three chapters about using satellite data to study mesoscale processes and phenomena, it should be evident that many applications have already matured

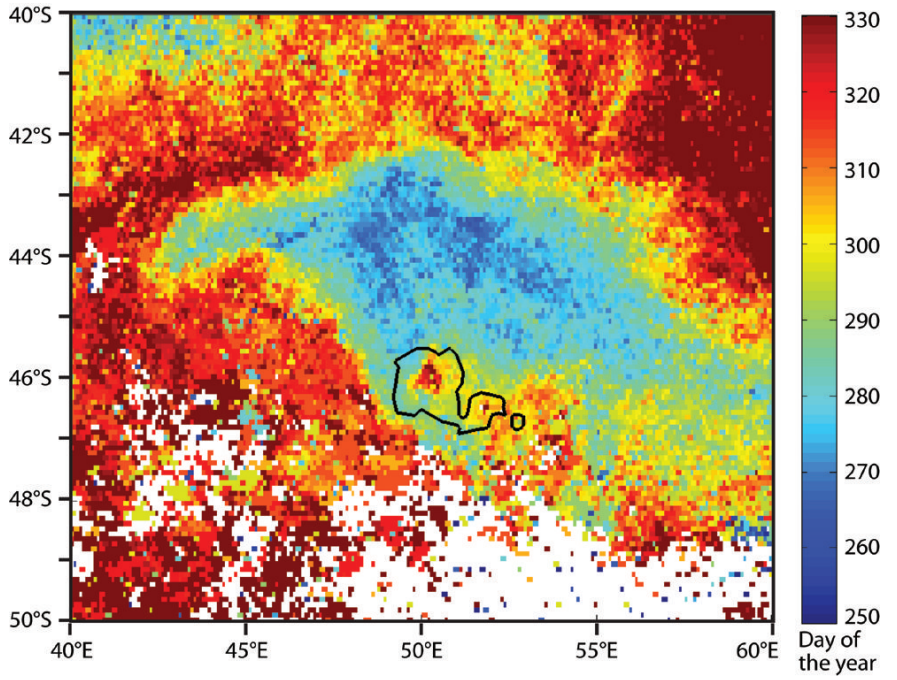


Figure 5.21. Map showing how the timing of bloom initiation (i.e., the average day of the year—over 1997–2005—when chlorophyll concentration measured from SeaWiFS first exceeds 1 mg/m^3) varies with position over the Crozet Plateau (image provided by Hugh Venables, British Antarctic Survey).

into tools used in mainstream oceanography. It is encouraging to find examples, as in the previous subsection, where the creative use of satellite data applied in conjunction with conventional *in situ* hydrography is providing the extra insight that leads to innovative results and better understanding of complex features. The availability of ocean color data that can be compared with SST and SSH is also opening up new understanding of how biological and chemical processes are steered by water circulation.

These chapters have not probed all aspects of the subject. For example no mention has been made of the technique of maximum cross-correlation (MCC) in which large-scale flow is detected from the movement of small-scale structures of SST or color between one overpass to the next (see, e.g., Domingues *et al.*, 2000; Barton, 2002; Bowen *et al.*, 2002; Emery *et al.*, 2003). Although the promise of ADT from altimetry is very exciting, in practice these two methods for defining the details of ocean circulation in an eddy-infested ocean are likely to go well together as complementary tools.

We have not entirely left mesoscale processes behind, since the same or similar phenomena crop up in later chapters, but from a different perspective. A field which now seems ripe for more research is to proceed beyond simple analysis and basic

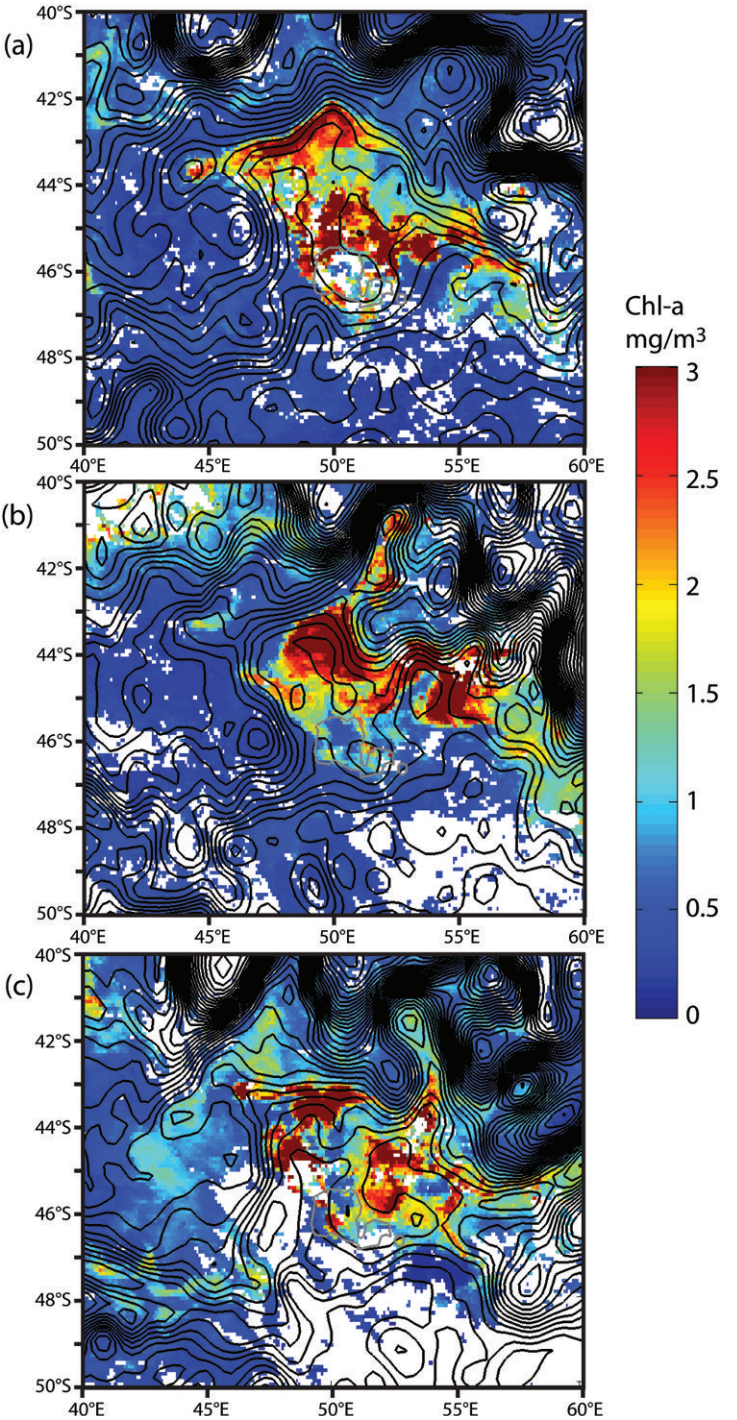


Figure 5.22. Contours of the ADT field, equivalent to streamlines of the surface absolute velocity field, overplotted onto the satellite-derived chlorophyll-*a* field surrounding the Crozet Plateau, for the weeks of (a) October 24–31 , 2002, (b) November 1–8 , 2005, and (c) October 24–31 , 2006 to show how chlorophyll varies consistently with the varying flow field from one year to another (image provided by Hugh Venables, British Antarctic Survey).

understanding of individual mesoscale processes to mapping their climatological variability. The climatology of frontal occurrence already mentioned in Chapter 4 shows that this is possible once a long time series of image data is available. We should expect to construct similar climatologies in relation to eddy statistics or the strength of upwelling The increasing availability of well-processed datasets, often using carefully specified analyses that blend data from different sensors in order to improve SST, ocean color, or SSH products, should pave the way for generating climatologies of variability. These should contribute to developing our understanding of how climate change affects the detailed behavior of the ocean at local scales.

5.8 REFERENCES

Abraham, E. R., C. S. Law, P. W. Boyd, S. J. Lavender, M. T. Maldonado, and A. R. Bowie (2000), Importance of stirring in the development of an iron-fertilized phytoplankton bloom. *Nature*, **407**, 727–730.

Aristegui, J., P. Tett, A. Hernandez-Guerra, G. Basterretxea, M. F. Montero, K. Wild, P. Sangra, S. Hernandez-León, M. Canton, J. A. Garcia-Braun, and E. D. Barton (1997), The influence of island-generated eddies on chlorophyll distribution: A study of mesoscale variation around Gran Canaria. *Deep-Sea Res. I*, **44**(1), 71–96.

Ballastero, D., and E. Coen (2004), Generation and propagation of anticyclonic rings in the Gulf of Papagayo, Costa Rica. *Int. J. Remote Sensing*, **25**(11), 2217–2224.

Barth, J. A., and P. A. Wheeler (2005), Introduction to special section: Coastal advances in shelf transport. *J. Geophys. Res.*, **110**(C10S01), doi: 10.1029/2005JC003124.

Barth, J. A., S. D. Pierce, and R. L. Smith (2000), A separating coastal upwelling jet at Cape Blanco, Oregon and its connection to the California Current System. *Deep-Sea Res. II*, **47**, 783–810.

Barth, J. A., S. D. Pierce, and R. M. Castelao (2005), Time-dependent, wind-driven flow over a shallow midshelf submarine bank. *J. Geophys. Res.*, **110**(C10S05), doi: 10.1029/2004JC002761.

Barton, E. D. (2001), Turbulence and diffusion: Island wakes. In: J. Steele, S. A. Thorpe, and K. Turekian (Eds.), *Encyclopedia of Ocean Sciences* (pp. 1397–1403). Academic Press, London.

Barton, E. D., P. J. Flament, H. Dodds, and E. G. Mitchelson-Jacob (2001), Mesoscale structure viewed by SAR and AVHRR near the Canary Islands. *Scientia Marina*, **65**(Suppl. 1), 167–175.

Barton, I. J. (2002) Ocean currents from successive satellite images: The reciprocal filtering technique. *J. Atm. Ocean. Tech.*, **19**(10), 1677–1689.

Bischof, B., A. J. Mariano, and E. H. Ryan (2004), *The North Equatorial Counter Current: Ocean Surface Currents*, available at *http://oceancurrents.rsmas.miami.edu/atlantic/north-equatorial-cc.html* (last accessed July 20, 2008).

Bouffard, J., S. Vignudelli, P. Cipollini, and Y. Menard (2008), Exploiting the potential of an improved multimission altimetric data set over the coastal ocean. *Geophys. Res. Letters*, **35**(L10601), doi: 10.1029/2008GL033488.

Bowen, M., W. J. Emery, J. L. Wilkin, P. Tildesley, I. J. Barton, and R. Knewtson (2002). Extracting multi-year surface currents from sequential thermal imagery using the Maximum Cross Correlation technique. *J. Atm. Ocean. Tech.*, **19**, 1665–1676.

Boyd, P. W., T. Jickells, C. S. Law, S. Blain, E. A. Boyle, K. O. Buesseler, K. H. Coele, J. J. Cullen, H. J. W. de Baar, M. Follows *et al.* (2007), Mesoscale iron enrichment experiments 1993–2005: Synthesis and future directions. *Science*, **315**, 612–617.

Buesseler, K. O., R. T. Barber, M.-L. Dickson, M. R. Hiscock, J. K. Moore, and R. Sambrotto (2003), The effect of marginal ice-edge dynamics on production and export in the Southern Ocean along 170°W. *Deep-Sea Res. II*, **50**, 579–603.

Caldeira, R. M. A., S. B. Groom, P. I. Miller, and N. Nezlin (2002), Sea-surface signatures of the island mass effect phenomena around Madeira Island, Northeast Atlantic. *Remote Sens. Environ.*, **80**, 336–360.

Castelao, R. M., and J. A. Barth (2006), Upwelling around Cabo Frio, Brazil: The importance of wind stress curl. *Geophys. Res. Letters*, **33**(L03602), doi: 10.1029/2005GL025182.

Castelao, R. M., T. P. Mavor, J. A. Barth, and L. C. Breaker (2006), Sea surface temperature fronts in the California Current System from geostationary satellite observations. *J. Geophys. Res.*, **111**(C09026), doi: 10.1029/2006JC003541.

Chelton, D. B., M. Freilich, and S. K. Esbensen (2000), Satellite observations of the wind jets off the Pacific Coast of Central America: Part I, Case studies and statistical Characteristics. *Monthly Weather Review*, **128**(7), 1993–2018.

Coale, K. H., K. S. Johnson, F. P. Chavez, K. O. Buesseler, R. T. Barber, M. A. Brzezinski, W. P. Cochlan, F. J. Millero, P. G. Falkowski, J. E. Bauer, and 40 others (2004), Southern Ocean Iron Enrichment Experiment: Carbon cycling in high- and low-Si waters, *Science*, **304**(5669), 408–414.

Comiso, J. C., N. G. Maynard, W. O. Smith, and C. W. Sullivan (1990), Satellite ocean color studies of Antarctic ice-edges in Summer and Autumn. *J. Geophys. Res.*, **95**(C6), 9481–9496.

Cooley, S. R., V. J. Coles, A. Subramaniam, and P. L. Yager (2007), Seasonal variations in the Amazon plume-related atmospheric carbon sink. *Glob. Biogeochem. Cycles*, **21**(GB3014), doi: 10.1029/2006GB002831.

Dietrich, D. E., M. J. Bowman, C. A. Lion, and A. Mestas-Nunez (1996), Numerical studies of small island wakes in the ocean. *Geophysical and Astrophysical Fluid Dynamics*, **83**(3/4), 195–231.

Domingues, C. M., G. A. Goncalves, R. D. Ghisolfi, and C. A. E. Garcia (2000), Advective surface velocities derived from sequential infrared images in the southwestern Atlantic Ocean. *Remote Sens. Environ.*, **73**, 218–226.

Dong, C., J. C. McWilliams, and A. Shchepetkin (2007), Island wakes in deep water. *J. Phys. Oceanogr.*, **37**, 962–981.

Emery, W. J., D. Baldwin, and D. K. Matthews (2003), Maximum cross correlation automatic satellite image navigation and attitude corrections for open ocean image navigation. *IEEE Trans. Geosc. Remote Sensing.*, **41**, 33–42.

Engelsen, O., H. Hop, E. N. Hegseth, E. Hansen, and S. Falk-Petersen (2004), Deriving phytoplankton biomass in the Marginal Ice Zone from satellite observable parameters. *Int. J. Remote Sensing*, **25**(7/8), 1453–1457.

ERD (2008), *ERD Coastal Upwelling Indices Web Page*. NOAA, available at *www.pfeg.noaa.gov/products/PFEL/modeled/indices/upwelling/NA/* (last accessed July 16, 2008).

Gabric, A. J., J. M. Shephard, J. M. Knight, G. Jones, and A. J. Travena (2005), Correlations between the satellite-derived seasonal cycles of phytoplankton biomass and aerosol optical depth in the Southern Ocean: Evidence for the influence of sea ice. *Glob. Biogeochem. Cycles*, **19**(GB4018), doi: 10.1029/2005GB002546.

Heide-Jørgensen, M. P., K. L. Laidre, M. L. Logsdon, and T. G. Nielsen (2007), Springtime coupling between chlorophyll a, sea ice, and sea surface temperature in Disko Bay, West Greenland. *Prog. Oceanogr.*, **73**, 79–95.

Hu, C., E. T. Montgomery, R. W. Schmitt, and F. E. Muller-Karger (2004), The dispersal of the Amazon and Orinoco River water in the tropical Atlantic and Caribbean Sea: Observation from space and S-PALACE floats. *Deep-Sea Res. II*, **51**, 1151–1171.

Lathuilière, C., V. Echevin, and M. Lévy (2008), Seasonal and intraseasonal surface chlorophyll-a variability along the northwest African coast. *J. Geophys. Res.*, **113**(C05007), doi: 10.1029/2007JC004433.

Legeckis, R. (1988) Upwelling off the Gulfs of Panama and Papagayo in the tropical Pacific during March 1985. *J. Geophys. Res.*, **93**(C12), 15485–15489.

Lubin, D., and R. Massom (2006), *Polar Remote Sensing: Atmosphere and Oceans* (xlii + 756 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Madsen, K. S., J. L. Høyer, and C. C. Tscherning (2007), Near-coastal satellite altimetry: Sea surface height variability in the North Sea–Baltic Sea area. *Geophys. Res. Letters*, **34**(L14601), doi: 10.1029/2007GL029965.

Marin, V. H., L. E. Delgado, and R. Escribano (2003), Upwelling shadows at Mejillones Bay (northern Chilean coast): A remote sensing, in situ analysis. *Invest. Mar., Valparaiso*, **31**(2), 47–55.

Martinez Diaz de Leon, A., I. S. Robinson, D. Ballastero, and E. Coen (1999), Wind driven circulation features in the Gulf of Tehuantepec, Mexico, revealed by combined SAR and SST satellite data. *Int. J. Remote Sensing*, **20**(8), 1661–1668.

McGillicuddy, D., F. E. Muller-Karger, and P. L. Richardson (1995), On the offshore dispersal of the Amazon's Plume in the North Atlantic: Comments on the paper by A. Longhurst, "Seasonal cooling and blooming in tropical oceans". *Deep-Sea Res. I*, **42**(11), 2127–2137.

Moore, J. K., and M. R. Abbott (2000), Phytoplankton chlorophyll distributions and primary production in the Southern Ocean. *J. Geophys. Res.*, **105**, 28709–28722.

Moore, J. K., M. R. Abbott, J. G. Richman, W. O. Smith, T. J. Cowles, K. H. Coale, W. D. Gardner, and R. T. Barber (1999), SeaWiFS satellite ocean color data from the Southern Ocean. *Geophys. Res. Letters*, **26**(10), 1465–1468.

Mustapha, M. A., and S.-I. Saitoh (2008), Observations of sea ice interannual variations and spring bloom occurrences at the Japanese scallop farming area in the Okhotsk Sea using satellite imageries. *Estuarine Coast. Shelf Sci.*, **77**, 577–588.

Nash, J. D., and J. N. Mourn (2005), River plumes as a source of large-amplitude internal waves in the coastal ocean. *Nature*, **437**(September), 400–403.

Nykjaer, L., and L. Van Camp (1994), Seasonal and interannual variability of coastal upwelling along northwest Africa and Portugal from 1981 to 1991. *J. Geophys. Res*, **99**(C7), 14197–14207.

Palacios, D. M., and S. J. Bograd (2005) A census of Tehuantepec and Papagayo eddies in the northeastern tropical Pacific. *Geophys. Res. Letters*, **32**(L23602), doi: 10.1029/2005GL024324.

Pollard, R. T., R. Sanders, M. I. Lucas, and P. J. Statham (2007a), The Crozet natural iron bloom and export experiment (CROZEX). *Deep-Sea Res. II*, **54**(18/20), 1905–1914.

Pollard, R. T., H. J. Venables, J. F. Read, and J. T. Allen (2007b), Large-scale circulation around the Crozet Plateau controls annual phytoplankton bloom in Crozet Basin. *Deep-Sea Res. II*, **54**(18/20), 1915–1929.

Pradhan, Y., S. J. Lavender, N. J. Hardman-Mountford, and J. Aiken (2006), Seasonal and inter-annual variability of chlorophyll-a concentration in the Mauritanian upwelling: Observation of an anomalous event during 1998–1999. *Deep-Sea Res. II*, **53**, 1548–1559.

Rees, W. G. (2006), *Remote Sensing of Snow and Ice* (pp. 285 + xix). Taylor & Francis/CRC Press, Boca Raton, FL.

Santos, M. P., A. S. Kazmin, and A. Peliz (2005), Decadal changes in the Canary upwelling system as revealed by satellite observations: Their impact on productivity. *J. Marine Res.*, **63**, 359–379.

Sasamal, S. K. (2007), Island wake circulation off Maldives during boreal winter, as visualised with MODIS derived chlorophyll-a data and other satellite measurements. *Int. J. Remote Sensing*, **28**(5), 891–903.

Smith, W. O., and D. N. Nelson (1986), Importance of ice edge phytoplankton production in the Southern Ocean. *BioScience*, **36**(4), 251–257.

Stewart, R. H. (2008), *Introduction to Physical Oceanography*. Texas A & M University, e-book, available at *http://oceanworld.tamu.edu/home/course_book.htm*

Strub, P. T., and C. James (2000), Altimeter-derived variability of surface velocities in the California Current System, 2: Seasonal circulation and eddy statistics. *Deep-Sea Res. II*, **47**, 831–870.

Strub, P. T., P. M. Kosro, and A. Huyer (1991), The nature of the cold filaments in the California Current System. *J. Geophys. Res.*, **96**(C8), 14743–14768.

Strub, P. T., T. K. Chereskin, P. Niiler, C. James, and M. Levine (1987), Altimeter-derived variability of surface velocities in the California Current System, 1: Evaluation of TOPEX altimeter velocity resolution. *J. Geophys. Res.*, **102**(C6), 12727–12748.

Thomas, A. C., J. L. Blanco, M. E. Carr, P. T. Strub, and J. Osses (2001), Satellite-measured chlorophyll and temperature variability off northern Chile during the 1996–1998 La Niña and El Niño. *J. Geophys. Res.*, **106**(C1), 899–915.

Van Camp, L., L. Nykjaer, E. Mittelstaedt, and P. Schlittenhardt (1991), Upwelling and boundary circulation off northwest Africa as depicted by infrared and visible satellite observations. *Prog. Oceanogr.*, **26**, 357–402.

Venables, H. J., and C. M. Moore (2010), Phytoplankton and light limitation in the Southern Ocean: Learning from high nutrient high chlorophyll areas, *J. Geophys. Res.*, **115**, C02015, doi: 10.1029/2009JC005361.

Venables, H. J., R. T. Pollard, and E. Popova (2007), Physical conditions controlling the development of a regular phytoplankton bloom north of the Crozet Plateau, Southern Ocean. *Deep-Sea Res. II*, **54**(18–20), 1949–1965.

Walker, N. D. (1996). Satellite assessment of Mississippi River plume variability: Causes and predictability, *Remote Sens. Environ.*, **58**, 21–35.

Wu, Y., I. K. Peterson, C. L. Tang, T. Platt, S. Sathyendranath, and C. Fuentes-Yaco (2007), The impact of sea ice on the initiation of the spring bloom on the Newfoundland and Labrador Shelves. *J. Plankton Res.*, **29**(6), 509–514.

Xie, S. P., W. T. Liu, Q. Liu, and M. Nonaka (2001), Far-reaching effects of the Hawaiian Islands on the Pacific Ocean–Atmosphere System. *Science*, **292**(5524), 2057–2060.

Zamudio, L., H. E. Hurlburt, E. J. Metzger, S. L. Morey, J. J. O'Brien, C. Tilburg, and J. Zavala-Hidago (2006), Interannual variability of Tehuantepec eddies. *J. Geophys. Res.*, **111**(C05001), doi: 10.1029/2005JC003182.

6

Planetary waves and large-scale ocean dynamics

6.1 PHENOMENA SEEN BEST FROM SATELLITES

Every field of natural science is constrained by the capacity of experimental observations to test the hypotheses and theoretical models on which understanding of the subject is based. The science of oceanography is more than a century old and throughout that time its pioneers sought to match the scope of their research to the planetary scale of the world's ocean. Yet the scarcity of observations made on a truly global scale and repeated over correspondingly long timescales hindered the development of global ocean dynamics towards that level of scientific maturity where theory could be properly tested by observation and experiment. The ready availability of satellite data has changed that situation, particularly in relation to a certain class of large-scale ocean processes.

Most of the new phenomena discovered by 20th-century oceanography, and the dynamical theories to explain them, were limited to local and regional horizons. Only in the last two decades of the century was this weakness remedied as the global datasets carefully compiled from satellites started to provide a firmer basis for testing the scientific theories of large-scale ocean wave dynamics. These are phenomena of large extent whose propagation is constrained by the rotation and the spherical form of the planet. This is the reason for their generic name, planetary waves. Their lengthscales and timescales are large enough that they respond to the way in which the Coriolis parameter (associated with Earth rotation) varies with latitude. They merit a special chapter because our modern understanding of them would have been impossible without the advent of satellite oceanographic methods.

This chapter looks first (in Section 6.2), at the ways in which global satellite datasets need to be prepared in order to reveal planetary wave phenomena, and then reveals the evidence for large-scale propagating features, found in global datasets of several independent variables representing very different ocean properties. In Section

6.3 we recall the basic theoretical concepts of Rossby wave[1] propagation, while Section 6.4 presents the data analysis technique used to measure the speed of waves. Section 6.5 considers how confronting well-established theory with new observations has led to better scientific understanding and appreciation of the potential importance of these phenomena, has prompted the need to improve the theory of Rossby waves, and has stimulated interest in the wider importance of this class of ocean dynamics which can only be appreciated fully from the vantage point of a sensor on an Earth-orbiting satellite. In this we find a clear example of the scientific method at work, further confirmation that remote-sensing methods have an essential role in making the observations needed to test fluid-dynamical theories on the planetary scale.

Finally Section 6.6 takes a brief look at some other large-scale ocean-dynamical phenomena which, like Rossby waves, benefit from application of satellite data processed to reveal their propagation characteristics. These are equatorial Kelvin waves, tropical instability waves, the Madden–Julian Oscillation, and the Antarctic circumpolar waves.

6.2 DETECTING PLANETARY WAVES FROM SPACE

In order to observe planetary wave motions, which occupy characteristic lengthscales of hundreds to several thousands of kilometers and timescales of weeks to several years, it is necessary to adopt a different approach to analysis of satellite data than that presented in the previous three chapters on mesoscale processes. Even when an ocean variable is mapped from satellites on a global or basin scale, the phenomena we are searching for may not immediately be apparent in single "snapshots" of that variable, or even in a time series of such snapshots presented in a movie loop. It may be that the magnitude of mesoscale processes, seasonal variability, or natural geographical variations are much greater than the amplitude of planetary waves, which can then be revealed only by looking for coherence over suitably long timescales. Even the random noise of the primary satellite measurement method may obscure large-scale waves. To overcome this, global datasets need to be prepared carefully. First, the signature, sometimes dominant, of high-frequency and seasonal variability can be suppressed by producing composites from several overpasses and then evaluating the anomaly in relation to the climatological mean (as explained in Section 6.2.1). Next, the technique of preparing longitude–time plots of data is described in Section 6.2.2. Experimental results when these methods are applied to sea surface height, sea surface temperature, and ocean color data are then presented in Sections 6.2.3 to 6.2.5.

[1] Throughout this chapter, no distinction is intended between the terms "planetary wave" and "Rossby wave". In general the former term is used when referring to the actual phenomenon as observed. The latter is used when describing either the theoretical concept of planetary wave motion, which was initially defined by Carl-Gustav Rossby, or later amendments of the theoretical wave model.

6.2.1 Producing composite anomaly datasets

It should be evident already from the preliminary discussion in this chapter that our attention is firmly fixed on large-scale processes, phenomena that may span thousands of kilometers. In order to observe these a sampling resolution as coarse as tens of kilometers and several days is not only adequate but preferable, as long as any variability that is detected at shorter lengthscales and timescales by individual overpasses of sensors has been smoothed out. Level 3 composite datasets fulfill this role.

As outlined in Section 2.3.6, to create composite datasets the measurements of an ocean parameter are first retrieved as level 2 products at the full resolution of the sensor. The measured data may be in the form of a two-dimensional image sampled over a grid aligned locally along and across the ground track of the satellite (e.g., an infrared or ocean color radiometer) or regular point samples along the satellite ground track (e.g., an altimeter). These data should be processed as usual to remove atmospheric effects, detect cloud, and estimate the appropriate ocean variable (as discussed in Chapter 2).

The space-time characteristics of the composite are defined by specifying the cells of a regular grid larger than the sensor's spatial sampling interval (as illustrated in Figure 6.1). To facilitate the detection of planetary waves, the new grid should be defined in regular geographical co-ordinates (latitude and longitude) with spacing of

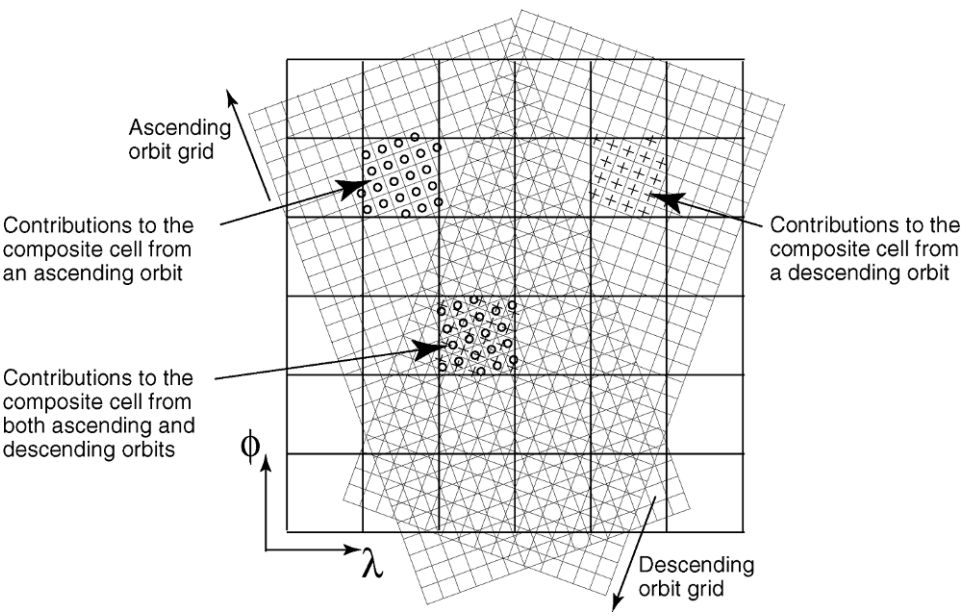


Figure 6.1. Schematic illustrating how values of an ocean variable, sampled at full resolution on a level 2 grid in sensor co-ordinates, are allocated to the corresponding level 3 geographical grid. Values accumulated from several overpasses are averaged to produce a composite dataset.

at least 1/12° latitude. Samples are also accumulated from each overpass within a given time window that is typically significantly longer than the sensor revisit interval for level 2 data. For example, SST data from a sensor like AVHRR with a 12-hour revisit interval, but subject to data loss from cloud cover, may be formed into a level 3 composite product using a time-sampling window of several days.

The composite is typically formed from all the level 2 data accumulated within each space-time cell. However, now that some level 2 data products come with reliable quality values that provide error estimates for each individual pixel (as explained in Section 14.4.2) there may be an opportunity to select objectively only the most reliable level 2 observations to produce the composite. The selected data are then averaged to a single value. This reduces the random variability arising from instrument noise and from unbiased errors in atmospheric correction or other processing. The longer the averaging time interval the larger the number of samples and, in principle, the lower the noise in the final composite time series. This is especially so for sensors affected by cloud cover. Moreover, unlike the production of composites for mesoscale applications, in this case there is a benefit from averaging out the true ocean variability at lengthscales shorter than a few tens of kilometers and timescales less than a few days, since these signals may have a greater magnitude than the low-frequency planetary waves being sought.

Ultimately, the preprocessing of composites should be optimized to match the spectral characteristics (i.e., the dominant length and timescales) of the phenomenon to be studied. Note that this does not simply mean that the sampling interval can be infrequent. If the ocean signal contains high variability then it needs to be sampled as completely as possible in order to obtain a true average without aliasing, which could otherwise introduce a bias. Other factors may have a biasing effect and must be dealt with intelligently. The details of how to approach this will vary depending on the variable (as discussed later in this section).

In practice the time series of composite images is generally not optimal for clearly revealing planetary waves. This is because it still contains large-scale signals describing the geographical distribution of surface properties of the ocean and their regular annual (seasonal) variability, which may have an amplitude that is considerably larger than the perturbations associated with planetary waves, and which would tend to dominate a viewer's perception of what the data reveal. Thus we need to present observed ocean properties as perturbations from their expected climatological value at the given location and time within the seasonal cycle. This is referred to as the anomaly value.

The regular monitoring of ocean properties from space and their assembly into a time series of composite images on a regular grid provides an ideal basis for generating the corresponding anomaly images. Figure 6.2 illustrates schematically the way this is achieved. A timespan of several years of composited images is needed. It does not matter what the time interval is for the composites, as long as it is repeated on an annual cycle. Thus, for example, composites could be integrated every month, every week, or every 4 days. In the latter cases, which do not divide exactly into one year, there must be some adjustment made to ensure that the annual cycle repeats precisely. Thus in the case of weekly composites, the 52nd "week" of each year will

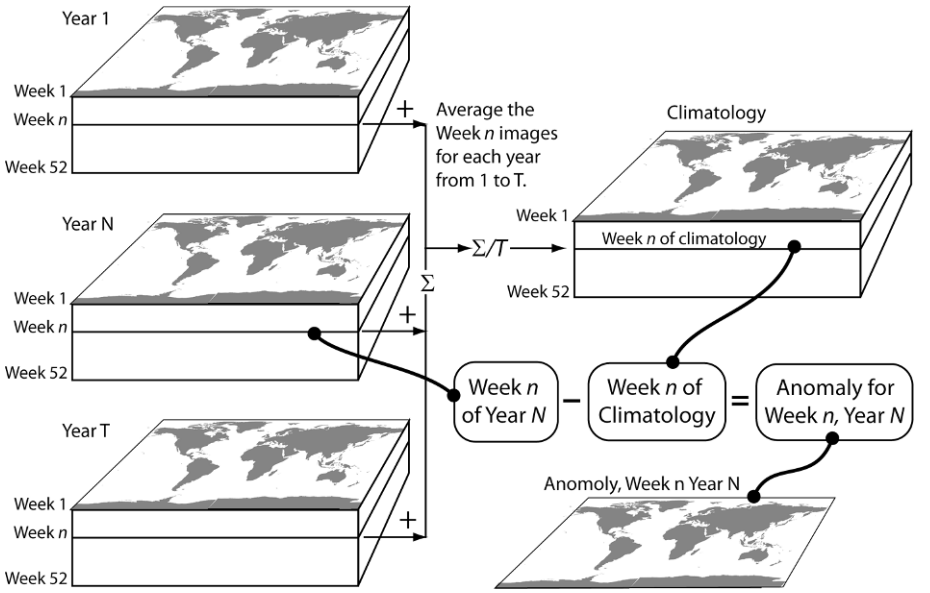


Figure 6.2. Schematic for creating a climatology and anomalies based on weekly composites. Starting with 52 weekly composite images spanning T years, extract the nth weekly composite from each of T years and average them to produce the climatology for week n. Repeat for each of the 52 weeks to produce 52 weekly climatology images. To produce the anomaly image for Week n, Year N, subtract the T-year climatology for Week n from the image for Week n, Year N.

consist of 8 days (9 days for a leap year) so that the nth composite image of every year always occupies exactly the same day numbers every year.

The images corresponding to a particular stage in the annual cycle (e.g., every nth week image from every year in the full span of data) are then averaged to produce the climatological distribution of the ocean property for the nth week. When this is repeated for all weeks (or months or other time interval) of the year, the result is a time series of images spanning 1 year and showing how, averaging over several years, the ocean property distribution varies during the year. This is annual climatology.

The anomaly image associated with the composite image for a given time period of a particular year is produced simply by subtracting from the composite image the climatological image for the same time period. The resulting time series of anomaly images therefore shows how an ocean property evolves in space and time, but with the underlying geographical distribution and its typical seasonal variations removed. The time series of anomaly maps for a variable which never departs from the same seasonal cycle would always be zero. Thus even a small departure from the norm stands out clearly in the anomaly.

Several points are worth noting in relation to anomalies. First, although it is easier to simply subtract a long-term time average without attempting to create an

annual climatology, this removes only the mean geographical distribution; the resulting anomalies still contain basic seasonal variations over an annual cycle. In the case of sea surface height, these are small, but for temperature and color they are likely to dominate the time series and tend to hide any nonannual signals present. Figure 6.3 shows samples at four times throughout the year from the global SST climatology produced by the NOAA National Center for Environmental Prediction (NCEP) combining data from several sources using optimal interpolation (Reynolds *et al*., 2002). These are actually climatological monthly means, although climatology is also produced at finer time resolution than this. There are quite large changes in global SST patterns between different seasons, as highlighted in the lower panel of Figure 6.3 where temperature transects along 170°W are shown for the four different months. Differences of up to 10°C between March and September are found at midlatitudes.

Second, the number of years over which a climatology is generated needs to be sufficient that a strong anomaly in a single year does not overly influence the mean. The length of the time series is a problem when a new dataset is being produced from a new sensor system, but a 10-year data span is generally long enough for generating anomalies that reveal planetary waves. This issue is discussed again in Chapter 11 in relation to the detection of climate anomalies like El Niños.

Third, when the annual climatology has already been established from a long span of continuous observations, then anomaly maps can be produced at the same time as the composites, as a near real–time product of the observing system. This capacity to generate anomaly maps is one of the benefits of satellite remote sensing that is often taken for granted, but which was unavailable to a former generation of oceanographers. When samples are infrequent and irregularly spaced it becomes difficult, if not impossible, to distinguish between measurement errors, natural seasonal variations, individual ocean phenomena localized in space and time, and planetary waves with a coherent propagation history. This is one reason why planetary waves were not unambiguously observed for several decades after their occurrence had been predicted theoretically.

6.2.2 Producing Hovmöller diagrams to reveal propagating features

Planetary waves do not have a particular spatial structure that makes them stand out clearly in an anomaly map of a particular ocean parameter. What characterizes them is the way they propagate with a given speed and direction. Thus if we are to find evidence for them in the time series of anomaly maps, produced as described in the previous subsection, we need to look at the time evolution of anomalies more than at individual snapshots. One way of doing this is to generate movie sequences of maps, which rely on the subjective interpretation of our brain–eye combination to pick out coherent moving structures. However, if we want a more objective and analytical approach to determining whether planetary wave features exist, the use of Hovmöller diagrams is appropriate.

A Hovmöller diagram is the name given to a plot which maps the variation of an ocean property as a function of longitude and time, at a fixed latitude, in contrast

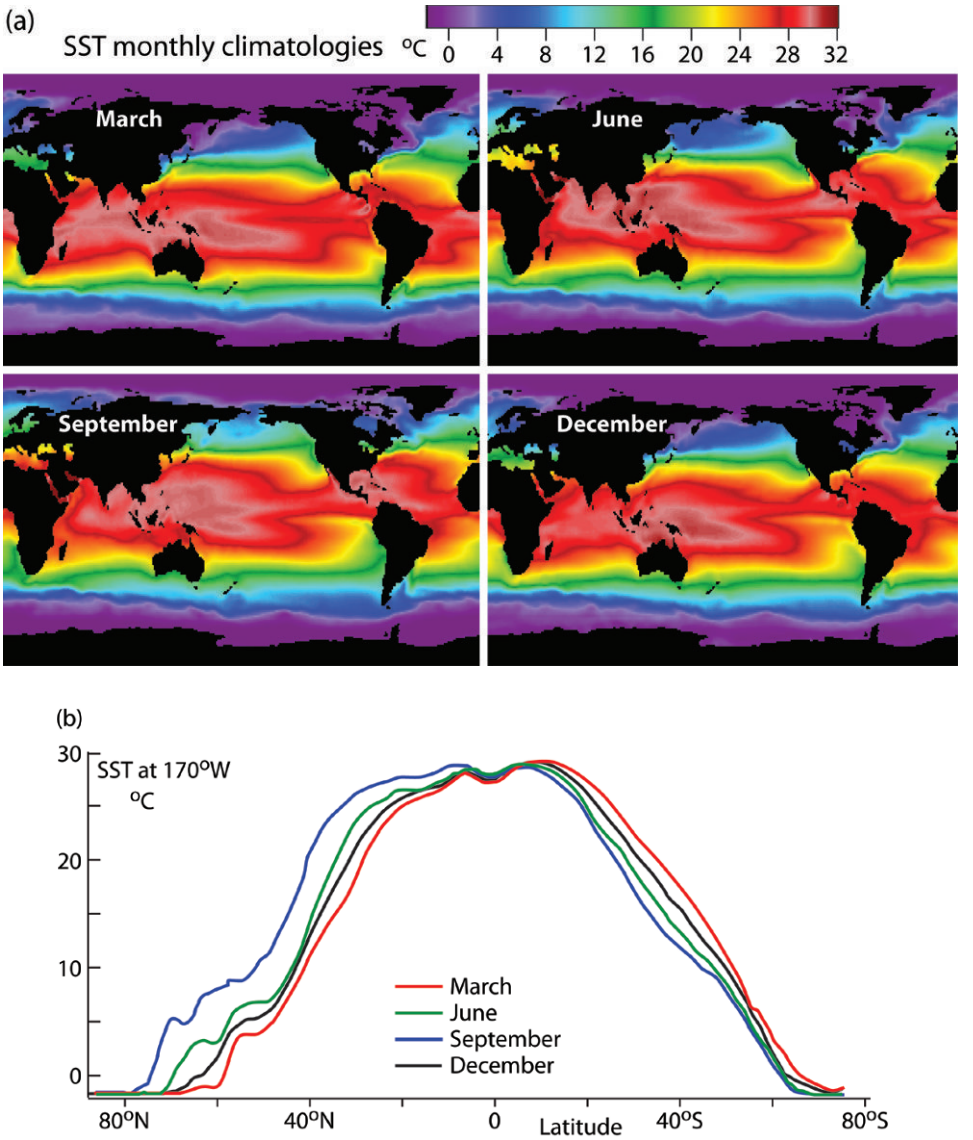


Figure 6.3. NCEP monthly SST climatology, sampled at March, June, September, and December. (a) Maps of global climatology. (b) Temperature distribution along 170°W.

with a normal map that plots the property as a function of longitude and latitude at a fixed time. If we consider the time series of anomaly data as a three-dimensional cube of data, in which two-dimensional geographical maps from successive time intervals are stacked up vertically (as shown in Figure 6.4), then the Hovmöller plot is derived as a vertical slice through the cube, cut along a particular line of

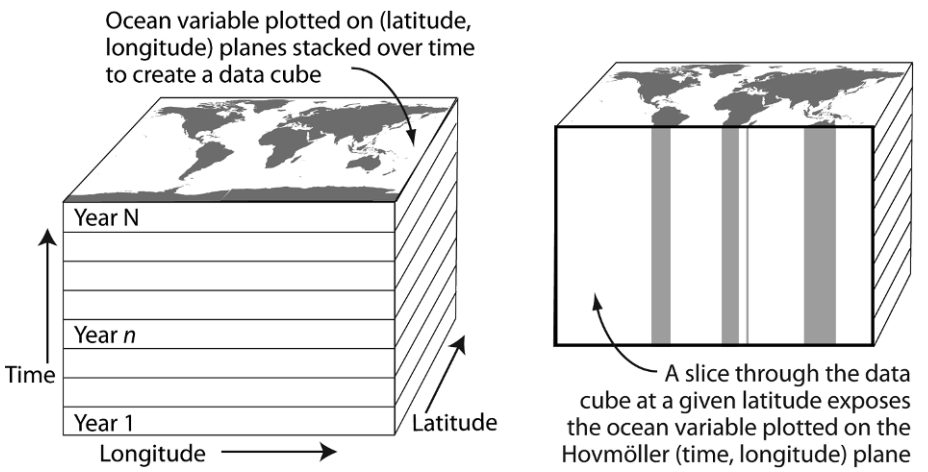


Figure 6.4. Schematic of the "data cube" produced by vertically stacking successive two-dimensional maps of a satellite data time series. A vertical cut through the cube at a given latitude exposes a "face" which is a map of how the ocean property varies with longitude and time at the chosen latitude.

latitude. The pixel size in the time dimension is determined by the integration time for composite maps from which the data cube is constructed.

It is also possible to take a slice along a line of longitude to produce a two-dimensional plot of how the ocean property is distributed with latitude and time. Indeed a cut into the data cube can be made along any straight or curved line drawn on the geographical map, producing a section which displays how the property varies with time along that line. The reason why planetary waves are normally sought on longitude–time plots is that they are expected to propagate in a direction parallel to lines of latitude.

6.2.3 Altimetry reveals the first compelling evidence of planetary waves

It is in the observations of the sea surface height anomaly (SSHA) from radar altimetry that planetary waves have been most clearly observed. While there were several reports of Rossby wave–like signals in the data from Geosat in the late 1980s (White *et al.*, 1990; Jacobs *et al.*, 1993; Le Traon and Minster, 1993), the possibility of tidal aliasing in the Geosat data record, and the marginal height resolution of Geosat in relation to the planetary wave surface amplitude of a few centimeters, left some room for doubt. The first compelling evidence for the ubiquitous existence of planetary waves came when a sufficient length of TOPEX/Poseidon record had been acquired. The complete elimination of tidal aliasing (see sections 11.3.2 and 11.4.3 of *MTOFS*) meant that there could be no spurious transfer of tidal energy into the long periods ($\sim$1 year) characteristic of planetary waves, while the height resolution of better than 3 cm left no room for doubt that what was being detected in zonal

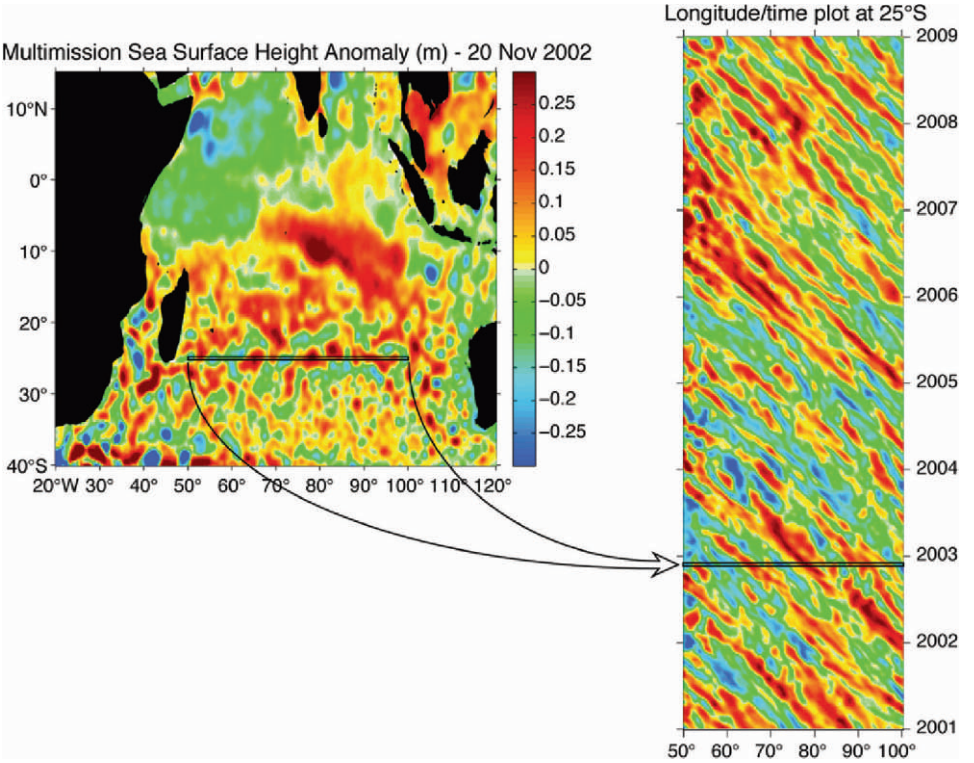


Figure 6.5. *Left panel*: map of sea surface height anomaly (SSHA) in the Indian Ocean derived from one 10-day cycle of the TOPEX/Poseidon altimeter on November 20, 2002. The schematic shows how a row of pixels from the image at latitude 25°S between 50°E and 100°E is extracted to form a row in the Hovmöller plot. *Right panel*: diagonal lines provide the evidence of planetary waves (created by P. Cipollini using altimeter data products produced by SSALTO/DUACS and distributed by Aviso with support from CNES).

sections across many parts of the ocean were indeed planetary waves (Chelton and Schlax, 1996).

Figure 6.5 shows how an example of a Hovmöller plot is constructed from a succession of SSHA images of the Indian Ocean. The evidence of planetary waves is to be found in the diagonally sloping patterns on the Hovmöller plot. What these show is that perturbations of SSHA, with an amplitude of order 20 cm above and below the mean level, gradually propagate westwards with time, taking nearly 2 years to travel about 3,000 km. The horizontal scale of these features is between 5° and 10° in longitude (500–1,000 km), and an observer at a fixed latitude would estimate that they have a characteristic timescale (from peak to peak) of about 6 months. These make up the dominant signals in the plot, although there are other perturbations present, corresponding to more random variations of sea surface height.

The factors that imply evidence of planetary waves are the dominance of a westward-propagating signal revealed as a diagonal line, the slope of which corresponds to a speed of propagation that approximately matches the expected speed of Rossby waves (as Sections 6.3 and 6.4 will demonstrate), and the coherence of perturbations along the diagonals. Note that there are no equivalent eastward-trending patterns which might be expected if the diagonal features were simply an artifact of the Hovmöller construction process. For those not already familiar with the theory of Rossby waves, a plot such as Figure 6.5 provides valuable experimental information about the behavior of the phenomenon. Although there are dominant lengthscales and time scales to be found by examining horizontal or vertical transects across the plot, these do not produce a regular set of wave motions, implying that there is not a regular periodic source of the wave-like energy. On the other hand, the strongest coherence of the pattern is found along the diagonals, which implies that once a pulse of energy is fed by some means into a perturbation of SSHA, it propagates westwards in a tightly constrained way at a given speed. These are useful insights which demand a satisfactory explanation by Rossby wave theory if it is to be confirmed as the explanation for the observed phenomenon.

6.2.4 Sea surface temperature signatures

In addition to SSHA data, other satellite-derived ocean datasets have also clearly demonstrated Rossby wave–like behavior when presented in Hovmöller plots. The first of these is sea surface temperature (SST). Initially Cipollini *et al.* (1997) observed similar diagonal patterns in Hovmöller plots of SST in the same location (34°N in the northeast Atlantic) as a strong planetary wave signature was found in the altimeter record. Then a global study (Hill *et al.*, 2000) using the SST monthly anomalies derived from the first 4 years of data from the along-track scanning radiometer (ATSR) demonstrated that planetary wave–like behavior is ubiquitous at tropical to midlatitudes in this high-quality, low-noise SST dataset composited in bins of 0.5° latitude × 0.5° longitude × 1 month.

Figure 6.6 shows an example of the type of planetary wave signatures they found. While this is noisier than the altimeter record, and displays a lot of patchiness, nonetheless the strong parallel stripes with an amplitude of about 0.5 K show up in most parts of the image. The patches of higher or lower temperatures represent localized, stationary departures from SST climatology which are probably caused by variable air–sea interaction processes, such as wind speed or cloud cover affecting insolation and mixed layer depths which then perturb SST from its normal annual cycle. Nonetheless, planetary wave propagation appears to be unaffected by these events and, although linear signatures sometimes disappear when a larger static SST anomaly occurs, they reappear again when the anomaly has subsided.

It may be that a more thorough screening of the SST record will be able to reduce patchy noise in the anomaly record. For example, if diurnal warming is allowed to contribute to the SST composite, then a season of lower-than-average winds leading to stronger diurnal warming would manifest itself as a warm zone in SST anomaly Hovmöller plots. This could be eliminated by selecting only nighttime

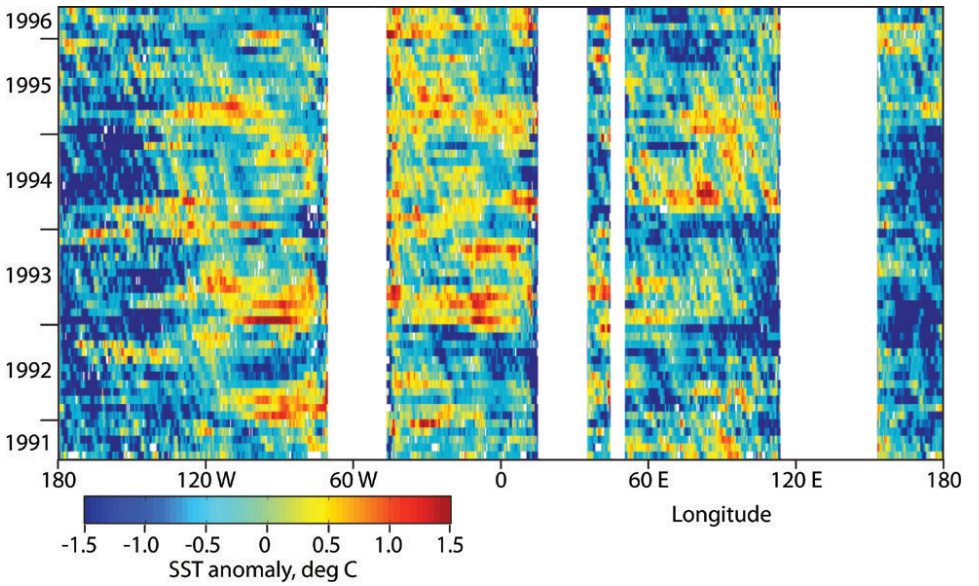


Figure 6.6. Hovmöller plot at 25°S of the SST anomaly derived from the ATSR. Note the diagonal striping that dominates some parts of the image and is evidence of westward planetary wave propagation. The white vertical bands correspond to landmasses in this plot that span the world from 180°W to 180°E. (based on Hill *et al.*, 2000).

data. Alternatively a rejection threshold in the cloud detection procedure that is not strict enough could lead to a slightly cool bias that emerges as a negative anomaly in Hovmöller images.

It should be noted that Hovmöller plots derived from the basic SST record contain seasonal and geographic signatures many times larger than the features in Figure 6.6 and these almost completely obscure planetary wave signatures. Because wave amplitude in the SST record is typically no more than about 0.5 K, it is essential that the SST is consistently measured to an accuracy that is much better than this. The long-term stability and low noise of the ATSR series of sensors has proved to be the most suitable measuring system for this purpose, although strong planetary wave signatures can also be detected in SST anomaly datasets derived from the AVHRR record. Sections 7.5.1 and 7.5.3 of *MTOFS* (Robinson, 2004) discuss in more detail the way in which procedures used to derive SST from satellite infrared radiometry can enhance or suppress unwanted artifacts in the SST anomaly record.

Planetary wave signatures have also been found in SST anomaly fields derived from microwave radiometry (Quartly *et al.*, 2003). Although the radiometric resolution of microwave sensors is presently poorer than the best infrared sensors, their capacity to measure through clouds can significantly reduce the sampling noise that occurs when producing composite images from infrared sensors. The similarity between SST anomaly signatures of planetary waves in the different Hovmöller

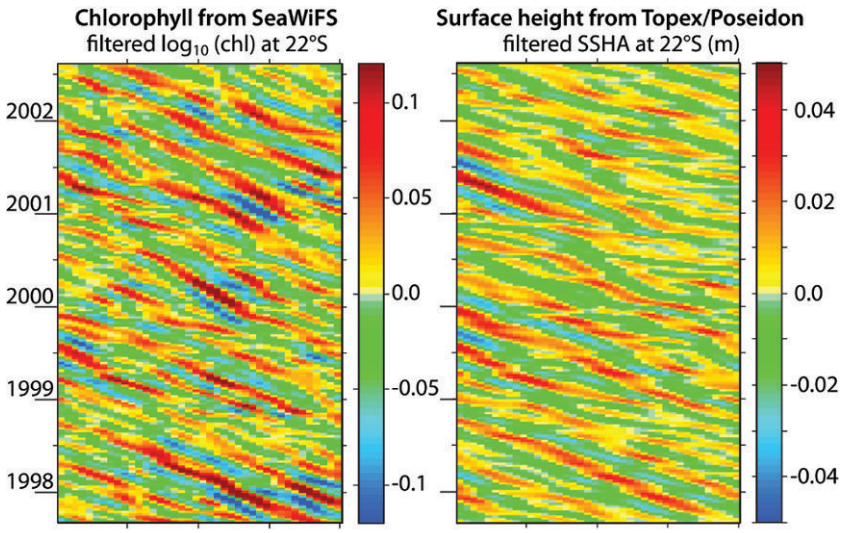


Figure 6.7. Hovmöller plots, between 130°W and 110°W and from 1997 to 2002, of $\log_{10}$(Chlor) on the *left* and SSHA on the *right* (courtesy P. Cipollini, adapted from figures previously published as figures 5a, b in Killworth *et al.*, 2004).

plots derived from infrared and microwave radiometers gives confidence that these are not merely artifacts of SST processing methods (Challenor *et al.*, 2004).

6.2.5 Evidence of planetary waves in ocean color

The other type of satellite-derived ocean data that has been found to manifest Rossby wave–like signatures is ocean color (Cipollini *et al.*, 2001). This might seem surprising because, unlike SSHA, neither the color of the sea nor the concentration of chlorophyll has any direct influence on the dynamics of the ocean. Nonetheless, the evidence of planetary wave signatures is plainly revealed in Hovmöller plots of log(chlorophyll) derived from SeaWiFS (as shown in Figure 6.7) from 22°S in the East Pacific, where comparison with the corresponding SSHA suggests that the same phenomena are responsible for both sets of signatures. Some of the strongest features are almost identical, although there are intriguing differences of detail. The possible ways in which planetary waves can affect the satellite-derived, near-surface chlorophyll record are discussed in Section 6.3.

6.3 THE CHARACTERISTICS OF ROSSBY WAVES

6.3.1 A summary of planetary wave theory

The basic mechanism that constrains Rossby waves is fairly simple. Parcels of water that are displaced northwards (or southwards) acquire negative (positive) vorticity in

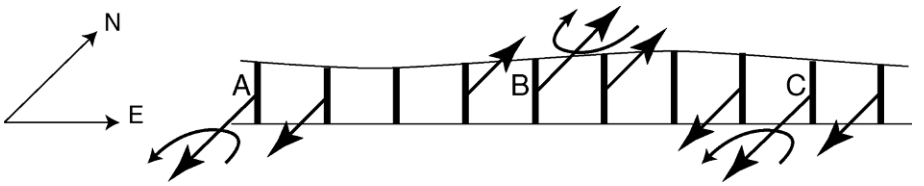


Figure 6.8. Schematic illustrating the torque experienced by water columns moving south or north at a tropical north latitude.

order to balance the gain (loss) of planetary vorticity and thus satisfy the conservation of potential vorticity which is required in the absence of any shear forces. Figure 6.8 illustrates this in a barotropic case (ignoring density stratification). We consider a situation where the flow is meridional, and varies sinusoidally with longitude. Where a patch of water is flowing northwards, as at B, the clockwise torque resulting from vorticity conservation tends to accelerate its northward velocity at the west of the patch and decelerate it at the east. For southward-flowing patches of water at A and C, the counterclockwise torque tends to accelerate the southward flow at the west and decelerate it at the east. Consequently, the positions of both the maximum northward and southward flow tend to drift towards the west.

Although this is an oversimplification of the complete dynamics represented in the full theoretical solution of planetary wave propagation, it shows clearly the asymmetry of the process which allows the wave-like distribution of velocity to propagate westwards, but not eastwards. The figure also indicates how surface height would vary with longitude to allow pressure gradients to balance the Coriolis force. It is drawn for the northern hemisphere. However, in the southern hemisphere, although surface slopes would be reversed, the "handedness" of the process which requires the pattern of velocity distribution to drift westwards does not change. This type of wave depends not on the Coriolis parameter itself but on its rate of change with latitude, which is always the same sign and is a maximum at the Equator.

In fact barotropic planetary waves propagate too rapidly to be evident: the phenomenon of interest in this chapter, which gives rise to the slow westward propagation of perturbations revealed in Figures 6.5 and 6.6 is that of baroclinic planetary waves. In baroclinic waves, motion varies with depth and internal pressure forces which maintain geostrophic balance are controlled not only by surface height but also by the varying height of density layers in the ocean. The first baroclinic mode can be thought of as the motion of a two-layer ocean, one layer above the thermocline and another, slightly denser layer below it (as shown in Figure 6.9). In this mode, flow is in opposite directions above and below the thermocline. The perturbations of thermocline height are much greater than sea surface height, and are opposite in sign. Thus pressure gradients in the lower layer, and corresponding geostrophic velocities, are very much smaller than those in the upper layer. In second-mode planetary waves, horizontal velocity in a column of water changes sign twice with depth and there are effectively three layers of flow. Higher modes have increasingly complex vertical dynamical structures.

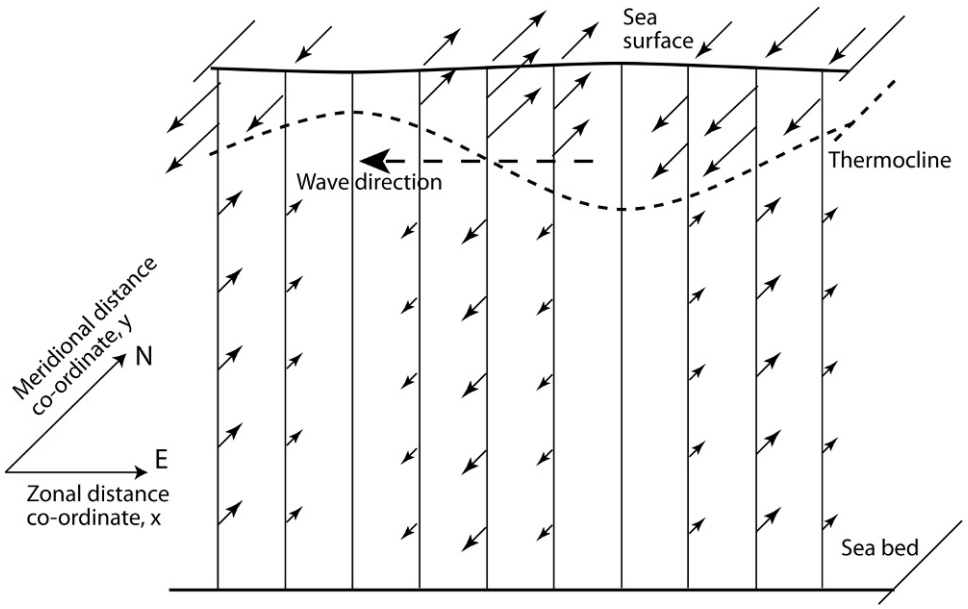


Figure 6.9. Schematic of the first baroclinic mode of planetary waves in the northern hemisphere.

Perturbations of thermocline depth create much weaker pressure gradients than corresponding slopes in sea surface height, because the density difference between the layers is typically much less than 1/100 of the actual density. The "reduced gravity" forces in baroclinic flows cause baroclinic waves to respond much more slowly than the barotropic case. This gives baroclinic planetary waves the slow propagation speeds and long characteristic timescales that make them so difficult to observe without the long-term monitoring capacity of satellite remote sensing.

Planetary waves were first studied in the atmosphere. Rossby (1940) solved a simplified form of the full equations of motion which approximated variation of the Coriolis parameter, f, with latitude as $f = f_0 + \beta y$. This so-called "β-plane approximation" linearizes the variation of f over a narrow range of latitudes centered at ϕ_0 where $f = f_0$, and where the origin $y = 0$ of the meridional distance co-ordinate, y, is located. Note that f is defined as $2\Omega \sin \phi$ where Ω is the Earth rotation rate and θ is the latitude. Therefore

$$\beta = \frac{\partial f}{\partial y} \propto \cos \phi. \tag{6.1}$$

For a full development of the equations of motion and their solution, including the special treatment of the equatorial zone where $f \to 0$ but β remains finite, readers should consult standard texts (e.g., Gill, 1982; Pedlosky, 1987). The response of the ocean to disturbances of a fairly large spatial and temporal extent is examined by assuming periodic perturbations of key dynamic variables with a zonal (west–east)

wavenumber of k per unit distance in x, a meridional (south–north) wavenumber of l per unit distance in y, and a frequency ω per unit time. Thus we look for solutions which represent key variables, V, such as height of the sea surface, thermocline displacement, or components of the velocity field, in the form:

$$V = A \exp[i(kx + ly - \omega t)], \tag{6.2}$$

which represents a propagating wave–like phenomenon if k, l, and ω are real.

Equation (6.2) is a valid solution of the fundamental equations of motion on the β-plane only if the following condition is satisfied:

$$\omega_n = \frac{-k\beta}{k^2 + l^2 + \lambda_n^{-2}}, \tag{6.3}$$

where λ_n is the Rossby radius of deformation; and ω_n is the frequency for the nth mode. The Rossby radius represents the distance traveled by perturbations in the time it takes them to spin up to geostrophic balance. This is discussed in Section 3.2.2 in relation to mesoscale eddies. Here the Rossby radius provides a horizontal lengthscale that represents the shortest wavelengths that can satisfy the assumption of geostrophy. Its size depends on the vertical density structure of the water column and varies according to the mode of wave propagation.

From (6.3) it is easy to derive an expression for the phase speed, c_x, of zonal propagation for the nth mode:

$$c_{nx} = \frac{\omega_n}{k} = -\frac{\beta}{k^2 + l^2 + \lambda_n^{-2}}. \tag{6.4}$$

The first thing to note about this is that for wave-like propagation (when k and l are real numbers), phase speed is always negative, because B is always positive in both hemispheres. As x is defined to be positive eastwards, this means that phase velocity must be always towards the west. However, at first sight the expression in (6.4) seems to be quite complex, and is dispersive; that is, zonal speed varies with both zonal and meridional wavenumbers, k and l. However, it should be noted that the lengthscale of the wave-like phenomena detected in SSHA and SST Hovmöller plots is of the order of hundreds of kilometers which is generally much larger than the typical Rossby radius. Therefore the horizontal wavenumbers are small and it is reasonable to assume that

$$k^2, l^2 \ll 1/\lambda^2. \tag{6.5}$$

Given (6.5), Equation (6.4) can be reduced to the much simpler, and nondispersive form:

$$c_{nx} = -\beta\lambda_n^2. \tag{6.6}$$

This theoretical prediction that a disturbance of sufficiently large horizontal dimension propagates nondispersively in a western direction is entirely consistent with satellite observations of features that propagate coherently for thousands of kilometers and for many months at the same speed. For nondispersive waves, phase velocity and group velocity are the same, and independent of wavelength. Thus, once sufficient energy has been fed into a large-scale perturbation of the layered structure

of the upper ocean, a planetary wave motion (such as shown in Figure 6.9) is likely to develop. This then propagates westwards at a steady speed, carrying energy with it until some other process interferes to extract the energy or until it reaches the western coast of the ocean. If the wave propagation process were dispersive, then the propagation speed would depend on the spatial dimension of the initial perturbation, and the energy associated with different scales would tend to flow westwards at different speeds. This would almost certainly prevent the appearance of coherent diagonal parallel lines in Hovmöller plots.

It appears that satellite data provide evidence of planetary waves which behave in a similar way to those predicted by Rossby more than 50 years before satellite sensors were deployed. However, whether satellite observations detect true Rossby waves (i.e., waves whose characteristics are essentially those predicted by Rossby wave theory) requires further quantitative analysis of the data. In particular we need to measure propagation speed, and how it varies with latitude, in order to compare observations with theory. From Equation (6.6) the approximate dependence on westward wave speed can be estimated. Using (6.1), and also noting from (3.1) that the Rossby radius varies inversely with f, we conclude that phase speed varies with latitude as $\cos\phi/\sin^2\phi$, provided the density distribution in the water column remains uniform. We shall return to comparison between observed and theoretical wave speeds in Section 6.5 following a discussion in Section 6.4 of how to measure feature propagation speeds from Hovmöller plots. First, we need to consider why it is that planetary waves should be seen at all from Earth-orbiting satellites.

6.3.2 How can Rossby waves be seen at the sea surface?

We have already concluded from propagation characteristics implied by diagonal patterns on Hovmöller plots that they are the signatures of planetary waves, but that still leaves open the questions, "Why can we detect planetary waves at all?" and, "Could there be planetary waves present that are not being detected by satellite data?" Because we are looking for the baroclinic form of Rossby waves, where the largest vertical perturbations due to the waves are expected at the thermocline rather than at the sea surface, is it not rather foolish to expect satellite oceanography methods to offer anything at all?

In fact it is relatively easy to explain the SSHA signatures of planetary waves. Implicit in the outline theory of planetary waves in the previous subsection, and particularly Figure 6.9, is that associated with these waves there is a displacement of sea surface relative to geopotential surface. Its magnitude is of the order of only a few centimeters, compared with a very much larger displacement of the thermocline and isopycnal surfaces inside the ocean, but this is the reason why Rossby waves were explicitly sought in the altimeter record as soon as measurement uncertainties for SSHA were reduced to a few centimeters. Because they are such slow and gradual phenomena, planetary waves do not stand out clearly without the painstaking removal of the tidal signal and other factors which might artificially create signatures of comparable magnitude. As noted in Section 6.2.3 the first attempts to use Geosat

data in the 1980s never had a sufficiently accurate measure of surface height to be able to conclusively demonstrate detection of planetary waves. That has all changed since the TOPEX/Poseidon mission and more recently Jason-1, which have transformed altimetry over the ocean and enabled other concurrent altimeters to also achieve comparable accuracies of a few centimeters (see chapter 11 of *MTOFS*).

It should be noted that the inability of altimetry to measure absolute heights and currents until the geoid is independently known is not a serious drawback for the detection of planetary waves, since these are, by definition, perturbations from the ocean's mean state. Of course, even with processing to optimize the disclosure of slow wave propagation (as discussed in Section 6.4), we can only measure those wave signatures which stand out above the noise level of the SSHA record. Careful filtering (as discussed in Section 6.4) can help to improve estimation of wave speed, but it remains likely that there are weaker signatures of smaller amplitude planetary waves which are being missed at present.

The discovery of planetary wave signatures in SST anomaly datasets came as more of a surprise. Because horizontal density distribution is not part of the primary Rossby wave mechanism (as discussed in Section 6.3.1), there was no direct reason to expect Rossby wave–like characteristics in the density or temperature field. The most simple explanation is that the north–south currents associated with planetary waves are advecting meridional gradients of climatological temperature distribution. Thus poleward velocities in either hemisphere, acting on the background gradient of temperature reducing towards the poles, will tend to increase temperature and equatorward velocities will decrease it. Considering Figure 6.9, this would imply that the maximum SSTA should occur with a 90° phase lead ahead of maximum SSHA. This appears to be the case when wave speeds are the same in SSTA and SSHA plots (see, e.g., Quartly *et al.*, 2003; Challenor *et al.*, 2004).

A similar mechanism could account for the ocean color signature, although this is more problematic since the climatological distribution of chlorophyll concentration has a more complex latitude dependence than SST. An alternative explanation of the apparent dependence of color/chlorophyll on the passage of Rossby waves relies on changes in mixed layer depth associated with planetary waves. Where the mixed layer is shallow (corresponding in a first-mode Rossby wave to where the layer interface is highest and the surface level is at its lowest, see Figure 6.9) the chlorophyll signature might be expected to be greatest. The reasoning is that either the deep chlorophyll maximum is raised to a depth where it is more readily visible from an ocean color sensor, or there might even be increased upwelling associated with the shallow mixed layer which raises nutrients and enhances production (Uz *et al.*, 2001).

In some cases it appears that different remote-sensing methods detect a different mode of planetary wave, as indicated by propagation speed. When this is observed, typically SST and/or color signatures propagate more slowly than the SSHA (Cipollini *et al.*, 1997; Quartly *et al.*, 2003; Challenor *et al.*, 2004). This implies that the process which perturbs temperature or color is more sensitive to higher modes of planetary wave, if present, than is the SSHA. The prospect of being able to distinguish different modes of planetary waves in a given location revealed by different ocean parameters may point the way to understanding both the imaging

mechanisms for waves and the impact waves may have on the ocean through which they pass. For such research to be based on quantitative rather than qualitative analysis, analytical tools needed to be designed especially for the unique characteristics of Hovmöller plots. These are described in the next section.

6.4 ESTIMATING PLANETARY WAVE SPEED

6.4.1 Methods for analyzing Hovmöller diagrams

The most useful information that can be extracted from Hovmöller plots showing planetary waves is the propagation speed of the waves. This is estimated by measuring on the image the inclination from the vertical of wave characteristic lines, defined as angle θ in Figure 6.10. If the source of the data is a time series of composite images with a spatial resolution of d km in the east–west direction and sampled once every b days, then the resolution of cells in the resulting Hovmöller image is d km $\times$ b days. The westward speed, V_p, of planetary waves is then:

$$V_p = \frac{d}{b}\tan\theta \quad \text{km/day}. \tag{6.7}$$

The simplest way of estimating wave speed is by measuring the inclination of signatures directly from the Hovmöller image using a ruler and protractor. This method is subject to error, mainly because it is difficult to draw a line representing the average or typical orientation of the stripes on a Hovmöller plot. Although in the examples shown in Figures 6.5, 6.6, and, 6.7 there is little doubt about the general trend of the signatures, their slope is variable, and the magnitude of the signal varies along the lines. When first produced from the original composite time series, Hovmöller plots may be quite noisy, with various other processes tending to mask the characteristic diagonal stripes signifying planetary waves.

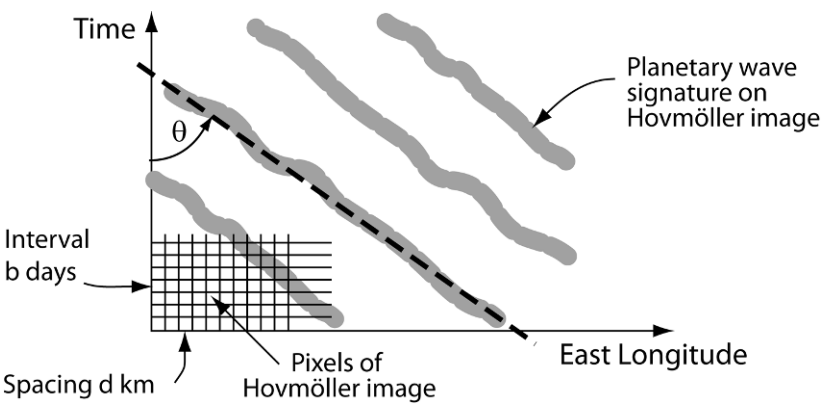


Figure 6.10. Schematic to show how planetary wave speed depends on the slope of wave signatures in time–longitude (Hovmöller) plots.

Low-pass spatial filters can be applied to the plot to remove high-frequency noise, while special filters can remove stationary features at a given longitude or sudden events spanning the longitude range of the plot at particular times. Indeed spectral filters can be designed to suppress all features that are not propagating westwards or that have wavenumbers and frequencies outside a given range (see, e.g., Cipollini *et al.*, 2001). Note, however, that overzealous prefiltering of the data should be avoided, since taken to the limit it would tend to exclude any outcomes other than the sought-after signal at the expected speed! Even after the plots have been filtered, measurements "by eye" are subjective; different observers would end up with slightly different estimates of wave speed. To eliminate this, a more objective method is needed.

The use of Fourier analysis is a possibility. The two-dimensional fast-Fourier transform (2D-FFT) is a well-developed numerical tool available for analysis of images. It generates a spectral plot of signal energy in (longitude wavenumber, frequency) space. If peaks can be identified, the speed of the features can be estimated as the ratio of frequency/wavenumber, and different peaks may be interpreted as evidence of different Rossby wave modes (Cipollini *et al.*, 1997; Subrahmanyam *et al.*, 2001). However, this approach is not as successful as might be expected because planetary waves are not always regular in their frequency and wavenumber. Disturbances having different longitudinal extent will propagate at the same speed because planetary waves are nearly nondispersive. They do not consist of a regular train of troughs and crests like swell waves, which would show up as a sharp peak on an FFT spectrum. Their spacing on the Hovmöller plot may be quite irregular, so that no spectral peak is found in FFT results, even though the uniformity of speed means that the different crests still appear as parallel lines. Moreover, to achieve the spectral precision needed to calculate speed accurately in this way, the 2D-FFT must be applied to a wide extent of data spanning many troughs and crests. Consequently this method is not optimal for measuring the speed of isolated or irregular planetary waves (Cipollini *et al.*, 2006b). To do this requires a spatial analysis tool designed specifically for identifying the orientation of parallel streaks on an image, irrespective of wavenumber or frequency (as described in the next subsection).

6.4.2 Radon transform

The Radon transform (Deans, 2007) was developed specifically for the task of determining the orientation of alignments in image fields. An image, $f(x,y)$, is projected onto a frame of axes (x',y') rotated through an angle θ from the original Cartesian (x,y) plane in which $y = x' \sin\theta + y' \cos\theta$ and $x = x' \cos\theta - y' \sin\theta$ (as shown in Figure 6.11). The transform is then simply defined as:

$$p(x',\theta) = \int f(x,y)\, dy'. \tag{6.8}$$

Equation (6.8) sums all the values of the image along the perpendicular to the x'-axis (parallel to the y'-axis) for each x'. For a single value of θ the result is a one-dimensional function, $p(x')$. If the rotation, θ, of the baseline is such as to align the

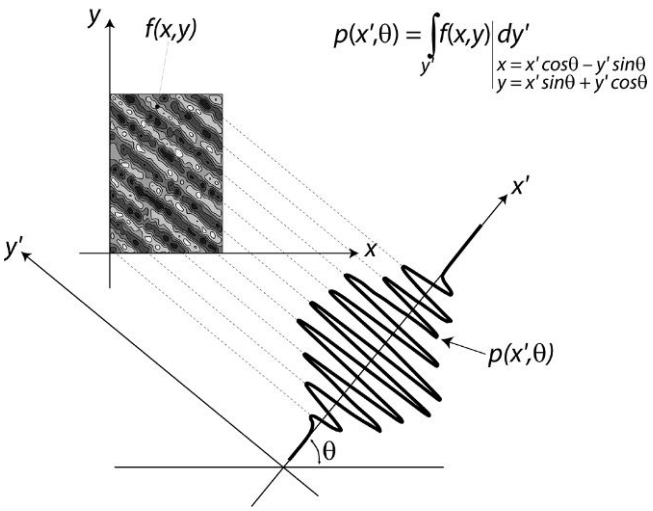


Figure 6.11. The principle of the Radon transform. Example of the one-dimensional function $p(x')$ when the transform is applied at a single orientation, θ (figure provided by P. Cipollini, based on figure 1 in Challenor *et al.*, 2001).

y'-axis with linear features in the image, this function separates out the troughs and crests of the image (as illustrated in Figure 6.11), which is an artificially generated plot with linear structures at about 40° to the horizontal. For other values of θ, each integration path cuts across troughs and crests in the image, and the resulting function $p(x')$ is very much smoother with little if any structure. The full two-dimensional Radon transform, $p(x', \theta)$, is constructed by evaluating the integral across a range of θ, resulting in an image such as Figure 6.12, in which the enhanced structure is apparent for $\theta \approx 40°$.

When applying the Radon transform to the study of planetary waves, the final stage of the process is to evaluate the variance of $p(x')$ for each value of θ. When this

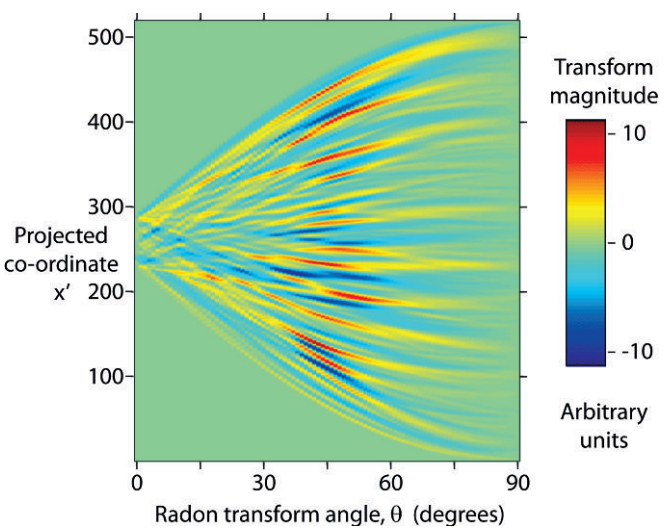


Figure 6.12. Example of the full two-dimensional transform $p(x', \theta)$ for a Hovmöller field such as Figure 6.11 (adapted from Cipollini *et al.*, 2006b).

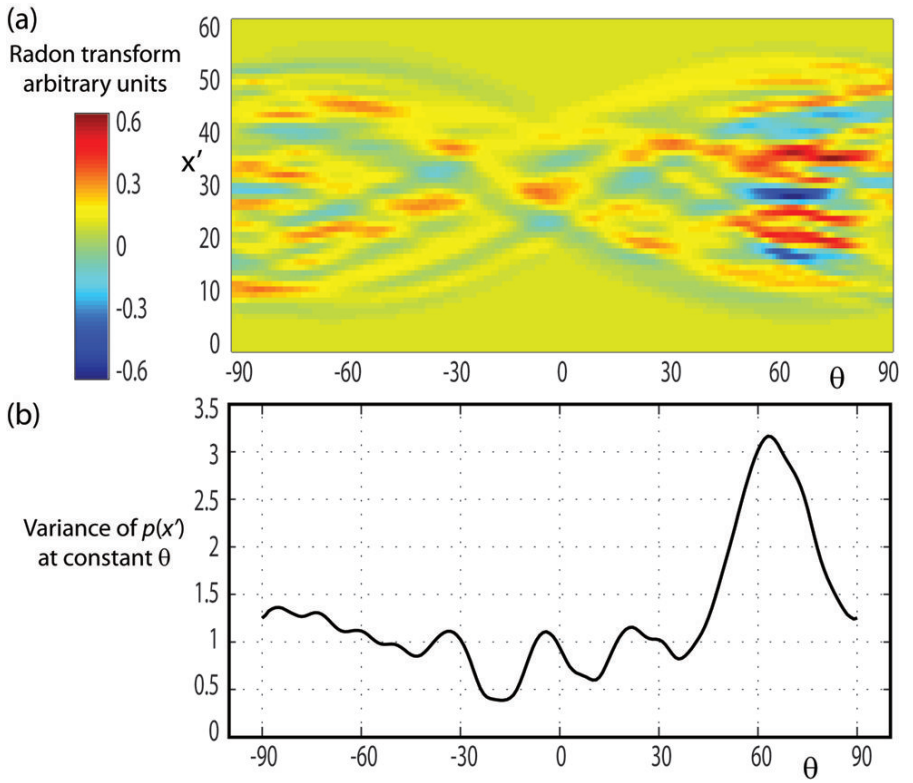


Figure 6.13. (a) Example of the Radon transform. (b) The corresponding variance–direction plot. This is the Radon transform for AATSR monthly data at 95°E, 25°S (as shown in Figure 6.6 and analyzed by Hill *et al.*, 2000).

is plotted as a function of θ (as shown in Figure 6.13) the peaks in variance are found at values of θ corresponding to the alignment of planetary wave signatures in Hovmöller plots. The width of the peaks indicate the variability of the wave speed across the segment of the plot that was analyzed, and provide an objective estimate of not only wave speed but also the error or uncertainty attached to that measure. This is a far more acceptable scientific approach than subjective estimates by eye.

The example in Figure 6.13 shows an unambiguous peak, which is consistent with the dominant direction of the stripes in Figure 6.6. This is not always the case, and care must be taken when interpreting peaks in a Radon transform variance direction plot. Smaller peaks may not be significant in comparison with the background. In some studies prefiltering has been applied to Hovmöller diagrams, to remove energy propagating east or at wavelengths or frequencies inappropriate for Rossby waves, but this should be done cautiously; overzealous filtering could spuriously detect what the filter permits rather than the propagating signal that is actually present.

6.4.3 Mapping the speed of planetary waves

Given a set of analytical tools based on the Radon transform approach outlined above, it is relatively easy to analyze entire ocean basins to identify the occurrence of planetary waves and to map their westward propagation speeds. Hovmöller plots are generated from the original time series of global maps, segmenting oceans zonally into slices of, typically, 1° latitude. For each latitude slice, the Radon transform is applied successively to vertical slices of a Hovmöller plot, each corresponding to a limited longitude range (typically 10°). If tapered prefiltering is used the resulting maps are smoother (Hill *et al.*, 2000). For long time series of data it is also possible to segment the timespan and so detect any changes of wave speeds over time, but it is advisable to apply the Radon transform over at least a 3-year up to a 5-year span.

When this is repeated for all zonal slices, the resulting measurements of westward wave speed can be presented as global maps (such as those in Figure 6.14). Note that no speeds are presented within 5° latitude of the Equator. This is because planetary wave speeds increase towards the Equator and the first-mode wave speed is fast enough that characteristic propagation patterns are almost parallel to the x-axis of the Hovmöller plots (θ in Equation 6.7 tends to 90°). This makes them either indistinguishable from instantaneous perturbations across the longitude window, or severely reduces the accuracy with which orientation can be measured. Outside the equatorial zone, the speed reduces with increasing latitude as expected, although there are distinct variations with longitude in some regions.

When zonal mean speed is considered (averaged across all longitudes for a given latitude) longitudinal variability is hidden, but it allows the latitude dependence of speed to be explored and compared with theoretical predictions. Figure 6.15 shows how zonal mean speed (plotted logarithmically on the y-axis) varies with latitude for various different measurements. Those derived from TOPEX/Poseidon SSHA and those from ATSR SST are very similar except north of 35°N where SST signatures indicate a significantly lower speed than SSHA signatures. This may be because the SST is revealing a higher wave mode at these latitudes. The speeds derived from ocean color (SeaWiFS) signatures at a limited number of locations are also plotted, not as zonal averages but as individual sample points. In the Indian and Atlantic Oceans they are similar to the SSHA and SST, but more divergence is found in the Pacific Ocean. Comparison with theoretically predicted curves is discussed in Section 6.5.

6.4.4 Meridional components of planetary wave propagation

So far it has been assumed that planetary waves are likely to propagate due west, although it is theoretically possible for their propagation velocity to have a meridional component. In fact the methodology presented above to measure the speed using Hovmöller time–longitude plots and two-dimensional Radon transforms is incapable of detecting whether a meridional component of propagation velocity is present. To explore this possibility further, one way to proceed would be—considering the data cube illustrated in Figure 6.4—to cut different time slices along

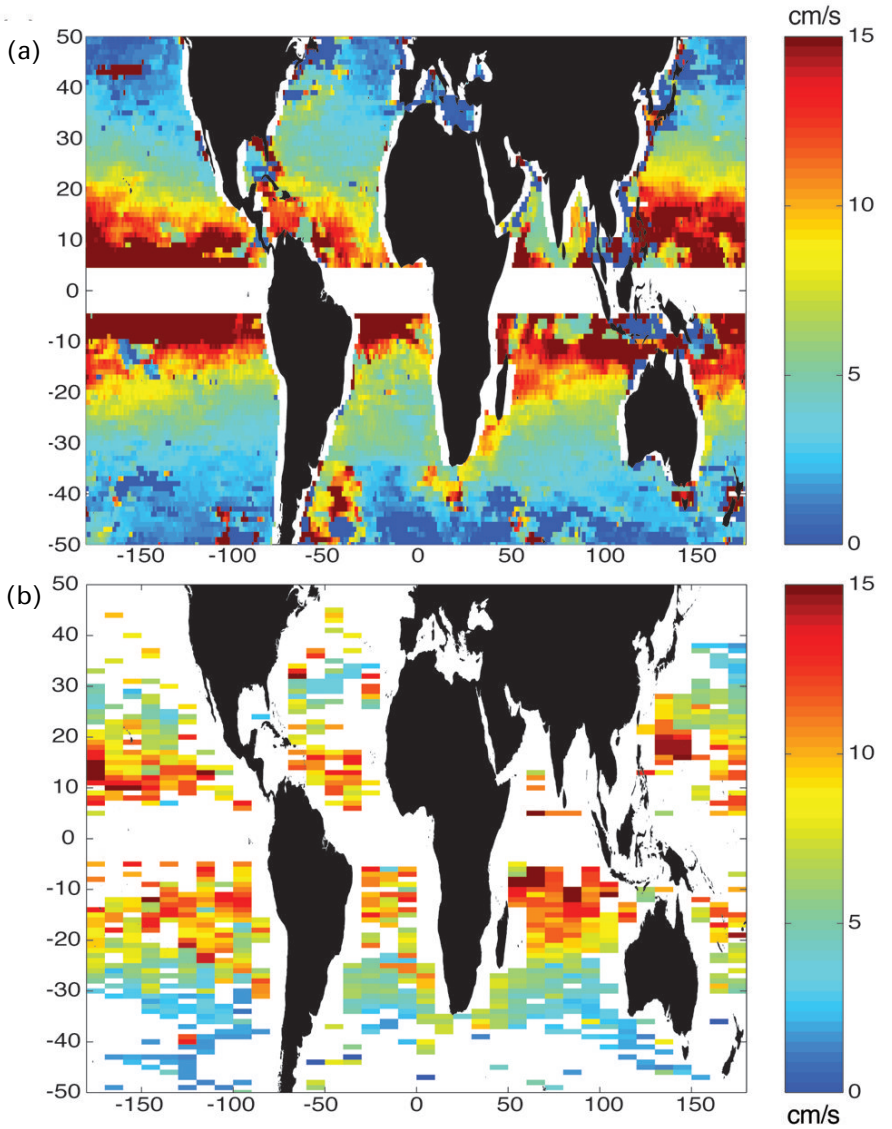


Figure 6.14. Global maps of planetary wave speed measured using Radon transforms of Hovmöller plots from (a) SSHA from combined TOPEX/Poseidon and ERS altimetry data for the period 1992–2002 (Cipollini *et al.*, 2006b) and (b) SST from ATSR for 1991–1995 (Hill *et al.*, 2000).

trajectories slicing diagonally across lines of both latitude and longitude, before applying the two-dimensional Radon transform to the resulting time–length plot.

In fact the principle of the Radon transform can be applied directly to a small cuboid volume within the data cube, searching for dominant lineation across a three-

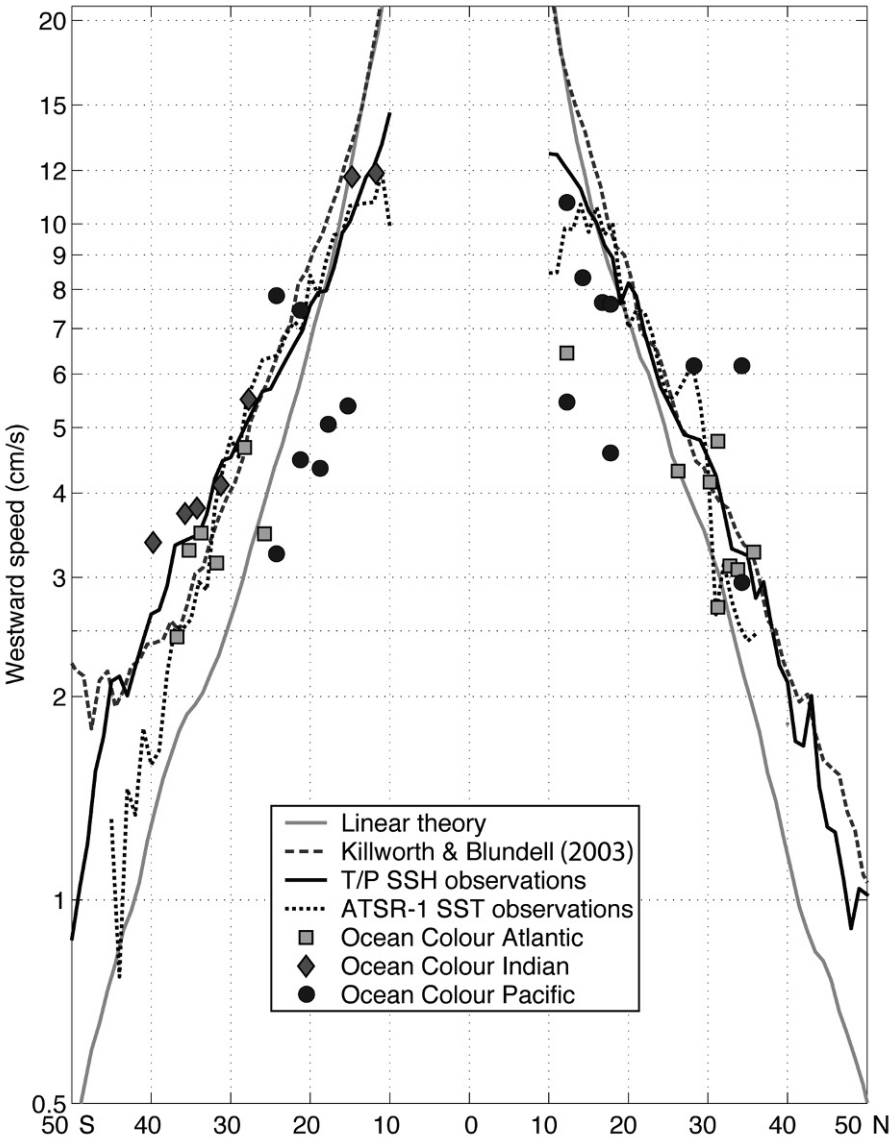


Figure 6.15. Zonal mean speed of planetary waves detected by their signature in different data types, and comparison with estimates based on Rossby wave theory (derived from Challenor *et al.*, 2004).

dimensional range of directions. This is the basis of the three-dimensional Radon transform method (Challenor *et al.*, 2001). Although the transform is applied three-dimensionally, the result can be presented in a two-dimensional plot representing the true geographic directions. Figure 6.16 is an example from TOPEX/Poseidon data.

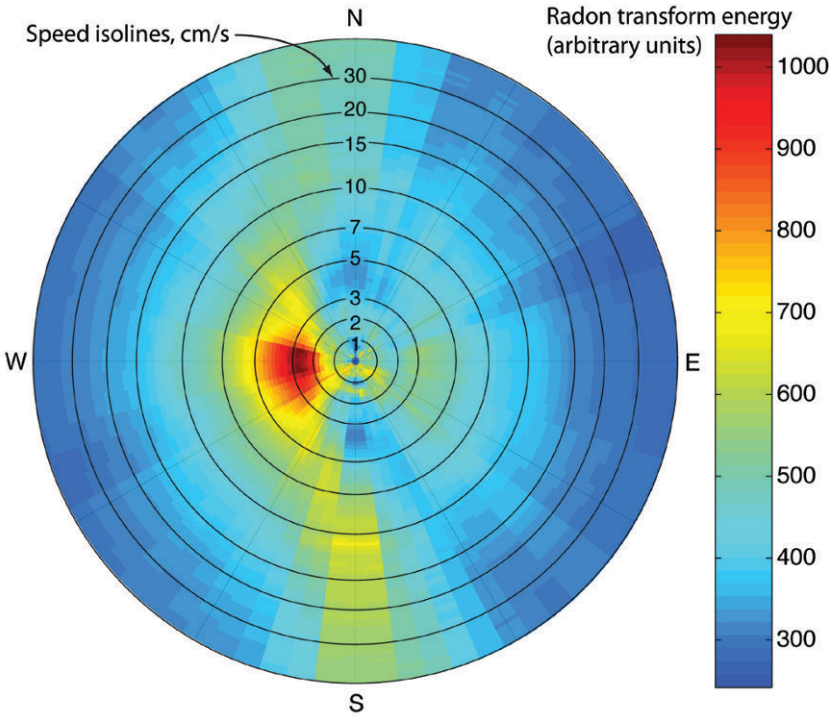


Figure 6.16. Polar plot of energy from the three-dimensional Radon transform of TOPEX/Poseidon SSH anomalies over the region of 5°(latitude) × 11°(longitude) centered at 34°N, 36°W and a timespan of 234 TOPEX/Poseidon cycles. Due north is at the top of the polar plot, and so the "compass angle" or azimuth yields the direction of wave propagation. Radial distance from the center is the elevation angle, whose corresponding speed in centimeters per second is shown by the concentric lines. A strong energy peak can be seen at a speed of 2.6 cm/s and about 265° (measured clockwise from north) implying that waves are propagating in a direction a few degrees south of due west (derived from Challenor, *et al.*, 2001).

The field in this polar plot is in arbitrary units representing the relative energy of the signal detected by the transform, based on variance of the Radon transform on the plane normal to the viewing direction. The radius from the center corresponds nonlinearly to the velocity magnitude. Where there are peaks of energy, it implies that a signal has been detected propagating in that direction with that speed.

In this example, the dominant signal, as expected, is westwards. The spread to north and south must partly be a consequence of a lack of precision in the method, with energy leaking from the dominant beam into adjacent directions. However, the asymmetry of the main peak may be indicating a slightly southward component of the main westward propagation. In addition, there do appear to be a few places on the figure where some propagation is occurring with a significant meridional component. This analysis and plot must be repeated for all cuboids making up the data cube, showing how the directionality of wave propagation varies geographically and

also with time during the whole data record. Challenor *et al.* (2001) were able to produce maps of the North Atlantic indicating regions where some deviation from purely westward propagation occurs. Although this type of analysis has not been applied extensively, it shows promise and can be expected to be developed in future.

6.5 UNDERSTANDING ROSSBY WAVES BETTER

6.5.1 Satellite data confirm the existence of Rossby waves

From the foregoing sections of this chapter it is clear that satellite data provide a reliable means of detecting and measuring the speed of large-scale, baroclinic, planetary waves in the ocean. In this section we consider the extent to which satellite data have been instrumental in advancing our scientific understanding of the phenomenon, starting with the question of their actual existence.

Since its development in the 1930s the theory of Rossby waves, developed and validated in relation to atmospheric dynamics, was expected to apply within the ocean as well as the atmosphere. Certain baroclinic motions were assumed to be a consequence of oceanic Rossby waves. However, because of their large horizontal scales and relatively weak surface manifestation, it was not possible unequivocally to demonstrate their existence in the ocean from observations. Several decades elapsed before firm evidence of Rossby waves was found in time series observations of ocean baroclinic structures (Emery and Magaard, 1976; White, 1977). Even then, the difficulty of monitoring the time evolution of baroclinic motions over a wide area, when limited to *in situ* observations, prevented any opportunity of testing the accuracy of propagation characteristics predicted by Rossby wave theory. The advent of satellite altimetry began to change that, since it provided regularly repeated measurements of the ocean, spatially detailed sampling, and global ocean coverage.

Early altimeter results showed the potential for remote sensing to solve the space-time sampling problem for the study of Rossby waves (Fu and Chelton, 2001). Then from 1992 onwards (as already discussed in Section 6.2.3) the altimetric performance of TOPEX/Poseidon opened up a full study of planetary wave behavior, made possible because the measurement accuracy and space-time sampling capability of the sensor system at last met the requirements for observing the phenomenon. Finally, once processing techniques had been developed to reveal planetary waves in the altimeter record, the idea came to search for them in SST and ocean color records, with the positive results described in Sections 6.2.3 and 6.2.4.

What this brief history demonstrates is that remote-sensing methods are no different from all other oceanographic observing techniques; when designed to meet the required measurement accuracy and achieve the appropriate sampling capability, they will lead to improved scientific understanding of ocean phenomena and open up new fields of research exploration.

6.5.2 Revisiting Rossby wave theory

The new understanding of planetary waves gained so far from the use of remote-sensing techniques has concerned their propagation characteristics and in particular the validity of predictions of basic Rossby wave theory. It is inherent in the scientific method that when a theory is tested by confronting its predictions with observations, either the theory is upheld or else any significant discrepancies are used as the basis for revising the theoretical model. The remote sensing of Rossby waves has been no exception. Because Radon transform analysis allows wave speeds to be measured to known confidence limits, it is possible to test this aspect of the theory.

Figure 6.15 shows the comparison between zonal mean wave speeds observed by different remote-sensing methods, and speed predicted by theory. It is evident that there is a definite mismatch between observation and the original linear Rossby wave theory. Poleward of latitude 20° the observed speed is significantly greater than the theory predicts, by a factor of 2 and even 3 in places. The first reliable evidence of this discrepancy immediately stimulated a response from ocean dynamics theorists (Killworth *et al.*, 1997), who concluded that in ignoring background baroclinic flow through which waves propagate the original theory had underestimated the speed. A number of other papers added further refinements, the details of which are beyond the scope of this book, and eventually the revised model of Killworth and Blundell (2003a, b) obtained the much closer agreement shown by the black dashed line (labeled 2002) in Figure 6.15.

Model predictions require estimates of the baroclinic Rossby radius and so depend on ocean hydrography and background flow, leading to a map of predicted, first-mode wave speeds. When this is matched with observations from 6 years of altimetry, the ratio between theory and observation is that shown in Figure 6.17.

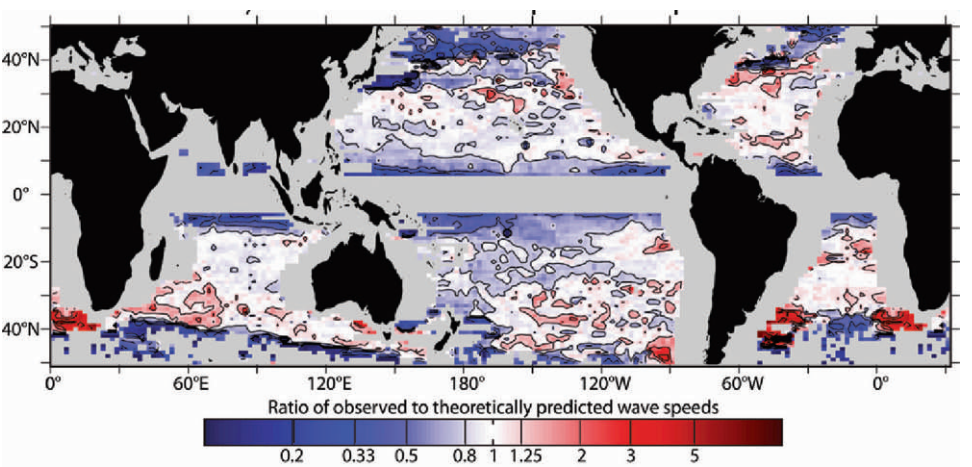


Figure 6.17. Ratio between planetary wave speed measured from multimission altimetry data and the first-mode, theoretical Rossby wave speed predicted by the revised model of Killworth and Blundell (2003a, b, 2004, 2005) and based on *in situ* data from the 2005 version of the *World Ocean Atlas* (figure provided by P. Cipollini (as used in Cipollini *et al.*, 2010).

There is generally good agreement in zones between latitudes 10° and 35° in both hemispheres. Equatorward of this observations are slower than predicted. At higher latitudes the situation is much noisier, presumably because the amplitude of waves here is very small and measurements may be unreliable. Negative values occur where eastward-propagating waves have been detected in the altimeter record. This happens in regions associated with major ocean currents where there may be strong barotropic flow advecting planetary waves eastward even though they still propagate westward relative to the flow.

Further insights about Rossby wave behavior can be gained from comparisons between the planetary wave signatures on Hovmöller plots of SSHA, SST, and ocean color. For example, the apparently different speeds of SST and SSHA signatures suggest that SST is detecting a higher mode wave. Phase relationships between different signatures is also a fertile field for further research (see, e.g., Quartly *et al.*, 2003; Cipollini *et al.*, 2006b). Fuller explanations of how Rossby waves produce a signature in surface color and temperature fields will lead to better understanding of whether, or how, the passage of Rossby waves through the ocean promotes vertical mixing and so changes near-surface hydrography, nutrients, or primary production. While there is as yet little indication that the ''Rossby rototiller'' effect (Siegel, 2001) is more than a very isolated phenomenon, there is scope for more investigation into this subject, by both data analysts and theoreticians. An important consideration is the rate at which any perturbation to ocean surface properties decays back to its climatological equilibrium, in relation to the period of the Rossby wave itself. It may be possible to estimate this from comparison of phase relationships between different signatures (Killworth *et al.*, 2004).

Another technical development which is likely to bring more illumination to the propagation of planetary waves is a means for identifying, isolating, and tracking individual waves (Cipollini *et al.*, 2006a). This is an important development because it recognizes that, although use of the term ''waves'' often implies a regular train of oscillations with a particular wavenumber and frequency, planetary waves tend to consist of isolated pulses of energy propagating with Rossby wave characteristics. Thus when comparing different signatures of planetary waves, or attempting to forecast their future influence on the ocean, it may be more helpful to consider each wave crest and trough as an individual phenomenon rather than as part of a wider field of waves. Having the digital tools to plot automatically the trajectory of each wave, and to track merging or bifurcation of crests or troughs, will facilitate an interesting new research direction.

It is also interesting to note that the team who first confirmed the presence of Rossby waves in the altimeter record are now questioning whether some of the low-frequency energy ascribed to Rossby waves is in fact associated with large ocean eddies (Chelton *et al.*, 2007). They automatically track the propagation of eddies in the altimeter record using a technique developed by Isern-Fontanet *et al.* (2003, 2006), which was briefly mentioned in Section 3.4.3. The implication of this is that the energy in eddies would propagate with somewhat different characteristics than Rossby waves.

6.5.3 The importance of Rossby waves

In the end we must ask whether Rossby waves really matter. It could be argued that their impact on the ocean is so small that they can barely be detected, even after a large amount of complex data analysis and processing. They propagate extremely slowly. They appear to lack the capacity of mesoscale dynamics to stir and mix the ocean. Mariners who avoid hurricanes and respect the power of wind waves and swell have no cause to worry about Rossby waves! The cynic might ask whether recent interest in improving Rossby wave theory is simply an academic response to the scientific challenge of testing a 60-year-old theory? While all these comments are valid, there is potentially much more significance to Rossby waves than has yet been demonstrated. This concerns our capacity to forecast medium-range changes in the ocean and their impact on regional climate change.

Climatic conditions which influence weather patterns depend on interactions between the sea and the air above it. In the complex interplay of processes which couple atmospheric and ocean dynamical systems, the ocean tends to have a stabilizing role. Being far more massive than air, the sea tends to absorb or release heat in response to changes in air temperature, or surface wind forcing, without itself changing very much. Nonetheless, when the ocean is subject to a persistent, large-scale change in atmospheric forcing, anomalies of temperature or vertical density structure may develop over wide areas. While the magnitude of such anomalies may be small in terms of SST or SSH, significant amounts of energy may be stored baroclinically. When any dynamical system is perturbed, one consequence is that wave motions are initiated which carry information about the perturbation, as energy, to other parts of the system. The trajectory of that information, its speed, and direction is dictated by permissible wave modes. For large-scale ocean perturbations spanning hundreds of kilometers, Rossby waves represent the primary ocean response.

For this reason if we are to be able to predict the response of the ocean when subject to major atmospheric anomalies, an understanding of planetary wave dynamics is essential. It has long been understood (e.g., Gill, 1982) that they provide the only means for transferring information about forcing anomalies from the east side of an ocean basin to the west, if no east–west topographic barrier exists to support boundary waves. Because the information is carried so slowly by planetary waves (with typical speeds between 1 and 10 cm/s depending on latitude) it can take several years to cross the ocean. This means that the ocean may absorb an excess input of heat from the atmosphere during, for example, an anomalously warm summer off its eastern margin and effectively "bury" it within its baroclinic structure. This perturbation travels slowly westward with very little interaction with the atmosphere until, several years later, it reaches the western margin of the ocean and starts to interact strongly with the western boundary current there. It may change the strength or route of currents such as the Gulf Stream or Kuroshio Current (Jacobs *et al.*, 1994), and feedback to the atmosphere may then become much stronger. It may be critical what time of year this interaction occurs. Thus knowing the exact speed of the planetary wave acquires great importance.

For this reason, it is not merely of academic interest to determine whether planetary waves behave in the way Rossby wave theory predicts. The improvements to Rossby wave theory that were first stimulated and later validated by new information from remote sensing are bound to have beneficial impacts on the modeling of large-scale, air–sea interactions and hence medium-term climate forecasting. Just as interesting is the possibility of recording altimeter measurements of the peaks and troughs of Rossby waves from the past few years right up to the present, and using that to forecast the likely propagation path of these waves in future. Moreover, if the SSTA signatures of individual planetary waves are also monitored, their phase and trajectory may be used to forecast the likely impact of ocean feedback on the atmosphere. These and other ideas are expected to provide a basis for important research in the coming years in which the role of oceanic Rossby waves in the climate system is explored in more depth.

In summary, the use of satellites to monitor planetary waves has provided the 21st-century climate scientist with (a) an improved theoretical model of Rossby wave behavior, (b) a means of observing the ocean response to large-scale climate perturbations in the atmosphere, and (c) a tool to track the trajectory of resulting ocean anomalies. One goal of future research must surely be to use this knowledge and these monitoring tools to improve medium-term forecasting of climate changes over a few months to years.

6.6 OTHER LARGE-SCALE PROPAGATING PHENOMENA

Planetary waves provide a strong example of how effective satellite remote sensing is for monitoring large-scale ocean processes, but they are not the only phenomenon benefiting from the availability of satellite data. In this final part of the chapter a number of other examples will be introduced, although there is space only to mention each topic briefly and point to the key literature where more information can be found.

What these phenomena have in common with planetary waves is that they are not immediately apparent in basin-scale or global maps of ocean properties. In fact they are hidden in the time series of global image data and need special analytical techniques to reveal them. They span a range of frequencies, wavelengths, and propagation speeds. Equatorial Kelvin waves are fairly fast-moving phenomena which benefit from techniques such as singular value decomposition to pick out their characteristic propagation signatures. Tropical instability waves are slower than Kelvin waves but faster than Rossby waves, and their mechanism depends on air–sea interaction processes. Phenomena such as Antarctic circumpolar waves and the Madden–Julian Oscillation have lower frequencies, and need the accumulation of a long time series of consistent data before they can be detected against a background of higher frequency signals. Another example for this section would be the El Niño phenomenon, but this is left to Chapter 11 where it is considered along with other large-scale ocean features that have a strong human impact.

6.6.1 Equatorial Kelvin waves

Kelvin waves are a theoretical model of long waves in a rotating ocean in which wave currents flow only in vertical planes parallel to the propagation direction of the wave. Geostrophic forces are balanced by a surface slope in the direction normal to the propagation direction, so that the amplitude of surface displacement due to the waves increases exponentially to the right (left) of the wave propagation direction in the northern (southern) hemisphere. Such waves propagate with the same speed, $(gh)^{1/2}$, where h is the water depth at which long waves would travel in a nonrotating frame of reference.

Baroclinic versions of such waves also behave in a similar way, with a large thermocline downward displacement matching a surface upward displacement (and *vice versa*) so that below the thermocline there is no horizontal pressure gradient associated with the Kelvin wave and almost all the motion is contained in the surface layer. The ratio of surface displacement to thermocline displacement is approximately $\delta\rho/\rho_{\text{mean}}$, where $\delta\rho$ is the density difference between the two layers; and ρ_{mean} is mean water column density. Baroclinic Rossby waves travel at correspondingly much slower speeds because of the reduced gravity effect (Gill, 1982). However, this theoretical model can apply only to real ocean phenomena if the exponentially increasing amplitude is constrained by a boundary. Thus waves of this type are found propagating along ocean margins, with the coast to the right of travel in the northern hemisphere. The one exception to this requirement for Kelvin waves to be boundary waves is found at the Equator where the Coriolis parameter changes sign with latitude. As shown in Figure 6.18 it is possible here for a wave to propagate from west to east, with its amplitude decreasing exponentially with increasing latitude to both north and south, and its maximum amplitude at the Equator where there is no geostrophic force. The situation could be considered as two Kelvin waves, one in each hemisphere, using each other in place of a coastal boundary. Such waves, trapped at the Equator, are called equatorial Kelvin waves (EKWs) and are not capable of propagating from east to west, which is the opposite of Rossby waves (discussed in Section 6.3.1).

Can such waves be detected in satellite data like Rossby waves can? Since they are not associated with any meridional currents, EKWs are not expected to have a strong SST or color signature, but they are defined by perturbations in sea surface height. Therefore the most promising method for detecting them is in Hovmöller plots of SSHA at the Equator, measured by satellite altimetry. Consider the speed of EKWs. In an ocean of depth 4 km, the propagation speed of barotropic EKWs would be 200 m/s. At this speed, the waves would take about 14 h to travel 10,000 km (90° of longitude). Given the resampling time of 10 days for the TOPEX/Poseidon and Jason altimeter systems, there is no possibility of detecting such waves from satellites. However, baroclinic EKWs are likely to have speeds which are two orders of magnitude slower. At a speed of 2.5 m/s a wave crest would take about 50 days to travel across 90°. This corresponds to five cycles of TOPEX/Poseidon and therefore five rows of pixels in the Hovmöller plot. If the SSHA is mapped onto a 1° grid, a signal propagating at the speed of the first

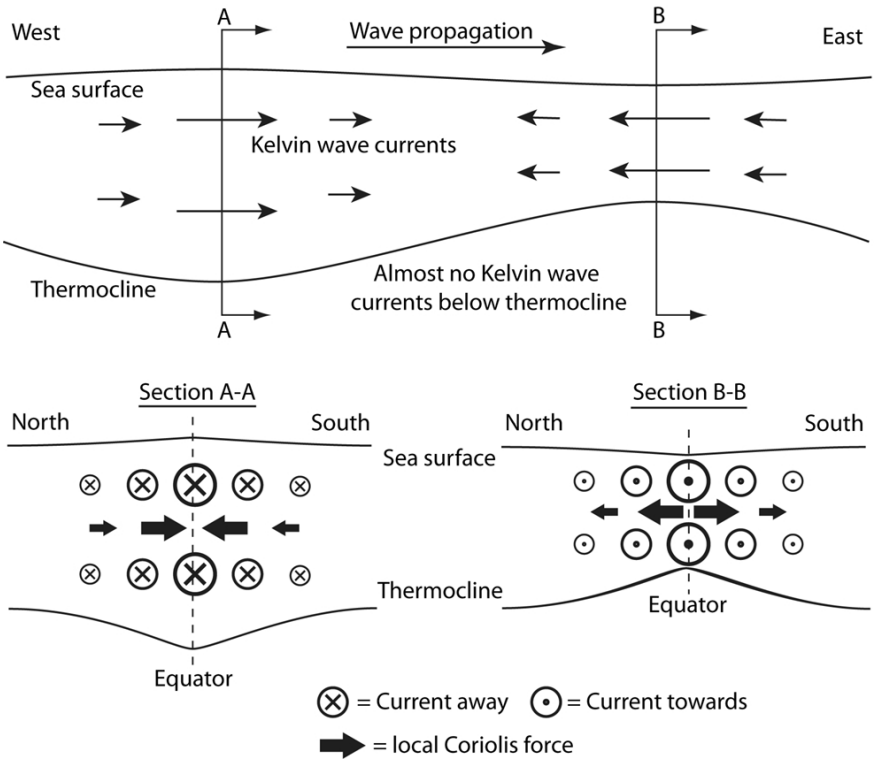


Figure 6.18. Schematic of currents, vertical displacements, and Coriolis forces in a baroclinic, equatorial Kelvin wave. The *upper diagram* is a section looking north, the lower views are sections looking east.

baroclinic mode EKW would have a slope of 5 time steps in 90 spatial samples, which should be detectable. However, the Atlantic Ocean is only about 50° wide at the Equator and EKWs would cross it in around 30 days, barely detectable at a resolution of 10 days.

The Pacific Ocean is therefore the place to look for altimetric evidence for EKWs. Figure 6.19 shows an example of TOPEX/Poseidon data, plotted as longitude *vs.* time, which reveals patches of raised and lowered sea level between 170°E and 90°W. If these occurred simultaneously across the basin, the contours between them would be vertical in this figure (note it is flipped and rotated 90° compared with previous examples of Hovmöller plots). Instead the contours are slightly inclined, showing that changes in SSHA start in the west and propagate to the east at a fast but detectable rate. A number of publications (e.g., Delcroix *et al.*, 1991) reported attempts to identify EKWs in Geosat data, but only since TOPEX/Poseidon data became available have there been convincing observations (Boulanger and Menkes, 1995; Boulanger and Fu, 1996). Recently EKWs have been detected in the Atlantic Ocean too (Pollo *et al.*, 2008).

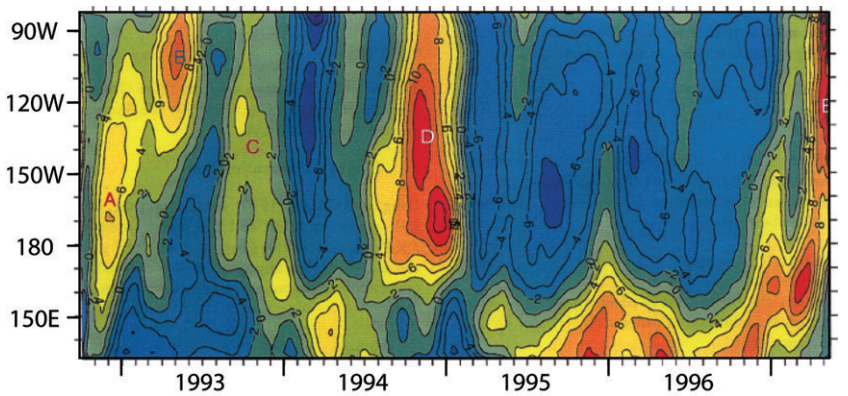


Figure 6.19. SSHA from TOPEX/Poseidon plotted in longitude, time at the Equator. The contour interval is 2 cm. The red zones are $> +12$ cm and the dark-blue zone is < -8 cm. Several major peaks can be seen, as well as several other linear features which slope slightly indicating eastward-propagating equatorial Kelvin waves (data plot from Zheng *et al.*, 1998).

So what is the value of satellite measurements of EKWs? First, it is important to confirm that the wave-like behavior of the ocean predicted by theoretical models at these large scales is consistent with observations. Second, EKWs are believed to be the mechanism by which the Pacific Ocean responds to perturbations in the western equatorial region. For example, an anomalous wind burst that is strong and persistent enough to significantly perturb the depth of the thermocline on the western side of the ocean will, when the wind relaxes, trigger a wavelike eastward propagation of the disturbance which will reach the eastern side of the Pacific in about 2 months. Such a process is highly relevant to understanding the mechanisms of the El Niño phenomenon (Picaut *et al.*, 2002) (see also Chapter 11). Moreover, if it could be detected while it is taking place, there is the potential to improve forecasting of the wider climatic phenomenon.

Because Kelvin waves are nondispersive, there is no particular wavelength or frequency attached to propagating signals. Availability of the altimetry time series allows the space-time structure of SSH perturbations to be studied using techniques such as singular value decomposition (Susanto *et al.*, 1998). It has also been possible to study the nonlinear behavior of large-amplitude disturbances which have been shown to adopt the form of solitons (Zheng *et al.*, 1998).

6.6.2 Tropical instability waves

Since they were first reported by Legeckis (1977), the cusp-shaped frontal waves illustrated in Figure 6.20 have been observed frequently in the equatorial Pacific and Atlantic. Now referred to as tropical instability waves (TIWs), although sometimes called Legeckis waves, the phenomenon is attributed to strong latitudinal shears between various components of the complex equatorial current system,

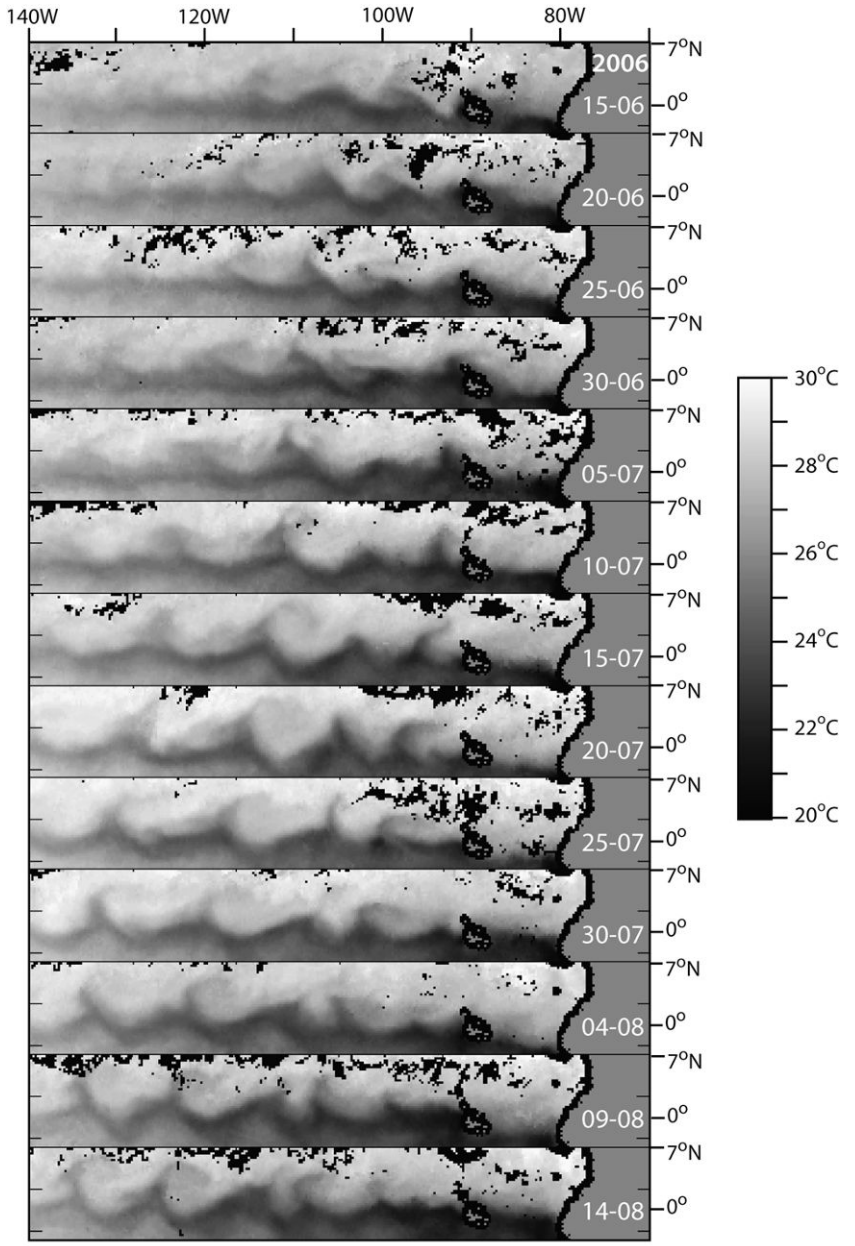


Figure 6.20. Sixty-day sequence at 5-day intervals (starting at the top of the figure) of AMSR-E 3-day composite SST images (June 15–August 14, 2006) in the equatorial Pacific Ocean, showing tropical instability waves. Each subimage spans from 7°N to 3°S. The AMSR-E data used here were produced by Remote Sensing Systems and sponsored by the NASA Earth Science REASoN DISCOVER Project and the AMSR-E Science Team (data are available at *www.remss.com*).

causing some flows to become unstable (Philander, 1978). Instabilities cause large perturbations of the SST front between colder upwelled water of the Pacific equatorial cold tongue and warmer water to the north (Kennan and Flament, 2000). These wave-like perturbations, with wavelengths of 1,000 km to 2,000 km and periods of 20 to 40 days, propagate westwards with a phase speed of about 0.5 m/s (Qiao and Weisberg, 1995). Thus they travel about 1,200 km in a month, much faster than Rossby waves, but about one-fifth the speed of EKWs and in the opposite direction.

In the 60-day period illustrated in Figure 6.20, SST derived from the AMSR-E microwave radiometer shows that TIWs are very active. The cold upwelling core is displaced both north and south, so that wave motion straddles the Equator, but the extreme cusp-like formations along the sharpest front occur only to the north. These cusps tend to roll up into eddies which stir cold water farther north as the waves migrate westwards and carry energy into a previously less disturbed region. A good example of an evolving vortex can be seen at 108°W on July 10, 2006, reaching 120°W by August 4. The black blobs are data dropout, probably due to heavy rain cells contaminating the microwave signal.

TIWs were first discovered in infrared thermal images from geostationary orbiting radiometers, and then studied extensively using other infrared sensors (e.g., Allen *et al.*, 1995), *in situ* data (Halpern *et al.*, 1988), and ocean models (Masina and Philander, 1999). Most recently, SST data retrieved from orbital microwave sensors, TMI, and AMSR-E (Chelton *et al.*, 2000; Hashizume *et al.*, 2001; Caltabiano *et al.*, 2005) have proved to be ideal for monitoring them. Although the coarse resolution (~50 km) of microwave radiometers misses some details that are detected by IR radiometers, the data are gridded every 1/4° of latitude and longitude, and they capture the dominant characteristics of the SST signature of TIWs without difficulty. The advantage of the microwave radiometer is its capacity to resample completely within 3 days, independently of cloud which often obstructs the view of IR sensors. This almost unimpeded sampling (apart from cells of rain) is important in order to capture the energetic variability of this phenomenon. SST seems to provide the clearest signatures of TIWs. SSHA data from altimetry can sometimes reveal the general propagation characteristics of TIWs, but lack both temporal and spatial resolution to define the detailed instantaneous spatial structure of waves, which is more akin to a field of mesoscale eddies (see Chapter 3).

In the same way as for Rossby and Kelvin waves the propagation characteristics of TIWs are revealed in time–longitude plots. Figure 6.21a, an example from the equatorial Atlantic Ocean, clearly reveals diagonal streaks consistent with the typical speed of TIWs. However, the TIW signal is obscured by other strong temperature signals, including the seasonal cycle, and filtering is recommended in order to isolate it from other SST clutter (as illustrated in Figure 6.21b). Caltabiano *et al.* (2005) derived this from the data shown in Figure 6.21a, employing a westward-only, two-dimensional, finite impulse response filter, following the method by Cipollini *et al.* (2001). The filter uses prior knowledge of the approximate range of the period and wavelength of waves in order to restrict output to a particular region of the frequency–wavelength spectrum. In this example a one-quarter length size sample kernel was used, with a bandpass of 5° to 20° in longitude and 20 to 40 days in time.

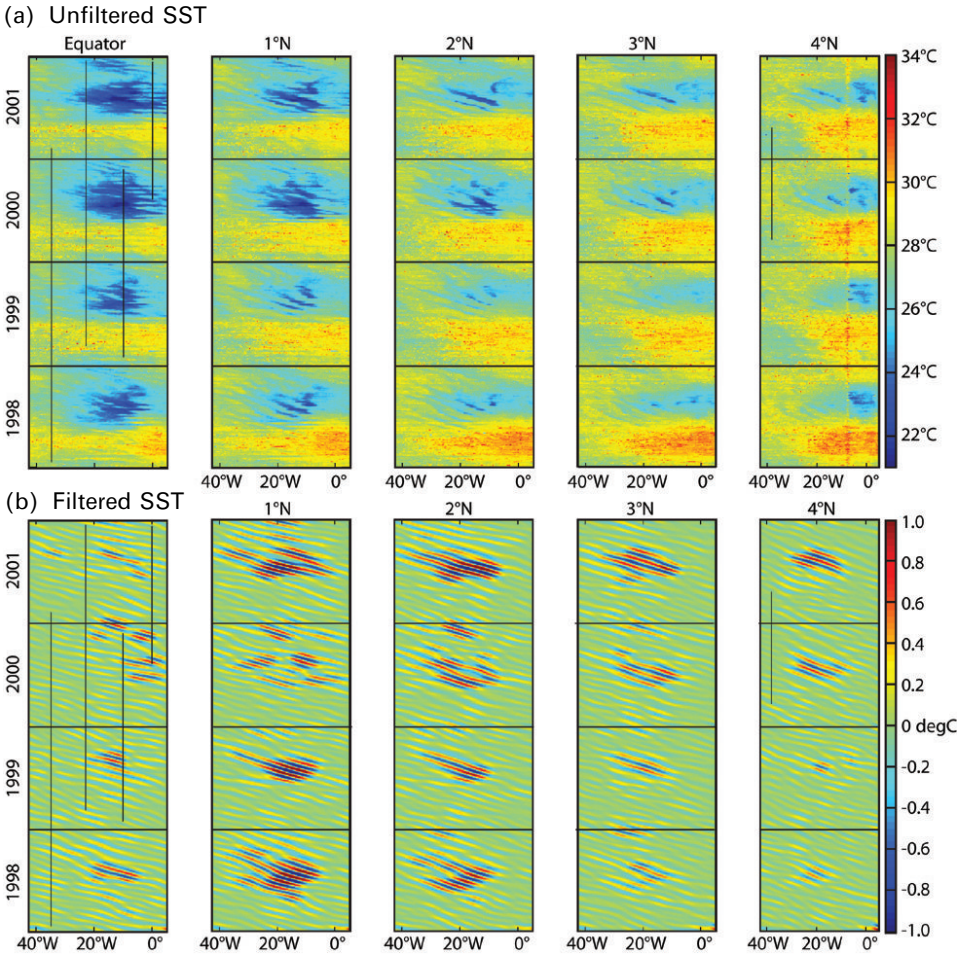


Figure 6.21. Time–longitude plots of temperature in the equatorial Atlantic for latitudes 0°N, 1°N, 2°N, 3°N, and 4°N, spanning 4 years from January 1998 to January 2002. (a) SST derived from TMI, showing bursts of TIW activity. The color scale is in degrees Celsius. (b) The result of band-pass filtering the record in (a). The color scale represents the SST anomaly in Kelvins (images from the PhD thesis of A. Caltabiano, University of Southampton, U.K.).

The result clearly shows the SST anomaly associated with TIWs plus any other signals or noise in the original record which share band-pass spectral characteristics. As intended, the timing and location of TIW activity are highlighted. Not only is this seasonal, occurring only during the months of northern-hemisphere summer, but there is considerable interannual variability. The longitude of maximum activity in the Atlantic appears to be always about 15°W to 18°W, but the energy spreads wider on either side of this in years when waves are strongest. The latitude of the activity is also variable. In general TIWs are strongest at 1°N and 2°N with a weaker signal at

the Equator. However, in 2000–2001 they are found to be equally strong out to 4°N whereas in 1998–1999 they barely extend north of 2°N. When these results are related to the original SST maps, it can be seen that the presence and strength of TIWs depends on development each year of the cold upwelling tongue.

It would be wrong to interpret the low-amplitude signal in all parts of Figure 6.21b as unequivocal evidence of background TIW activity, since the filter distorts the spectral and propagation characteristics of noise or any other processes present in the original signal. However, what the filtered figure does reveal is the value of an uninterrupted 3-day record of the SST field from satellite microwave radiometers allowing the complex variability of TIW activity to be tracked without difficulty.

TIWs are important because they appear to modify mean oceanic currents by reducing the shears between them (Hansen and Paul, 1984; Weisberg, 1984). They also interact with the atmosphere. While they are only indirectly dependent on the atmosphere through wind forcing of equatorial currents which create the horizontal shear flows that drive the instability, TIWs have an impact on the atmosphere, introducing variability with similar 20 to 30-day periodicities into the formation of cloud (Deser *et al.*, 1993; Hashizume *et al.*, 2001), air–sea heat flux (Thum *et al.*, 2002), and wind (Hayes *et al.*, 1989; Chelton *et al.*, 2001; Liu *et al.*, 2000; Hashizume *et al.*, 2002). Because wind, cloud, rain, and water vapor are also detectable from satellites (in some cases the same microwave radiometer that measures SST), it is relatively easy to correlate the variability of ocean and atmospheric variables, and thus confirm the strength and spectral characteristics of the ocean's impact on the atmosphere. For example, by applying the same band-pass filter to atmospheric fields as to SST in Figure 6.21b, Caltabiano *et al.* (2005) demonstrated strong coupling between the SST and the wind components, and pointed towards the influence that TIWs have on the intertropical convergence zone (ITCZ) over the Atlantic Ocean.

6.6.3 The Madden–Julian Oscillation

The Madden–Julian Oscillation (MJO) is strictly an atmospheric rather than an ocean phenomenon although it is thought to play a role in various large-scale air–sea interaction processes. It is mentioned here because recently it has been shown that remote-sensing measurements over the ocean are able to contribute to monitoring and better understanding the MJO.

Since it was first identified as an oscillation in the zonal wind of the tropical Pacific with a period of 40 to 50 days (Madden and Julian, 1971), meteorologists discovered that the MJO phenomenon is a pattern of atmospheric circulation in which a center of deep atmospheric convection, with associated rainfall, moves steadily eastward along the Equator at a speed of about 5 m/s, flanked to the east and west by zones of much weaker convection (Madden and Julian, 1994). Although most apparent over the Indian and Pacific oceans, the MJO circles the globe and its influence has been detected in many important aspects of tropical meteorology, including the variability of rainfall in the Pacific Islands, the monsoon regions of Asia, in North and South America and in Africa, the growth of tropical cyclones in

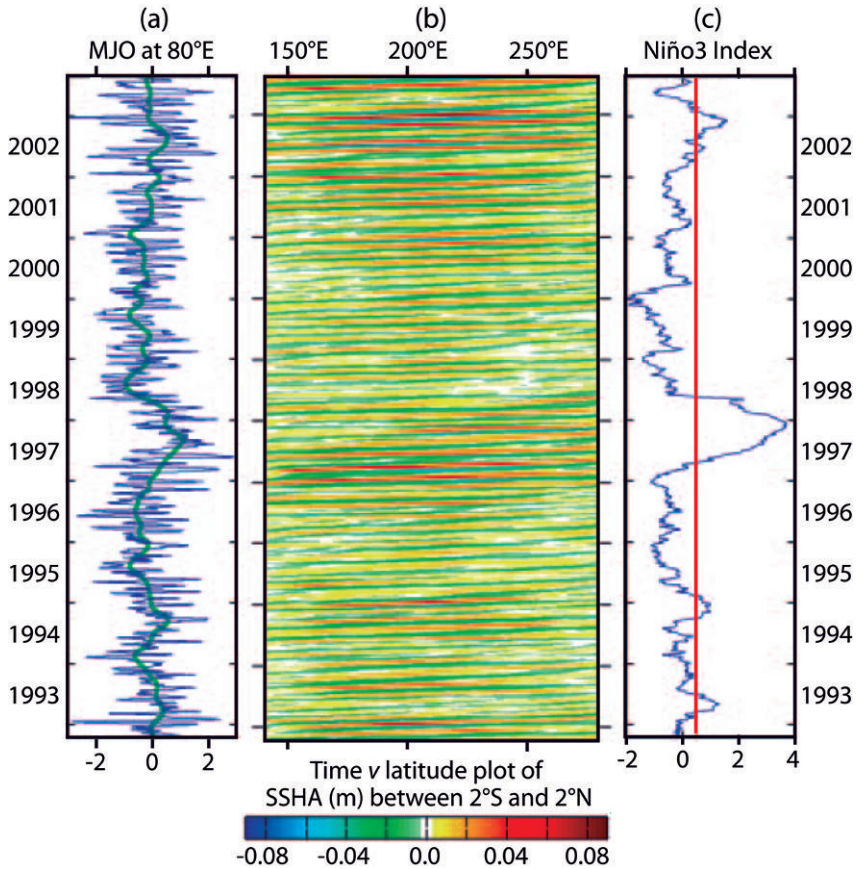


Figure 6.22. (a) MJO 80°E index over the period of study; the green curve is the MJO index filtered with a 160-day Hanning window. (b) Longitude–time plot of the 2°S to 2°N, averaged, SSHA filtered data showing oceanic Kelvin waves. (c) The Niño3 index, with the red line indicating the threshold for El Niño events (adapted from Edwards *et al*., 2006).

the Pacific and Caribbean, and equatorial winds over the Atlantic (Zhang, 2005). It is now recognized that the MJO is the dominant component of intraseasonal variability in the tropical atmosphere, even though it is proving to be difficult to forecast it using numerical models.

The MJO also has a strong influence on the ocean through winds affecting SST and rainfall, while connections with the El Niño–Southern Oscillation phenomenon (ENSO—see Chapter 11) have been identified (Zhang and Gottschalck, 2002). It is in this context that the methods of satellite oceanography have been brought to the study of the MJO. Edwards *et al*. (2006) analyzed the altimetric record of SSHA to confirm that equatorial Kelvin waves provide a connection between the MJO and ENSO. Figure 6.22 shows an example of their analysis. On the left is the MJO index varying with time along the vertical axis. To the right is the ENSO index and there is

evidently a lagged correlation of the ENSO to the MJO. The central panel is a Hovmöller plot of SSHA across the Pacific, filtered to highlight Kelvin wave propagation from west to east.. When there are peaks in the MJO index, Kelvin waves appear to have a higher amplitude, implying that the ocean propagates the signal from the MJO that influences the El Niño. This may have the potential to serve as an approach to ENSO forecasting, although more research is needed to determine whether this is feasible.

6.6.4 The Antarctic Circumpolar Wave

The presence of the large-scale Antarctic Circumpolar Wave (ACW) was first observed and characterized by White and Peterson (1996), who analyzed a time series from 1982 to 1995 of the monthly anomalies of SST, sea level atmospheric pressure, and sea ice extent over the Southern Ocean. They discovered that within a broad interannual frequency band (periods of 3 to 7 years) perturbations of these variables were correlated right round the Southern Ocean, implying some form of atmosphere–ocean coupling. The dominant energy was found to have a zonal wavenumber of 2 (i.e., there were two complete cycles of phase circling the southern hemisphere at mid to high latitudes) and a period of 4 years. This structure propagated slowly eastward around the Southern Ocean, each of the variables remaining in fixed phase with each other, taking about 8 years to circle the globe once, at a speed of about 8 cm/s. Relative phases of different variables are shown in Figure 6.22, this pattern rotating slowly eastwards relative to the continents.

Coincident with this discovery a similar correlation was found between SSH and SST (Jacobs and Mitchell, 1996), suggesting that variability in the Antarctic circumpolar current was also coupled to the phenomenon. Note that although the phase structure in Figure 6.22 travels east, it does so more slowly than the Antarctic Circumpolar Current, and thus the ACW actually propagates westwards relative to the water flow in the Southern Ocean.

Since the ACW was discovered, it has been the focus of fairly wide research interest. For example, it has been shown to be linked to weather patterns and rainfall in New Zealand and Australia. Its connection to the ENSO phenomenon has also been studied (White *et al.*, 2002). Several attempts have been made to explain the air–sea interaction processes which maintain the phenomenon. Similar behavior has been found in coupled ocean–atmosphere models, although the models also suggest that several other different types of behavior are possible. A key underlying question is where the forcing energy for the ACW comes from, and what controls its propagation. Since the Southern Ocean connects with all the other major oceans, a phenomenon like this has the potential to influence worldwide ocean circulation and climate. Yet there remains some controversy about how important the ACW really is. Because it is not immediately obvious in the observed data and appears only from correlation analysis in the interannual frequency band, not all scientists are convinced that it has a major role to play in understanding and forecasting short-term climate changes in the southern hemisphere. While further analysis of the observations has led to a refinement of the spectral characteristics of the ACW,

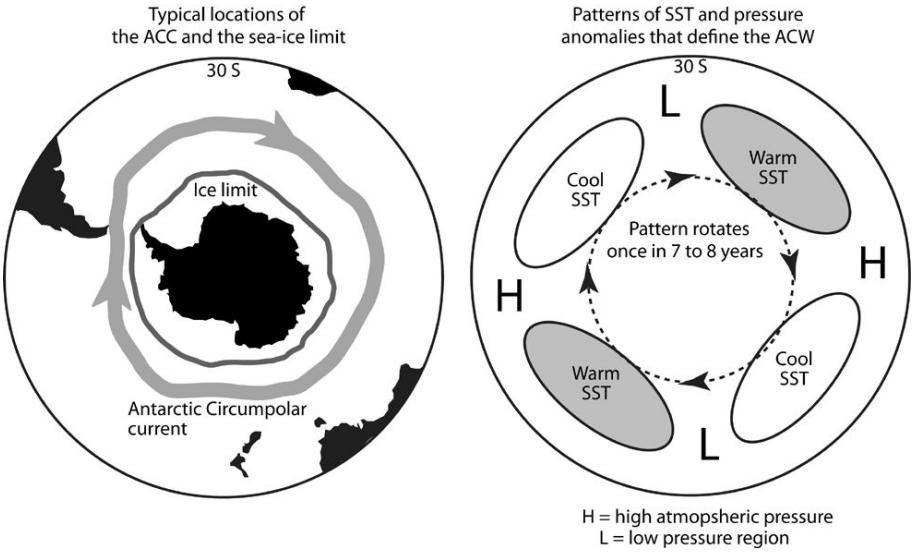


Figure 6.23. Sketch of the approximate positions of the ACC and sea ice limits around Antarctica (*left*) and the relative phase distribution of SST and surface pressure anomalies in the ACW phenomenon (*right*).

the analysis of extended records from before 1985 suggests that the stable ACW structure defined in Figure 6.23 is just one of many possible coupling states between the ocean and atmosphere. It appears to have dominated between 1985 and 1994, but other less organized patterns of behavior may have dominated at other times (Connolley, 2002).

So far, most observational studies of the ACW have used the global oceanic and atmospheric datasets available to climate scientists, derived from a variety of observational sources. However, three of the key variables, SST, SSHA, and sea ice extent, are derived from satellite ocean sensors. As the quality of these data steadily improves, and as the span of the data record steadily extends, there will be scope for further research into the ACW which can build on the experience of the types of analysis developed to study all the other ocean dynamic processes mentioned in this chapter. It is also important to ensure that the high quality of today's SST and SSHA measurement series is continued into the indefinite future, as far as possible without breaks. Phenomena of interannual variability may require many decades of quality-controlled satellite data acquisition before all their dynamical secrets can be revealed.

6.7 REFERENCES

Allen, M. R., S. P. Lawrence, M. J. Murray, C. T. Mutlow, T. N. Stockdale, D. T. Llewellyn-Jones, and D. L. T. Anderson (1995), Control of tropical instability waves in the Pacific. *Geophys. Res. Letters*, **22**, 2581–2584.

Boulanger, J.-P., and L.-L. Fu (1996), Evidence of boundary reflection of Kelvin and first-mode Rossby waves from TOPEX/POSEIDON sea level data. *J. Geophys. Res.*, **101**(C7), 16361–16372.

Boulanger, J.-P., and C. Menkes (1995), Propagation and reflection of long equatorial waves in the Pacific Ocean during the 1992–1993 El Niño. *J. Geophys. Res.*, **100**(C12), 25041–25060.

Caltabiano, A. C. V., I. S. Robinson, and L. P. Pezzi (2005), Multi-year satellite observations of instability waves in the Tropical Atlantic Ocean. *Ocean Science*, **1**, 97–112.

Challenor, P. G., P. Cipollini, and D. D. Cromwell (2001), Use of the 3D Radon transform to examine the properties of oceanic Rossby waves. *J. Atmos. Oceanic Tech.*, **18**(9), 1558–1566.

Challenor, P. G., P. Cipollini, D. D. Cromwell, K. L. Hill, G. D. Quartly, and I. S. Robinson (2004), Global characteristics of Rossby wave propagation from multiple satellite datasets. *Int. J. Remote Sensing*, **25** (7/8), 1297–1302.

Chelton, D. B., and M. G. Schlax (1996), Global observations of oceanic Rossby waves. *Science*, **272**, 234–238.

Chelton, D. B., M. G. Schlax, R. M. Samelson, and R. De Szoeke (2007), Global observations of large oceanic eddies. *Geophys. Res. Letters*, **34**(L15606), doi: 10.1029/2007GL030812.

Chelton, D. B., F. J. Wentz, C. L. Gentemann, R. de Szoeke, and M. G. Schlax (2000), Satellite microwave SST observations of transequatorial tropical instability waves. *Geophys. Res. Letters*, **27**, 1239–1242.

Cipollini, P., D. D. Cromwell, M. S. Jones, G. D. Quartly, and P. G. Challenor (1997), Concurrent altimeter and infrared observations of Rossby wave propagation near 34°N in the Northeast Atlantic. *Geophys. Res. Letters*, **24**, 889–892.

Cipollini, P., D. D. Cromwell, P. G. Challenor, and S. Raffaglio (2001), Rossby waves detected in global ocean color data. *Geophys. Res. Letters*, **28**, 323–326.

Cipollini, P., P. G. Challenor, and S. Colombo (2006a), A method for tracking individual planetary waves. *IEEE Trans. Geosc. Remote Sensing.*, **44**(1), 159–166.

Cipollini, P., G. D. Quartly, P. G. Challenor, D. D. Cromwell, and I. S. Robinson (2006b), Remote sensing of extra-equatorial planetary waves. In: J. F. R. Gower (Ed.), *Manual of Remote Sensing, Vol. 6: Remote Sensing of Marine Environment* (pp. 61–84). American Society for Photogrammetry and Remote Sensing, Bethesda, MD.

Cipollini, P., A. C. S. Sutcliffe, and I. S. Robinson (2010), Oceanic planetary waves and eddies: A privileged view from satellite altimetry. In V. Barale, J. F. R. Gower, and L. Alberotanza (Eds.), *Oceanography from Space, Revisited.* Springer Science/Business Media BV.

Connolley, W. M. (2002), Long-term variation of the Antarctic Circumpolar Wave. *J. Geophys. Res.*, **107**, doi: 10.1029/2000JC000380.

Deans, S. R. (2007), *The Radon Transform and Some of Its Applications* (304 pp.). Dover Publications.

Delcroix, T., J. Picaut, and G. Eldin (1991), Equatorial Kelvin and Rossby waves evidenced in the Pacific Ocean through Geosat sea level and surface current anomalies. *J. Geophys. Res.*, **96**, 3249–3262.

Edwards, L. A., R. E. Houseago-Stokes, and P. Cipollini (2006), Altimeter observations of the MJO/ENSO connection via Kelvin waves. *Int. J. Remote Sensing*, **27**(5/6), 1193–1203.

Emery, W. J., and L. Magaard (1976), Baroclinic Rossby waves as inferred from temperature fluctuation in the eastern Pacific. *J. Marine Res.*, **34**, 365–385.

Fu, L.-L., and D. B. Chelton (2001), Large-scale ocean circulation. In L.-L. Fu and A. Cazenave (Eds.), *Satellite Altimetry and Earth Sciences* (pp. 133–170). Academic Press, San Diego, CA.

Gill, A. E. (1982) *Atmosphere–Ocean Dynamics* (International Geophysics Series, Vol. 30, 662 pp.). Academic Press, San Diego, CA.

Halpern, D., R. A. Knox, and D. S. Luther (1988), Observations of 20-day period meridional current oscillations in the upper ocean along the Pacific Equator. *J. Phys. Oceanogr.*, **18**, 1514–1534.

Hashizume, H., S. P. Xie, W. T. Liu, and K. Takeuchi (2001), Local and remote atmospheric response to tropical instability waves: A global view from space. *J. Geophys. Res.—Atmos.*, **106**, 10173–10185.

Hill, K. L., I. S. Robinson, and P. Cipollini (2000), Propagation characteristics of extratropical planetary waves observed in the AATSR global sea surfacae temperature record. *J. Geophys. Res.*, **105**(C9), 21927–21945.

Isern-Fontanet, J., E. Garcia-Ladona, and J. Font (2003), Identification of marine eddies from altimetric maps. *J. Atm. Ocean. Tech.*, **20**, 772–778.

Isern-Fontanet, J., E. Garcia-Ladona, and J. Font (2006), Vortices of the Mediterranean Sea: An altimetric perspective. *J. Phys. Oceanogr.*, **36**(1), 87–103.

Jacobs, G. A., and J. L. Mitchell (1996), Ocean circulation variations associated with the Antarctic Circumpolar Wave. *Geophys. Res. Letters*, **23**(21), 2947–2950.

Jacobs, G. A., W. J. Emery, and G. H. Born (1993), Rossby waves in the Pacific Ocean extracted from Geosat altimeter data. *J. Phys. Oceanogr.*, **23**, 1155–1175.

Jacobs, G. A., H. E. Hurlburt, J. C. Kindle, E. J. Metzger, J. O. Mitchell, W. J. Teague, and J. Wallcraft (1994). Decade-scale trans-Pacific propagation and warming effects of an El Niño anomaly. *Nature*, **370**, 360–363.

Kennan, S. C., and P. J. Flament (2000), Observations of a tropical instability vortex. *J. Phys. Oceanogr.*, **30**, 2277–2301.

Killworth, P. D., and J. R. Blundell (2003a), Long extra-tropical planetary wave propagation in the presence of slowly varying mean flow and bottom topography, I: the local problem. *J. Phys. Oceanogr.*, **33**, 784–801.

Killworth, P. D., and J. R. Blundell (2003b), Long extra-tropical planetary wave propagation in the presence of slowly varying mean flow and bottom topography, II: Ray propagation and comparison with observations. *J. Phys. Oceanogr.*, **33**, 802–821.

Killworth, P. D., and J. R. Blundell (2004), The dispersion relation for planetary waves in the presence of mean flow and topography, Part I: Analytical theory and one-dimensional examples. *J. Phys. Oceanogr.*, **34**, 2692–2711.

Killworth, P. D.. and J. R. Blundell (2005). The dispersion relation for planetary waves in the presence of mean flow and topography, Part II: Two-dimensional examples and global results. *J. Phys. Oceanogr.*, **35**, 2110–2133.

Killworth, P. D., D. B. Chelton, and R. de Szoeke (1997), The speed of observed and theoretical long extra-tropical planetary waves. *J. Phys. Oceanogr.*, **27**, 1946–1966.

Killworth, P. D., P. Cipollini, B. M. Uz, and J. R. Blundell (2004), Mechanisms for planetary waves observed in satellite-derived chlorophyll. *J. Geophys. Res.*, **109**(C07002), doi: 10.1029/2003JC001768.

Legeckis, R. (1977), Long waves in the eastern equatorial Pacific Ocean: A view from a geostationary satellite. *Science*, **197**, 1179–1181.

Le Traon, P.-Y., and J.-F. Minster (1993), Sea level variability and semiannual Rossby waves in the South Atlantic subtropical gyre. *J Geophys. Res.*, **98**, 12315-12326.

Madden, R. A., and P. R. Julian (1971), Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708.

Madden, R. A., and P. R. Julian (1994), Observation of the 40–50 day tropical oscillation: A review. *Mon. Weather Rev.*, **122**, 814–837.

Masina, S., and S. G. H. Philander (1999), An analysis of tropical instability waves in a numerical model of the Pacific Ocean, 1: Spatial variability of the waves. *J. Geophys. Res.*, **104**, 29613–29635.

Pedlosky, J. (1987), *Geophysical Fluid Dynamics* (Second Edition, 710 pp.). Springer Verlag.

Philander, S. G. H. (1978), Instabilities of zonal equatorial currents. *J. Geophys. Res.*, **83**, 3679–3682.

Picaut, J., E. Hackert, A. J. Busalacchi, R. Murtugudde, and G. S. E. Lagerloef (2002). Mechanisms of the 1997–1998 El Niño–La Niña, as inferred from space-based observations. *J. Geophys. Res.*, **107**(C5), doi: 10.1029/2001JC000850.

Pollo, I., A. Lazar, B. Rodriguez-Fonseca, and S. Amault (2008), Oceanic Kelvin waves and tropical Atlantic intraseasonal variability, 1: Kelvin wave characterization. *J. Geophys. Res.*, **113**(C07009), doi: 10.1029/2007JC004495.

Qiao, L., and R. H. Weisberg (1995), Tropical instability wave kinematics: Observations from the Tropical Instability Wave Experiment. *J. Geophys. Res.*, **100**, 8677–8693.

Quartly, G. D., P. Cipollini, D. D. Cromwell, and P. G. Challenor (2003), Rossby waves: Synergy in action. *Phil. Trans. Roy. Soc. London A*, **361**, 57–63.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang (2002), An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1625.

Robinson, I. S. (2004), *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Rossby, C. G. (1940), Planetary flow patterns in the atmosphere. *Quart. J. Roy. Meteorol. Soc.*, **66**, 68–87.

Siegel, D. A. (2001), The Rossby rototiller. *Nature*, **409**, 576–577.

Subrahmanyam, B., I. S. Robinson, J. R. Blundell, and P. G. Challenor (2001), Rossby waves in the Indian Ocean from TOPEX/POSEIDON altimeter and model simulations. *Int. J. Remote Sensing*, **22**, 141–167.

Susanto, R. D., Q. Zheng, and X. H. Yan (1998), Complex singular value decomposition analysis of equatorial waves in the Pacific observed by TOPEX/Poseidon altimeter. *J. Atmos. Oceanic Tech.*, **15**(3), 764–774.

Uz, B. M., J. A. Yoder, and V. Osychny (2001), Pumping of nutrients to ocean surface waters by the action of propagating planetary waves. *Nature*, **409**, 567–600.

White, W. B. (1977), Annual forcing of baroclinic long waves in the tropical North Pacific Ocean. *J. Phys. Oceanogr.*, **7**, 50–61.

White, W. B., and R. G. Peterson (1996), An Antarctic circumpolar wave in surface pressure, wind, temperature and sea-ice extent. *Nature*, **380**, 699–702.

White, W. B., C.-T. Tai, and J. DiMento (1990), Annual Rossby wave chracteristics in the California Current region from the Geosat exact repeat mission. *J. Phys. Oceanogr.*, **20**, 1297–1311.

White, W. B., S.-C. Chen, R. J. Allan, and R. C. Stone (2002), Positive feedbacks between the Antarctic Circumpolar Wave and the global El Niño–Southern Oscillation wave. *J. Geophys. Res.*, **107**, 3165, doi: 10.1029/2000JC000581.

Zhang, C. (2005), Madden-Julian Oscillation. *Rev. Geophys.*, **43**(RG2003), doi: 10.1029/2004RG000158.

Zhang, C., and Gottschalck (2002), SST anomalies of ENSO and the Madden-Julian Oscillation in the equatorial Pacific. *J. Climate*, **15**, 2429–2445.

Zheng, Q., R. D. Susanto, X. H. Yan, W. T. Liu, and C. R. Ho (1998), Observation of equatorial Kelvin solitary waves in a slowly varying thermocline. *Nonlin. Processes Geophys.*, **5**, 153–165.

7

Ocean biology from space

7.1 INTRODUCTION

This chapter outlines the ways in which satellite data are being used by ocean biologists. Fifteen years ago, marine biology might have been considered the least likely of all the branches of oceanographic research to benefit from space measurements. In broad terms, biologists are concerned with understanding how particular organisms behave and react to their environment, and it must be admitted at the outset that satellites cannot see individual marine plants and animals.[1] Even the planet's largest animals, the cetaceans, have not been seen in the water from Earth-orbiting spacecraft. However, satellite ocean color data provide a reliable basis for estimating the concentrations of chlorophyll[2] associated with the phytoplankton of the upper ocean. Phytoplankton are a fundamental component of marine ecosystems, the primary producers of organic compounds from solar energy, and the base of the food web in the ocean. Thus measuring them globally and learning about their spatial distribution ought to provide the basis for global and regional ecosystem studies in the ocean. These in turn are of importance for harvesting the ocean's living resources.

Consequently the main contribution of remote sensing to ocean biological science is through the use of ocean color. The reader therefore needs some knowledge of the techniques of ocean color remote sensing (as set out in section 2.4.2 and more fully in chapter 6 of *MTOFS*—Robinson, 2004), which this chapter is intended to complement. However, other techniques in addition to ocean color can also be useful for ocean biology, and one purpose of this chapter is to point out the variety

[1] Although there have been reports that elephant seals basking on beaches can be individually counted using ultra high–resolution imagery from WorldView-1.

[2] Satellite ocean color data specifically provide an estimate of the concentration of chlorophyll-*a*, which is implied throughout this chapter when mention is made of "chlorophyll" or "Chl".

of ways in which satellite data provide marine biologists with unique measurements and novel insights in support of their science. Sometimes this exploits the synergy from combining different remote-sensing techniques. The underlying aim of the chapter is to persuade biological oceanographers to recognize the value of satellite data in the pursuit of their science.

At first this might seem difficult to do. After all, remote sensing offers little to someone interested purely in individual organisms, apart from providing a more complete view of the ocean environment in which the animals and plants live. But when it comes to studying the aggregate behavior of large populations of small organisms, whose distribution is better described in terms of a continuum of concentration rather than discrete individuals, the spatially detailed, wide coverage, and regularly repeated view from space provides precisely the new perspective needed to stimulate original lines of study. As we shall see, the advent of remote sensing has invigorated the study of marine phytoplankton as a field of research. Ocean color data promise the opportunity to characterize the spatial and temporal variability of the structure of phytoplankton communities in studies that link this to higher trophic levels as well as to ocean biogeochemistry (IOCCG, 2008). The availability of ocean color data has encouraged the use of numerical models in marine biology and thus promoted new interdisciplinary interactions between biologists, chemists, and physicists (IOCCG, 2008, chapter 3).

Because ocean color development somewhat lagged behind other remote-sensing methodologies, applications to marine biology are not as mature as those presented in some of the other thematic chapters. Nonetheless there are many examples of applications in the literature and there is room here for only a selective outline. First comes an overview in Section 7.2 of what can be learned about phytoplankton distribution using the latest ocean color measurements from space. Then in Section 7.3 follows an introduction to how primary production can be measured globally with the help of remote sensing. Section 7.4 considers application of satellite data to fisheries. Section 7.5 critically explores whether satellite data can be useful for the study of habitats in shallow tropical seas, while Section 7.6 opens a wider perspective on the use of satellite data for monitoring some environmental stresses on coral reef communities. The concluding section (Section 7.7) envisages what future trends may be. It is worth pointing out that factors of general relevance to ocean biology crop up in most chapters of this book, and there are specific biological applications mentioned in Chapters 5 (upwelling), 6 (do Rossby waves enhance primary production?), 11 (impact of El Niños and monsoons on primary production), 12 (enhancement of production by internal wave mixing), 13 (phytoplankton in shelf seas), and 14 (assimilation of satellite data in ecosystem models).

7.2 PHYTOPLANKTON BLOOMS

7.2.1 An unfolding new view of phytoplankton distribution

The limited quantity of CZCS data available in the early 1980s confirmed the value of ocean color data for those studying phytoplankton (Holligan *et al.*, 1983;

Mitchell, 1994; and associated papers), but it is only since 1997 when global maps of chlorophyll pigment concentrations started to be delivered reliably by SeaWiFS that they started to make a scientific impact. For example, Murtugudde *et al.* (1999) in a study of the whole Pacific and Indian Ocean basins using just the first year of observations were among the first to demonstrate insights to be gained from looking at chlorophyll as an indication of phytoplankton biomass over a wide region, identifying seasonal changes in spatial distribution, and relating both of these to physical oceanography and physical forcing as revealed in global datasets from other satellite sensors. Since 2001, further ocean color data have become available from the MODIS sensors on NASA's EOS Terra and EOS Aqua satellites, and MERIS on ESA's Envisat.

Figure 7.1 shows an example of the quality of information that SeaWiFS data are capable of providing. The image was produced by the author using level 2 Chl data retrieved for the North Atlantic region off Newfoundland from the SeaWiFS overpass on April 21, 2001. Level 2 data (see Table 2.1 and Figure 2.8 to explain the distinction between data-processing levels) are provided at the native resolution of

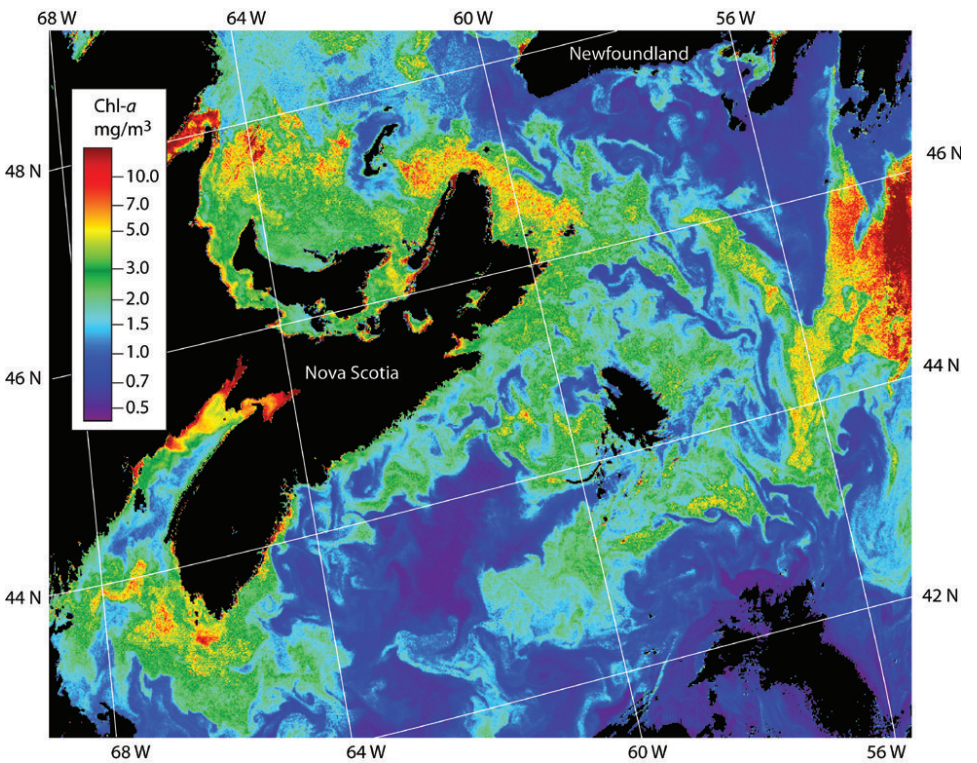


Figure 7.1. Image created from an extract of a level 2 chlorophyll image product from SeaWiFS, showing the spring bloom off Nova Scotia in the northwest Atlantic on April 21, 2001 (level 2 data acquired through the NASA Ocean Color website).

the sensor, 1.1 km at nadir. This image shows the spring bloom occurring west of the Grand Banks. The rich spatial detail shows the complexity and variability of phytoplankton bloom events. Much of the spatial structure is associated with local mesoscale dynamics (as discussed in Chapter 3), and the patchiness is important biological information in its own right. Clear, cloud-free images of blooms such as this are available only occasionally; most overpasses are partly or fully obscured by cloud. Moreover, it covers a relatively small part of the whole ocean. Although it is easy to acquire such data using the NASA Ocean Color website's level 1 and 2 browser to select and download individual files, there is some work involved in displaying HDF files, color-coding the maps, and identifying geographical co-ordinates since the data are in image co-ordinates rather than gridded to latitude and longitude.

It is therefore often more practical to refer, first, to level 3 data products which can be browsed as prerendered global images from the same website, and downloaded either as HDF files for further quantitative analysis or as colored graphics files for qualitative analysis and interpretation. These data are already condensed to 9 km resolution. Figure 7.2 shows an example of global daily images that can be acquired within a few seconds of entering the website. This is as it appears on the screen, with the addition of the title and the standard color scale applicable to all Chl images from SeaWiFS. At first glance this image may seem disappointing since it fails to give a clear view of the whole ocean. That is because it is made up of all the daytime segments of the SeaWiFS swath but these leave large gaps between them, especially at low latitudes. Cloud is also masked as black. However, viewing this image at full resolution on a computer screen (it has $4{,}320 \times 2{,}160$ pixels) gives a better impression than the reproduction in this book and delivers more useful

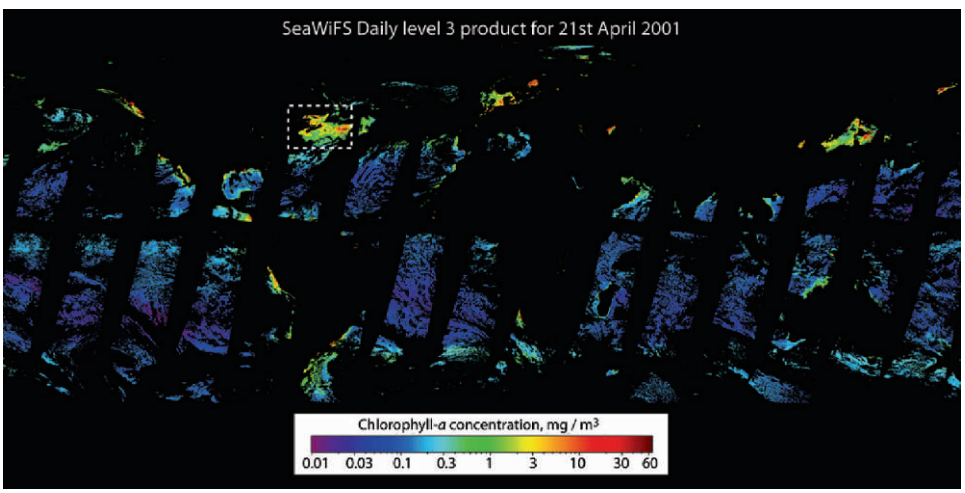


Figure 7.2. Global level 3 daily image showing chlorophyll concentration retrieved from SeaWiFS, for April 21, 2001.

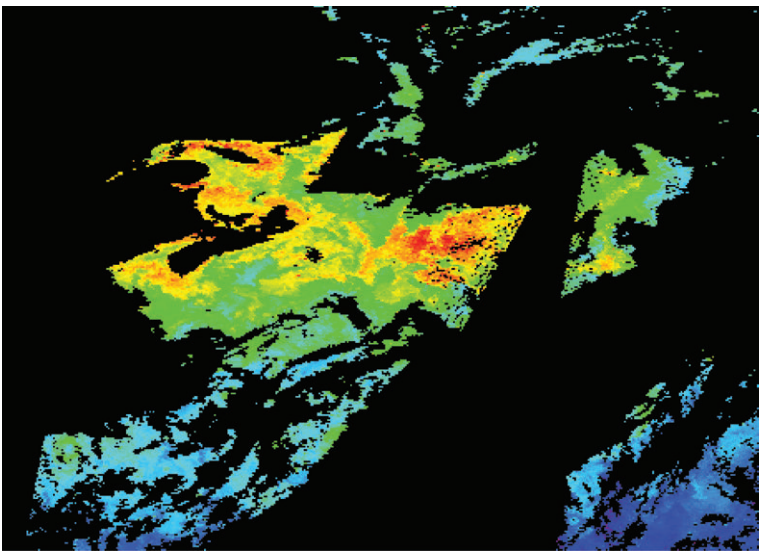


Figure 7.3. Part of global daily, level 3 chlorophyll image from SeaWiFS shown in Figure 7.2, at its full resolution of 9 km pixels (data obtained from NASA Ocean Color website).

information. In practice, images like this provide a quick-and-easy way to discover where there are any interesting cloud-free patches worth exploring with level 2 data. The author took less than 5 minutes browsing daily images online to identify the clear view of a spring bloom to be used to produce Figure 7.1. Figure 7.3 shows what this region looks like at the full resolution of level 3.

In order to build up a broad view of how the distribution of primary production evolves with time it is even more useful to make use of an 8-day composite dataset. By averaging clear pixels of best quality from every overpass during the integration period, data gaps between swaths are eliminated and data loss caused by cloud is reduced. Figure 7.4 provides an example of this for the 8-day period containing the data already presented in Figures 7.1 to 7.3. Although cloud cover still remains it is much easier to see where the main regions of high production are located worldwide during this period. As well as the main upwelling regions, there are spring blooms occurring in the North Atlantic and North Pacific oceans, while some production persists in late autumn on the Patagonian shelf and to a lesser extent throughout the Antarctic Circumpolar Current. When integration is taken a step further, to a monthly composite, most cloud is eliminated (as shown in Figure 7.5), while the overall global picture is similar, apart from some smearing.

Nonetheless, there are things to be learned from the 8-day composites that would be lost by consulting only monthly averages. This is illustrated in Figure 7.6, which shows a sequence of four 8-day composites of the northeast Atlantic Ocean equally spaced over the 9-week period between March 14 and May 8, 2001, and containing part of Figure 7.3. This reveals very clearly the northward progression of the spring bloom during this period, which is not evident in the monthly view. Since the timing of blooms may be just as important as their magnitude when relating primary production to its wider marine biological implications, it is

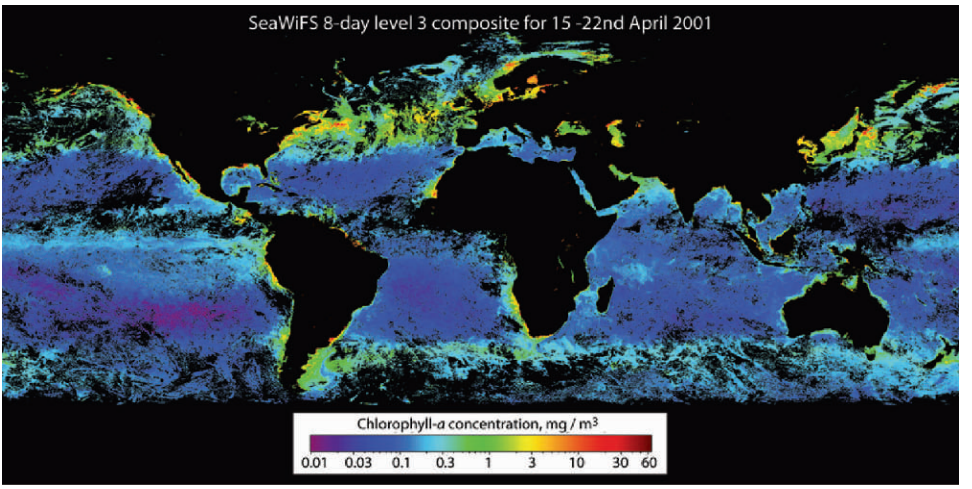


Figure 7.4. Global, level 3, 8-day composite image showing chlorophyll concentration retrieved from SeaWiFS, during the period April 15–22, 2001 (data obtained from NASA Ocean Color website).

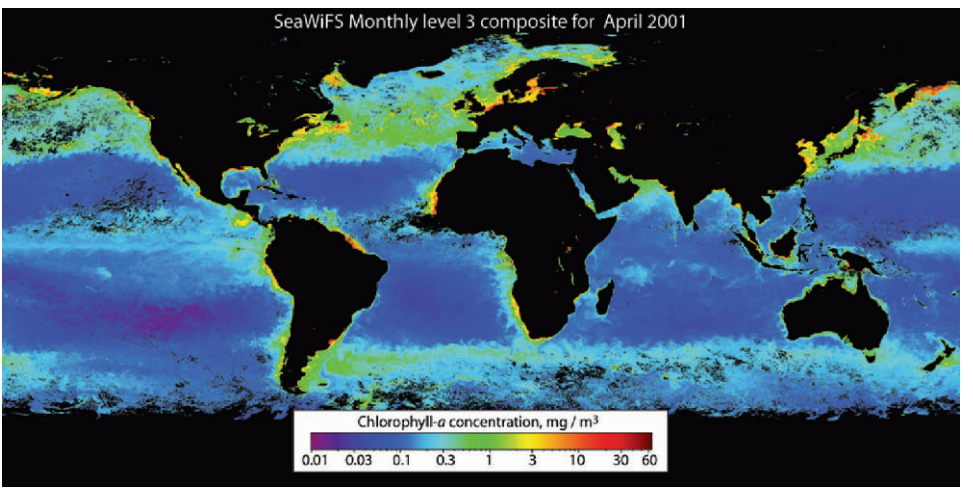


Figure 7.5. Global, level 3, monthly composite image showing chlorophyll concentration retrieved from SeaWiFS, for April 2001 (data obtained from NASA Ocean Color website).

important not to lose information that is available by choosing an inappropriate composite product.

Since 2002 a similar set of ocean color–derived products have become available from MODIS (at the same website as the SeaWiFS data) and from the European Space Agency's MERIS sensor (see Table 2.10 for Internet addresses for access to all these ocean color data sources). Apart from the advantage of more frequent cover-

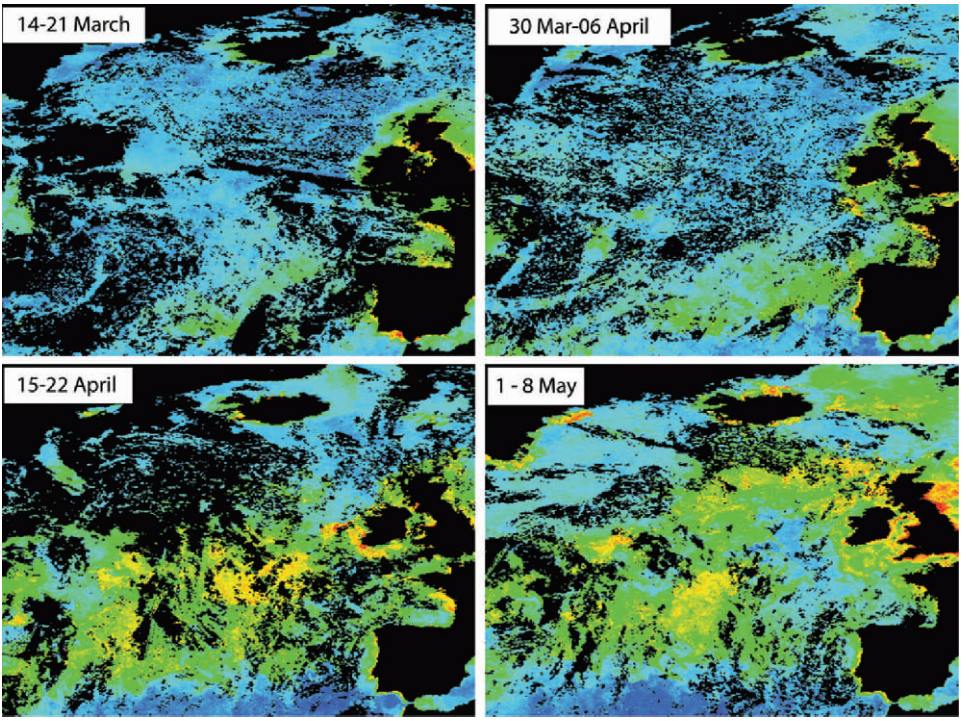


Figure 7.6. Northeast Atlantic extracts from alternate, 8-day, level 3 chlorophyll composites from SeaWiFS showing the northward progression of the spring bloom between March 14 and May 8, 2001 (data obtained from NASA Ocean Color website).

age from these new sensors, which also offer finer spatial and spectral resolution (see table 2.4 and Chapter 6 of *MTOFS*), users have an additional benefit from MODIS, which provides coincident chlorophyll and SST maps derived from the same instrument. This has made it even easier to analyze and interpret the distribution of phytoplankton in relation to physical conditions in the ocean implied by surface temperature. This is demonstrated by a number of the illustrations already used in Chapters 3 to 5 (e.g., Figures 3.3, 3.16–3.19, and 4.3–4.8 relating to mesoscale dynamics, and Figures 5.3, 5.4, 5.9–5.12, and 5.14 relating to coastal upwelling). Without satellite data, it would have been very much more difficult to establish the direct relationship between the spatial distribution of chlorophyll and the location of physical phenomena such as ocean eddies, fronts, or upwelling. In fact without satellite data it is unlikely that we would be aware of the complexity of the patterns of phytoplankton distribution at all. For an authoritative, inspiring, and wide-ranging perspective of the impact which more than a decade of routine ocean color monitoring from space had made on marine science, readers should consult the review by McClain (2009), one of the pioneers of ocean color remote sensing.

The rest of this section considers three aspects of what has been made possible by the wealth of satellite ocean color data now at our disposal. First, we consider the insights gained from the global distribution of chlorophyll and, by implication, of phytoplankton biomass. Next, there is an overview of the research into phytoplankton populations and distribution made possible by the availability of satellite data. Finally, we consider a particular type of phytoplankton, the coccolithophore, that has a characteristic ocean color signature.

7.2.2 The global distribution of chlorophyll

It can be argued that the greatest impact that satellite data are making on the way ocean biologists approach their research is the new global perspective that has been unlocked. The archetypal dataset for illustrating this is the long-term, global, chlorophyll composite of all available data from SeaWiFS or MODIS. Figure 7.7 shows the 10-year cumulative composite from SeaWiFS. The MODIS composite from 5 years of independent data is almost the same. Although any spatial detail of mesoscale variability and all seasonal variability have been eliminated from this image, it is difficult to overstate how important is the information it contains for the whole subject of biological oceanography. At a glance it tells us where the desert regions of the ocean are to be found, reveals the areas where the primary production by phytoplankton appears to be capable of supporting substantial ecosystems, and

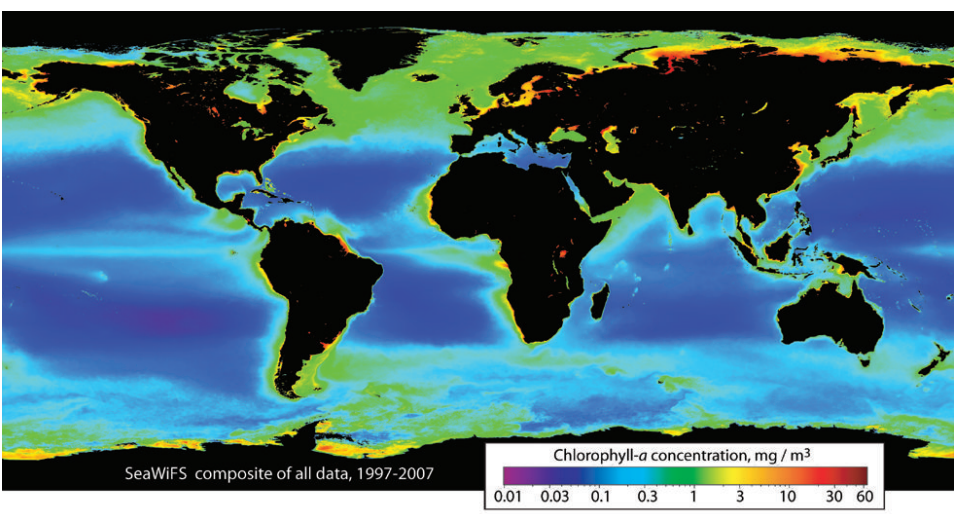


Figure 7.7. Global composite image of all SeaWiFS chlorophyll data acquired from the mission launch in September 1997 until the end of 2007. Color scale as for Figures 7.4 and 7.5 (data obtained from NASA Ocean Color website).

identifies some locations where production appears to be very large.[3] If the earliest biological oceanographers had been presented with such an image a century ago, it could have provided a context in which to make sense of the diversity of marine animals their global explorations discovered.

How can we account for this distribution? Fundamentally, phytoplankton thrive in regions where physical processes are able to supply them with nutrients while keeping them within the upper layer of the ocean where they can harvest sufficient light for photosynthesis. This map marks out regions where these conditions are satisfied. But the distribution is an interesting, almost quirky, combination of general smooth trends and very singular features with definite shapes in precise locations. Recalling that physically the ocean is a turbulent environment whose motion is driven by another turbulent fluid, the atmosphere, we can account for the broad distribution in terms of ocean circulation gyres, fuzzy edges being the consequence of time-averaging phytoplankton distributions dependent on mesoscale turbulent eddies (as mentioned in Chapter 3), or clustered around ocean fronts that are meandering (as shown in Chapter 4). There are other features that are geographically located but still quite smooth, found along the coasts and painting the Equator line on the Pacific Ocean. These are the evidence of coastal and equatorial upwelling (discussed in Chapter 5). However, where very precise patterns have emerged, or been preserved, despite averaging from over 1,000 different images, we must conclude that they are related to sea bed topography or land, which provides the only geographic constancy in a turbulent ocean. For example, in the South Atlantic and South Indian Oceans are some very clear patterns related to sea bed topography (already mentioned in Sections 4.6 and 5.6). The high chlorophyll concentrations in polar regions can be related to special conditions of stable stratification during the melt season for sea ice (as mentioned in Section 5.5), while the small pockmarks of enhanced production dotted apparently randomly around the ocean are found on closer inspection to correspond to isolated islands (see Section 5.4).

Because climatological composites are supplied at a resolution of 9 km (MODIS data come at 4 km) they repay closer inspection at a much finer resolution than this book can provide. As an example Figure 7.8 shows the Pacific Ocean northeast of Australia. This shows up small oases of phytoplankton in what is otherwise low chlorophyll. Mostly these are associated with islands (the land mask is shown in black) but in some cases they appear to be reefs with no, or a very small, island. However, it is also interesting to see there are some islands with no associated production. Previous generations of oceanographers would be amazed at how much information is now available from the archives of freely available satellite ocean color data. But, more than simply providing information, images such as these are capable of provoking questions in the mind of the thoughtful viewer and inspiring new areas of research.

[3] Care is needed when interpreting high chlorophyll in coastal zones since the SeaWiFS OC4v4 algorithm used to retrieve the chlorophyll concentrations shown in this image is not valid where there are land-derived coloring agents in the water. See Section 13.2.6 for a discussion of problems with ocean color product algorithms in coastal seas.

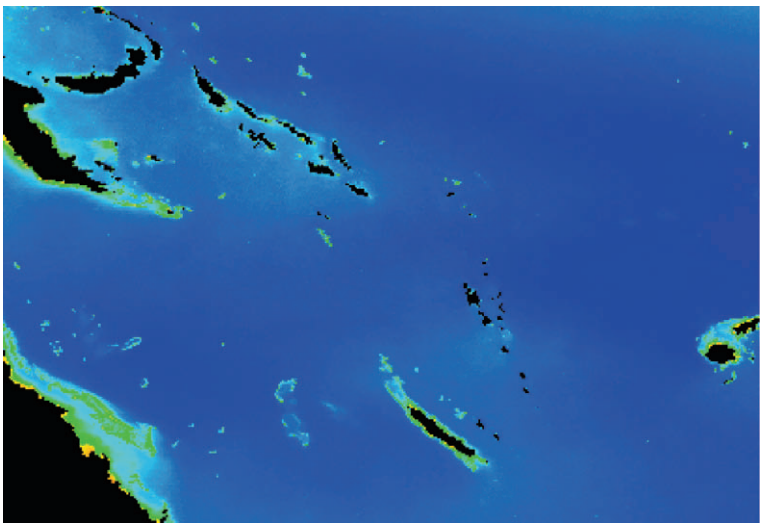


Figure 7.8. Full-scale extract of the Coral Sea from Figure 7.7 to the northeast of Australia, showing the production associated with islands, atolls, and reefs.

The full climatology based on SeaWiFS data divides the year into 46 octets (periods of 8 days), the 46th containing just 5 days and the 8th containing 9 days in a leap year. These intervals define the bins into which data are divided and then averaged over the 10 years (see Section 6.2.1), so that a detailed representation is available of typical, week-by-week variation of chlorophyll. The 8-day climatology still contains a noticeable amount of data dropout due to cloud, and in this respect monthly averaged climatologies are less corrupted. Figure 7.9 shows a section of the

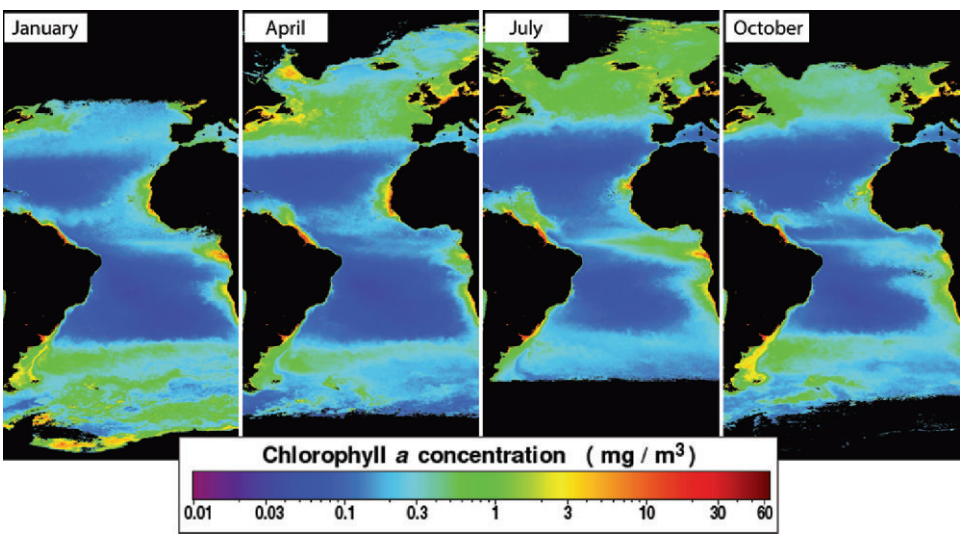


Figure 7.9. Atlantic Ocean extracts from four monthly climatologies of chlorophyll derived from SeaWiFS data between 1997 and 2007 (data obtained from NASA Ocean Color website).

monthly SeaWiFS climatology containing the Atlantic Ocean, for the months of January, April, July, and October. Seasonal migration of the centers of production can clearly be seen, particularly those associated with the spring bloom at mid to high latitudes, but some regions such as upwelling centers show little seasonal change. In the tropics, complex changes during the year reflect the variability of wind-driven equatorial currents and countercurrents, including the traces of the Amazon discharge which is discussed in Section 5.3.

There is a growing interest in the science of marine phenology, the study of the timing of events such as the spring bloom, because it provides a means of monitoring the impact of climate change on marine ecosystems (see, e.g., Edwards and Richardson, 2004; Richardson and Schoeman, 2004). The early or late development of a seasonal phenomenon needs to be understood in its spatial context, since localized observations based, for example, on a fixed mooring site can be misleading. Given the ever-extending timespan from which satellite-derived chlorophyll observations are available, and the capacity of these data to resolve timing of events to within a few days, there is potential for systematic study of phytoplankton phenology. The combination of satellite ocean color and SST data permits both physical forcing and the response of primary production to be related to *in situ* observations of the phenology of zooplankton and fisheries. A few examples of this kind are mentioned later in the chapter.

To complete this overview of global-scale information now available to all oceanographers from ocean color data, mention should be made of anomaly images that can be produced. As explained in Section 6.2.1, once-monthly or 8-day climatologies have been generated. Individual monthly or 8-day composites for a given date can be compared with the corresponding climatology to create an anomaly image. Figure 7.10 shows an example for April 2001. It has been produced by subtracting the global April climatology (represented numerically in its $\log_{10}$ Chl-*a* form) from the data presented in Figure 7.5 (also in its $\log_{10}$ Chl-*a* form), and then extracting the North Atlantic region. Justification for evaluating the chlorophyll anomaly in this way is that the concentration of chlorophyll in the ocean is approximately log-normally distributed (Campbell, 1995). The absolute values of chlorophyll span more than four orders of magnitude and if the anomaly were evaluated from absolute values it would be difficult to interpret. By basing the anomaly on the logarithm of chlorophyll, the inverse logarithm of the result represents the ratio between the actual value and the climatology (i.e., the lower scale shown in Figure 7.10). As with many anomaly scales, the central value is colored white and corresponds to the zero $\log_{10}$ anomaly where the current observation is the same as the climatology. Orange and red correspond to positive log anomalies, where the current Chl is higher than the climatology. Blue and purple are negative log anomalies, where the current value is smaller than the climatology.

Anomaly maps need to be interpreted carefully. In this case, where the map has white or pale colors, there is very little difference from the typical chlorophyll concentration at that location for that time of the year. In this North Atlantic extract, there are both large positive and large negative anomalies, denoting significant differences from average chlorophyll distribution. These could be the result of a

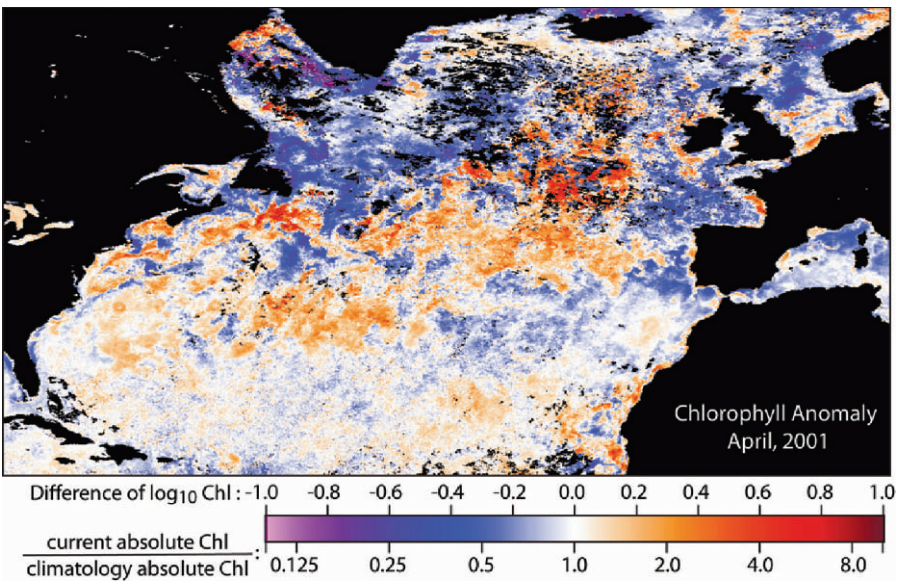


Figure 7.10. North Atlantic extract showing the monthly chlorophyll anomaly for April 2001, based on the difference between the $\log_{10}$ values of Chl for the current monthly composite and for climatology. The upper scale shows the difference in $\log_{10}$ Chl, the lower scale represents this as the ratio between the actual Chl and the climatological Chl at each location (created by the author using monthly composite and climatology data obtained from NASA Ocean Color website).

different timing of the spring bloom. They could also be partly related to a displacement of a major front such as the Gulf Stream which creates strong positive and negative anomalies close to each other. A fuller interpretation calls for careful analysis of previous and subsequent months and the plotting of time series of chlorophyll at particular locations for comparison with the mean climatological annual cycle at that location. The value of a single anomaly image, as displayed here, is that it highlights locations of departures from normal, where a more thorough analysis of the evolving time series of images is required.

Finally, having drawn extensively in this section from data that are freely available from the NASA Ocean Color website,[4] it is worth pointing out that the whole range of different composite and climatology image datasets are available as pictures or digital data for other level 3 products in addition to chlorophyll, including normalized, water-leaving radiance and some atmospheric parameters. A growing number of global demonstration products from the MERIS sensor, including chlorophyll and water-leaving radiances, is also available online from ESA as prerendered images.[5]

[4] *http://oceancolor.gsfc.nasa.gov/*
[5] *http://envisat.esa.int/level3/meris/*

7.2.3 Scientific exploitation of satellite ocean color data

Access to Coastal Zone Color Scanner (CZCS) data (1978–1986) stimulated scientific examination of the spatial distributions and growth of phytoplankton populations in order to characterize different types of behavior: for example, in the Gulf of Mexico and Caribbean (Gonzalez *et al.*, 2000), the U.S. East Coast (Yoder *et al.*, 2001), and the California Current (Thomas and Strub, 2001). It also led to more forward-thinking about the role for operational ocean color sensors. For example, Erickson and Eaton (1993) demonstrated the potential for assimilating ocean color data into ocean general circulation models supplemented with the capacity for modeling biogeochemical processes.

The nature of the CZCS dataset restricted investigations to fairly broad analysis of seasonal patterns because sampling frequency was generally insufficient to follow individual blooms. The advent of SeaWiFS data changed that, enabling a much more detailed definition of the temporal evolution of populations. For example, a sequence of SeaWiFS data over the Celtic Sea (Joint and Groom, 2000) showed clearly the appearance and decline of the spring bloom, with chlorophyll concentrations greater than $5\,mg/m^3$ indicated in some patches in March and April, falling to less than $0.5\,mg/m^3$ in August, although patches of continued production were identified where nutrient supply was maintained by processes, such as island stirring or mixing at shelf–sea tidal fronts. Thus SeaWiFS data made it possible to track the life history of a particular phytoplankton population through the spring bloom and beyond.

Kahru *et al.* (2000) used the SeaWiFS time series from its inauguration in September 1997 to measure changes in chlorophyll concentrations in the California Current system associated with the El Niño event of 1997–1998. The El Niño caused a decrease in biomass off the Californian coast, but an increase farther south off Baja California. Their ability to track changes in spatial structure as well as the magnitude of chlorophyll led to a much more informed conclusion than if they had relied on point measurements. Interestingly they were able to reanalyze the CZCS record and found a similar pattern during the 1982–1983 El Niño.

In the Alboran Sea, close to the western opening of the Mediterranean Sea, Garcia-Gorriz and Carr (1999) used the continuity provided by OCTS and SeaWiFS data to supplement previous observations with CZCS in order to draw climatological conclusions about phytoplankton distribution. Unlike other parts of the western Mediterranean, where a significant spring bloom occurs, the annual cycle here consists of two regimes: a bloom period during autumn and winter (November–March) and a nonbloom period during summer (May–September). They were able to relate this behavior to circulation and upwelling processes as the causes. More recently, analysis of the phytoplankton distribution pattern throughout the whole Mediterranean Sea using 6 years of SeaWiFS data has identified some longer term trends (Barale *et al.*, 2008). Lévy (2005) performed a similar analysis of production regimes in the northeast Atlantic.

The wide-scale coverage by SeaWiFS has enabled questions about the basin-scale distribution of phytoplankton, and its seasonal variability, to be addressed.

Thus Signorini *et al.* (1999) were able to examine the variability of sea surface chlorophyll concentration in the tropical and subtropical Atlantic using the first year of SeaWiFS imagery. In their comprehensive approach to understanding observed fluctuations of chlorophyll they also used an ocean circulation model in conjunction with dynamic heights from TOPEX/Poseidon and gridded wind stress data. They found regions of high surface chlorophyll that correlated with mesoscale and large-scale physical processes such as the strong upwelling off the west coast of Africa, relatively high oceanic production within the Guinea Dome region, and generation and propagation of anticyclonic eddies along the coast of South America, north of the Equator. Major river outflows such as the Amazon, Orinoco, and Congo also had strong signatures where chlorophyll appeared to be very high, although this measurement may have been affected by suspended particulates leading to Case 2 conditions. Interestingly they found that the expected autumn bloom was absent, perhaps the consequence of ENSO conditions at that time. A similar approach was adopted by Murtugudde *et al.* (1999) in a study of the tropical Indo-Pacific basin during the same period. This provided a much more comprehensive overview than was previously possible about the impact of El Niños on chlorophyll concentrations, showing not only reduced chlorophyll production where upwelling is switched off but its enhancement in other parts of the ocean. Remote sensing of the El Niño phenomenon is discussed further in Section 11.2.

SeaWiFS data have been used regularly in conjunction with field experiments studying phytoplankton distribution and primary production. In the Southern Ocean, as part of the U.S. JGOFS experiment, elevated chlorophyll levels were found to be associated with the Pacific–Antarctic Ridge, a surprising result since the crest of the ridge is never shallower than 2,000 m deep (Moore *et al.*, 1999). In the Bay of Bengal the seasonal pattern of chlorophyll distribution has been related to river discharges varying with monsoons (Gomes *et al.*, 2000). Off Hawaii, SeaWiFS data contributed to experiments examining interannual phytoplankton variability and its relation to mesoscale physical processes in the North Pacific Central Gyre (Leonard *et al.*, 2001). In the Irminger Basin of the northern Atlantic the timing of the spring bloom has been related to particular meteorological conditions each year (Henson, *et al.*, 2006). Farther south in the subpolar North Atlantic, the strength of blooms appears to be modulated by the wind (Ueyama and Monger, 2005). Interestingly, there have also been cases where the detection from satellite data of previously unexpected blooms in oligotrophic waters may act as a stimulus for new fieldwork plans to learn more about these puzzling events. One example is in the oligotrophic North Pacific (Wilson, 2003).

In this summary of research stimulated by the availability of satellite ocean colour data, the work of Platt and Sathyendranath (1999) serves as an example of how the theoretical framework for understanding pelagic ecosystems is also evolving in response to remote-sensing observations. They suggested that the appropriate way to define spatial structure in the oceanic ecosystem is in terms of dominant ecophysiological rate parameters characterizing autotrophic production. They argued that in this regard the ocean can be segmented into regions or *provinces* within which rate parameters are almost uniform, separated by boundary zones across which they vary

significantly, normally in response to physical forcing. While this is broadly consistent with previous ideas (Longhurst, 1998), they made the particular point that boundaries between provinces must be considered as dynamic partitions, and they proposed that ocean color measurements hold the key to defining their position at any particular time. They drew attention to the potential for using satellite chlorophyll composite datasets, averaged over different timescales of 8 days, monthly, and seasonally, to help identify variability in biological boundaries. It will be shown in Section 7.3.2 how knowledge of boundary positions can be important for estimation of ocean primary production.

An example of this approach, applied in the confluence zone of the Brazil and Malvinas currents (Saraceno *et al.*, 2004, 2005), is also mentioned in Sections 3.4.4 and 4.6.2 which note that other satellite-derived ocean datasets of temperature and eddy kinetic energy also contribute to distinguishing between different biogeographical ocean provinces. More recently a combination of multivariate statistics and classification techniques has been applied to a time series of satellite-derived, surface–ocean chlorophyll data from SeaWiFS (Hardman-Mountford *et al.*, 2008), leading to objective characterization and classification of ecological patterns over the whole ocean. This study also investigated the characteristic system properties of the broad-scale patterns observed to test whether the identified provinces behave as autonomous ecological systems. Analyses such as this provide evidence of new directions that marine ecologists are now able to take. Such work could not have been addressed without the availability of extended time series of global, fine-resolution maps of properties derived from ocean color and other satellite sensors.

Alongside the development of this systems approach to representing marine ecosystems has come the improvement of numerical ecosystem models which allow wider biogeochemical processes to be explored. Satellite ocean color data can contribute here as well. For example, SeaWiFs data have been used for calibrating or tuning coefficients used in numerical ecosystem models (Hemmings *et al.*, 2003, 2004, 2008). In comparison with use of a limited number of *in situ* observations for model parameter tuning, satellite data provide a much wider range of conditions and a broader spread of variables with which to confront a numerical model, testing it more thoroughly and leading to its refinement, as discussed further in Section 14.3.

Biogeochemical model development is also driving the move to partition phytoplankton into different functional types, each with a characteristic role in ocean biogeochemical cycling. Research is therefore being directed to find ways, if possible, of analyzing ocean color data in order to partition chlorophyll retrievals into different functional groups. Rather than differentiation by size, the preferred goal would be to distinguish between phytoplankton that are predominantly nitrogen fixers, silicifiers, calcifiers, or producers of dimethyl sulfide (DMS), since each of these functions has an impact on different biogeochemical fluxes (Nair *et al.*, 2008). Alvain *et al.* (2005) used an empirical approach to identify different groups from their spectral characteristics and Aiken (2007) used MERIS data to distinguish between different phytoplankton types in the Benguela ecosystem. Retrieval of inherent optical properties (IOPs) of the water from ocean color data (IOCCG, 2006) is also offering a promising step towards the goal of distinguishing functional groups.

Certain specific species or groups have distinctive spectra which may be readily identified. Coccolithophores, the main type of calcifier, and also a DMS producer, will be discussed in the next section. The cyanobacterium *Trichodesmium*, a nitrogen fixer, has been detected in SeaWiFS data (Subramaniam *et al.*, 2002), while algorithms have been proposed to distinguish diatoms in ocean color data (Sathyendranath *et al.*, 2004).

7.2.4 Coccolithophores

The particular area of phytoplankton research that seems to have most energetically adopted the use of satellite data is the study of coccolithophores, in particular the species *Emiliania huxleyi*. These produce external plates, called coccoliths, made of calcium carbonate, which make them highly reflective across all visible wavelengths. As a bloom develops and individual cells die, they leave behind free liths which are even more reflective. Consequently the color signature of coccolithophores is very different from other phytoplankton species (Balch *et al.*, 1991). There is strong backscatter even in the blue part of the spectrum, giving the sea a bright turquoise appearance. This has made them much easier to detect qualitatively (see Figure 7.11), not only from ocean color sensors but also using the broad-band, visible channel of AVHRR and the high-resolution Landsat that lacks the radiometric resolution to distinguish the subtle green coloration of the sea caused by most phytoplankton. Thus coccolithophores have been detected during the last 30 years, even when no ocean color sensors were operating. Consequently their global distribution has been mapped pretty thoroughly (Brown and Yoder, 1994; Brown and Podestá, 1997; Tyrell *et al.*, 1999).

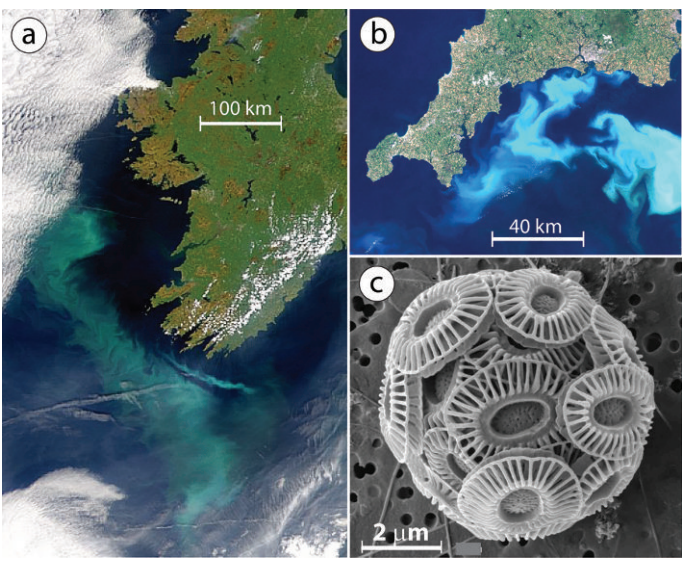


Figure 7.11. Images of coccolithophores. (a) SeaWiFS near real–color composite image of bloom along the continental shelf edge southwest of Ireland on May 10, 2000. (b) Landsat Advanced Thematic Mapper false-color composite image of bloom in the English Channel south of Cornwall, U.K. on July 24, 1999. (c) Electron microscope image of a single coccolithophore cell encased in coccoliths.

The very large blooms detected from satellites, extending for hundreds of kilometers, appear to be formed only by the species *E. huxleyi* and *Gephyrocapsa oceanica* (Iglesias-Rodríguez *et al.*, 2002). Inclusion of the electron microscope inset in Figure 7.11 is a reminder of the range of spatial scales spanned by the biology of *E. huxleyi*. Biologists need microscope images with a resolution of 10 nm to define the organism, and images with a coverage of 1,000 km to define the extent of the largest blooms, thus spanning an incredible 14 orders of magnitude! It is also fascinating to note how *E. huxleyi* cells generate liths internally, one at a time, and then push them outside the cell to form the protective structure seen in the microscope image, a feat of nanoengineering to match the technology that has placed in orbit the sensors that provide the satellite images!

Despite their ready visibility from space, quantifying the concentration of live cell numbers in coccolithophore blooms is difficult using ocean color satellite data. Because of enhanced scattering across the visible spectrum, the standard blue–green ratio algorithm for retrieving chlorophyll concentration does not work well. Overall brightness detected at different spectral bands may be used as a measure of the concentration of coccoliths themselves (Gordon *et al.*, 2001), but cannot distinguish readily between places where there are many live cells with attached liths and places where the cells have died and the liths have detached but remain in suspension.

Care must be taken not to assume automatically that any patch of bright reflectance in the ocean must be caused by coccolithophores since other phenomena can generate a similar signature in visible wavelength remote sensing. For example, off the Namibian coast extensive eutrophication events have caused seafloor release of hydrogen sulfide which forms a precipitate in surface waters that looks like a coccolithophore bloom (Weeks *et al.*, 2002). In the Bering Sea, some apparent blooms of *E. huxleyi* seen in winter satellite images turned out following analysis of *in situ* sampling to be empty diatom frustules that had been resuspended from the seabed by strong winds (Broerse *et al.*, 2003). Nonetheless summer blooms of coccolithophores do occur in the Bering Sea and other high-latitude regions.

Given increasing ocean acidification accompanying rising concentrations of atmospheric CO_2, it is expected that coccolithophores may start to decline, which would result in a reduction of their global role in calcification and sequestering of oceanic CO_2 (Riebesell *et al.*, 2000). However, there is recent laboratory evidence to suggest that coccolithophores may have a more complex and resilient response to changes in CO_2 and pH (Iglesias-Rodríguez *et al.*, 2008). The capacity to globally monitor the frequency and extent of coccolithophore blooms from satellites will therefore be of continuing importance during the next few decades.

7.3 PRIMARY PRODUCTION

7.3.1 Theoretical background

Primary production occurs when algal cells capture radiant energy and then convert it into chemical energy which is stored in the algal biomass by the production of new

plant cells. At the molecular level this is accomplished by the process of photosynthesis, through pigments such as chlorophyll-*a* and others more fully described by Kirk (1994). Care must be taken when discussing primary production in the context of satellite oceanography. Remote-sensing methods measure chlorophyll concentration per unit volume near the surface of the ocean. They do not measure how that concentration varies with depth. Neither can they readily determine the chlorophyll-to-carbon ratio, and so they cannot directly measure total phytoplankton biomass in the water column per unit surface area, although chlorophyll maps may be approximately interpreted as indicators of algal biomass distribution. Sometimes maps of chlorophyll distribution are loosely referred to in terms of primary production, but they must never be confused with the true measurement of production. As this section will show, using satellite data to measure primary production is not straightforward and requires more than a simple retrieval of chlorophyll concentration from ocean color.

The term primary production strictly refers to an energy exchange process, and its measurement is normally expressed in the quantity P, the primary production rate, although sometimes it is represented by PSR, the photosynthetically stored radiation rate. P is a measure of the rate at which carbon is fixed by photosynthesis, per unit time per unit volume. Units are $\mathrm{g\,C/(m^3\,s)}$. In some contexts, total production, P_{tot}, is integrated over depth through the water column and expressed per unit surface area, the units being $\mathrm{g\,C/(m^2\,s)}$. When integrated globally over the World Ocean, or over specific seas, units are g C/s, but usually expressed as Gt C/yr (note that $1\,\mathrm{Gt\,C/yr} \approx 31.6 \times 10^6\,\mathrm{g\,C/s}$).

PSR is a measure of the energy photosynthetically stored as chemical energy within the plant biomass throughout the water column per unit time per unit surface area. It has the units of $\mathrm{W/m^2}$. However, since the rate is typically averaged over a day, it is often reported as $\mathrm{J/(m^2\,day)}$.

The dependence of P on other environmental variables is based on the instantaneous growth rate equation, in which the rate of carbon fixation per unit volume per unit wavelength of light at depth z and wavelength λ is given (Kiefer and Mitchell, 1983) as:

$$P(z,\lambda) = 12C(z)a_c^*(z,\lambda)E_{\mathrm{PAR}}(z,\lambda)\phi_\mu(z,\lambda), \tag{7.1}$$

where C is chlorophyll concentration ($\mathrm{g\,Chl/m^3}$); a_c^* is the chlorophyll-specific light absorption coefficient ($\mathrm{m^2/g\,Chl}$); E_{PAR} is photosynthetically available radiation (sometimes called photosynthetically active radiation) within the visible waveband (400 to 700 nm); and ϕ_μ is quantum yield. It is important to note that most variables in Equation (7.1) vary with z and λ.

E_{PAR} is photon flux density, a measure of photons per unit area per unit time, and therefore it is expressed in quantum units (mol Q which are the same as Einsteins), so the units of E_{PAR} are $\mathrm{Ein/(m^2\,s)}$. The energy of a photon is hf, or hc/λ, where h is Planck's constant; c is the speed of light; and f is the frequency of the light. Therefore, photon flux can be calculated from optical energy flux (i.e., scalar irradiance E_0), but the spectral dependence of E_0 needs to be known since

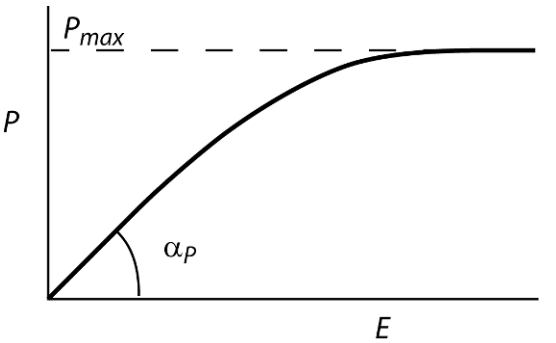


Figure 7.12. Typical variation of production rate, P, with irradiance, E.

there are fewer photons per unit of energy in the blue than in the red part of the spectrum. Section 7.3.3 considers how PAR is measured from satellites.

In Equation (7.1) we can consider $C.a_c^*$ as the probability that a photon entering the sea will encounter a chlorophyll molecule per unit distance along its path. Therefore the rate of photons encountering chlorophyll molecules is $Ca_c^*E_{\mathrm{PAR}}$ per unit volume. Only a proportion, ϕ_μ, of these photons is actually engaged in the photosynthesis process and converted into chemical energy, leading to capture of a carbon molecule from the water and its conversion into "phytoplankton carbon". This conversion as a proportion of available photons is represented by the quantum yield term in (7.1) the units of which are mol C/Ein. The constant of 12 in (7.1) comes from the molecular weight of carbon.

Quantum yield varies with the light level according to the empirically observed P–E curve illustrated in Figure 7.12. At low light levels P increases linearly with illumination, the initial slope of the graph, $\tan\alpha_P$, being the maximum quantum yield. However, this gradually reduces and approaches a saturation value $P_{\max}$ for high light levels. Particular values of α_P and $P_{\max}$ can vary with λ, with temperature, with the availability of nutrients required for cell growth, and with species and other aspects of phytoplankton population. Thus, evaluation of (7.1) is not straightforward.

If the total production rate over the water column, P_{tot}, is to be calculated, Equation (7.1) must be integrated with respect to depth through the euphotic zone and with wavelength across the spectrum, bearing in mind that $C(z)$ may be layered or continuously variable rather than uniform, and E may not have a simple functional dependence on depth. A lot of work has been done to explore the dependence of column-integrated production on different types of vertical distribution of C and to reduce them to a form based on readily measurable optical and biochemical properties (e.g., Platt *et al.*, 1988; Morel, 1991; Platt and Sathyendranath, 1993). It is worth pointing out that less than 2% of E_{PAR} entering the sea surface is ever utilized for primary production, and in oligotrophic waters this may be as small as 0.02%.

One other quantity of interest, especially in relation to the role of the ocean in global carbon budget, is that of *new production* (Platt *et al.*, 1992). Much carbon fixed by primary production is derived from the detritus and degradation products of

dead phytoplankton cells, and so is no more than a recycling of carbon previously fixed by the same population. Any additional production in excess of local community metabolism is described as new production, P_{new}. The ratio of P_{new} to P_{tot} is called the f-ratio.

7.3.2 Methods for estimating production from remote sensing

The fact that one of the terms in Equation (7.1) is C, the concentration of chlorophyll, gives a reason to look for ways to evaluate primary production from satellite data. However, the additional terms in the equation, plus the need to integrate down to depths which ocean color sensors cannot penetrate, means that primary production rates should not be expected to bear a simple relationship to ocean color remote-sensing measurements.

Empirical models

Nonetheless surprising success was gained initially by using simple empirical models to relate chlorophyll to primary production (Behrenfeld and Falkowski, 1997a; Joint and Groom, 2000). At their simplest these models demonstrate a relationship between primary production and chlorophyll within a given set of *in situ* observations. Typically a linear relation is sought between the logarithms of P_{tot} and C. If this relationship is sufficiently accurate, robust, and stable, it can be applied to satellite-derived chlorophyll in order to estimate primary production. There is scope to use other satellite data to refine such models. For example, since the availability of nutrients is an important factor for the physiological response of plankton to the availability of light (ϕ_μ in Equation 7.1) and because nutrients are often supplied through upwelling of deeper, cooler water, then satellite-derived SST may be included in the empirical model scheme to reflect the effect of nutrient availability on production rates.

When such empirical algorithms are successful it is because, for the sample data used to derive the relationship, there must be little variation of the vertical structure of C, little change in averaged properties of E_{PAR}, and consequently also insufficient variation in ϕ_μ that is independent of C to prevent a simple relationship being established. Nonetheless, although this approach may be satisfactory for using remote sensing of ocean color to interpolate in space and time between a series of *in situ* measurements of production within a limited region, it cannot be relied upon to extrapolate to locations or times for which no calibration data are available to tune the P_{tot} *vs.* C relationship. It certainly cannot be applied globally with any confidence.

Process-based models

More robust methods for utilizing satellite data in estimating primary production, which have credibility without reliance on coincident *in situ* production measurements, have adopted a physical or semiempirical modeling approach based on (7.1). Such methods analyze a large number of *in situ*, direct measurements of primary

production along with as detailed a knowledge as possible of the vertical distribution of C, the spectral form of the optical absorption coefficient, and the nature of the P *vs.* E curve, from a wide variety of locations and seasons. By running a forward model based on (7.1) they have explored the various dependencies of equation variables on measurable quantities, and have derived semiempirical relationships to represent that dependency. Since the form of semiempirical relationships is physically based, a more sound foundation is provided for applying resulting models across whole ocean basins and throughout the year. Within this broad approach two slightly different methods have developed, which can be labeled by their distinctive element as the "cross-section for photosynthesis" method (Morel, 1991; Antoine and Morel, 1996) and the "biogeographical provinces" method (Longhurst *et al.*, 1995; Sathyendranath *et al.*, 1995; Platt and Sathyendranath, 1999).

Cross-section for photosynthesis

Morel and co-workers adopted what they term the ψ^* approach (Antoine and Morel, 1996). This represents the integrated form of (7.1) as

$$P_{\text{tot}} = (1/J_C)\langle C\rangle_{\text{tot}} E_{\text{PAR}}(0^+)\psi^*, \tag{7.2}$$

where $\langle C\rangle_{\text{tot}}$ is column-integrated chlorophyll per unit surface area; $E_{\text{PAR}}(0^+)$ is photosynthetically available irradiance just above the sea surface; $J_C(= \text{PSR}/P)$ is the chemical energy equivalent of a unit mass of carbon fixed (in kJ/g C); and ψ^* has the units of $\text{m}^2/\text{g Chl}$ and so can be thought of as a cross-section for photosynthesis per unit of chlorophyll. ψ^* contains all the remaining variability arising from (7.1). Thus it represents not only the underwater optical processes inherent in (7.1), including the role of a_c^* and the effect of nonlinear interactions between the separate depth dependences of C and E_{PAR}, but also biochemical factors inherent in the P *vs.* E relationship. The factor ψ^* therefore varies according to a wide variety of environmental variables, including those that affect the light field such as Sun angle, sea state, and cloud cover, the shape of the vertical distribution of chlorophyll, and the photochemical response of the particular algal population (Morel and Berthon, 1989). Some slight simplification is achieved by evaluating ψ^* for integrations over a whole day.

The rationale for this approach is that the variables $\langle C\rangle_{\text{tot}}$ and $E_{\text{PAR}}(0^+)$ are recoverable from space, so (7.2) can be used to estimate primary production from satellite data, as long as ψ^* can be evaluated. $\langle C\rangle_{\text{tot}}$ is based on the chlorophyll concentration retrieved from satellite ocean color, such as the data illustrated in Section 7.2.2. $E_{\text{PAR}}(0^+)$ is also retrieved from satellite data (as discussed later in this section). The value of ψ^* required at a particular place and time can be determined by running a full spectral model of photosynthesis (Morel, 1991) with the appropriate vertical distribution of chlorophyll and other environmental variables. Then ψ^* could be treated as a climatological field whose space-time distribution is to be determined. In practice, lookup tables of ψ^* have been created, which are entered by specifying date, latitude, cloudiness, temperature, and remotely sensed chlorophyll concentration. Interpolation is possible between the precalculated grid of

values in the lookup table. Separate tables are used for well-mixed (uniform depth distribution of C) and stratified upper layers. In the latter, the profile of C with a subsurface maximum is based on statistical analysis of measured profiles. In this way the fairly complex sensitivity of production to a variety of environmental parameters is made manageable.

Biogeographical provinces

The approach adopted by Platt and Sathyendranath (1988) uses the concept of *biogeographical provinces*. The photosynthetic rate parameters α_P and $P_{\max}$ (as defined in Figure 7.12) and the vertical structure of C are explicitly specified for a given local region at a given time. The local production rate is then obtained by full integration of (7.1) over depth and wavelength (Sathyendranath *et al.*, 1989) using remote-sensing input to specify variability of the magnitude of C and E_{PAR} over spacescales smaller than the size of the local region and over timescales shorter than the season for which other parameters are assumed to be constant. The depth distribution of C is parameterized following (Platt *et al.*, 1988) as a Gaussian form

$$C(z) = C_0 + \frac{h_C}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(z - z_m)^2}{2\sigma^2}\right], \tag{7.3}$$

where C_0 is background pigment concentration in mg/m^3; h_C (in mg/m^2) determines the amplitude of the vertical structure; z_m defines the depth of the pigment concentration peak; and σ defines its width as indicated in Figure 7.13. Because ocean color data provide a depth-averaged value of C, only three of the four parameters used in (7.3) need to be specified.

The approach is extended to wider regions and other seasons by making the assumption that vertical structure parameters are stable and photosynthetic rate parameters can be considered constant during the whole of a season of the year, and across a reasonably large geographical region. From analysis of field measurements of photosynthetic rate and biomass structure parameters, it can be shown that these parameters in general vary slowly, except across boundaries between regions of significantly different oceanographic type. Thus an ocean basin can be segmented into a relatively small number of biogeographical prov-

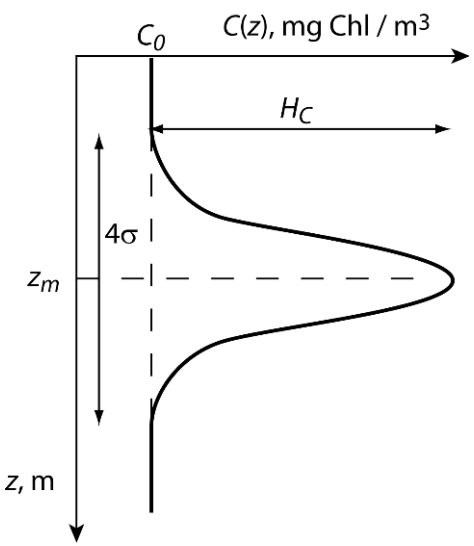


Figure 7.13. Idealized structure of the vertical distribution of biomass in the upper ocean, used in models of primary production. $H_C = h_C/\sigma\sqrt{(2\pi)}$.

inces, within which they are taken to be uniform for a particular season. Provinces differentiate between ocean boundary currents, ocean gyres, upwelling regions, and shelf seas and between different gyres and distinct coastal areas. Although their definition continues to evolve with further empirical observations (Brock *et al.*, 1998) and theoretical considerations (Platt and Sathyendranath, 1999), the approach has enabled global estimates of production to be made (Longhurst *et al.*, 1995).

New production

Both the above approaches can in principle be extended further to estimate new production, P_{new}. Sathyendranath *et al.* (1991) achieve this by estimating the new production ratio f from which P_{tot} is evaluated. Although it would be attractive to produce climatologies of f in the same way as ψ^*, or else determine f to match each biogeographical province and hence arrive at global estimates of P_{new}, this is not likely to be satisfactory. The processes controlling f are different from those that determine the parameters for each province. Moreover, even if an average f could be established, it is likely that events such as upwelling, which cause a burst of new production to occur, may be associated with anomalous values of f.

Whereas regenerated production simply uses up the nitrogen made available by recycled material in the photic zone, new production requires additional nitrogen to be supplied from outside the photic zone, such as vertical flux from below. It should therefore be possible to find a functional dependence of f on N, the nitrate concentration, as confirmed by Elskens *et al.* (1999). In the context of algal production in the upper ocean, the concentration of N can be related inversely to temperature, since upwelling cooler water is nutrient-rich. Thus Sathyendranath *et al.* (1991) used SST from AVHRR to derive estimates of N over Georges Bank off Nova Scotia, from which estimates of f were derived, and hence a measure of new production. Dugdale *et al.* (1997) used a similar approach off the California coast. Nonetheless, encouraging as these very local examples are, the prospect for extrapolating this method to a wider scale appears problematic. The relationships between SST and N and thence f are likely to be very sensitive to local conditions, and progress will require a lot of detailed, *in situ* measurements to explore their validity and stability (e.g., Henson *et al.*, 2003). However, as the availability of co-located SST and ocean color becomes more reliable, this would appear to be a field of research worth pursuing.

Ongoing research

Research to develop satisfactory methods to quantify primary production using satellite data continues rapidly, building on the basic approaches outlined in the previous section. One interesting development has been described as a carbon-based primary production model (Behrenfeld *et al.*, 2005) because, in addition to using satellite-retrieved chlorophyll as an index of chlorophyll phytoplankton biomass, they also use optical backscatter (retrieved as an inherent optical property of water from satellite ocean color data) as an index for particulate organic carbon. Hence they quantify the carbon:chlorophyll ratio from satellite data rather than

using uncertain assumptions, and claim that this leads to more reliable estimates of primary production.

It is beyond the scope of this book to go into further detail about the variety of other evolving methods, apart from pointing readers towards the most recent literature. Since 2000 there have been three ''round robin'' exercises in which blind intercomparisons were made between the different methods available, organized by NASA's Primary Productivity Working Group. The first two of these (Campbell *et al*., 2002) reached the interesting conclusion that the performance of the algorithms and the degree of correlation with each other were independent of algorithm complexity. The third (Carr *et al*., 2006) concluded that the available models based on ocean color were most challenged by high-nutrient, low-chlorophyll conditions, and extreme temperatures or chlorophyll concentrations. It recommended that further progress in primary production modeling requires improved understanding of the effect of temperature on photosynthesis and better parameterization of the maximum photosynthetic rate. Further outcomes from the third experiment focused on results from the tropical Pacific ocean (Friedrichs *et al*., 2009).

Another helpful review of the role of ocean color sensors in the estimation of primary productivity has been published (IOCCG, 2008, chapter 5) This also summarizes the state of the art for retrieving other important components of the global carbon budget from ocean color sensors (i.e., particulate organic carbon, particulate inorganic carbon—essentially calcium carbonate—and colored dissolved organic matter, or CDOM, in addition to phytoplankton carbon).

7.3.3 Estimating PAR from space

Reliable knowledge of PAR is essential for calculation of primary production. Global estimates of primary production therefore require a source of information about the global distribution of solar irradiance at sea level, which can come from remote-sensing satellites. PAR is also needed by ecosystem modelers in order to drive photosynthesis with realistic values of the light field. The most accurate estimation of short-wave radiation reaching the ground is based on radiative transfer models (Ellingson and Fouquart, 1991). Solar irradiance at the top of the atmosphere is calculated from the solar constant and Earth–Sun geometry for a particular time, day of the year, and location on the Earth. Meteorological satellites provide information about cloud conditions, water vapor, and atmospheric aerosols which are needed by atmospheric transmission models. The models then estimate sea level irradiance under both cloud-free (Gregg and Carder, 1990) and cloudy conditions.

The particular case of PAR turns out to be the easiest to deal with because it is limited to the visible wavelengths of light (Gautier, 1995). In this narrow waveband, the absorption of light by clouds is virtually zero. Thus no matter how thick the cloud or how much scattering occurs, irradiance leaving the cloud base towards the ground must be the difference between solar downward irradiance at the top of the cloud and upward irradiance leaving the cloud. Thus PAR can be readily derived from a knowledge of cloud albedo (Gautier *et al*., 1980; Frouin *et al*., 1988, 1989).

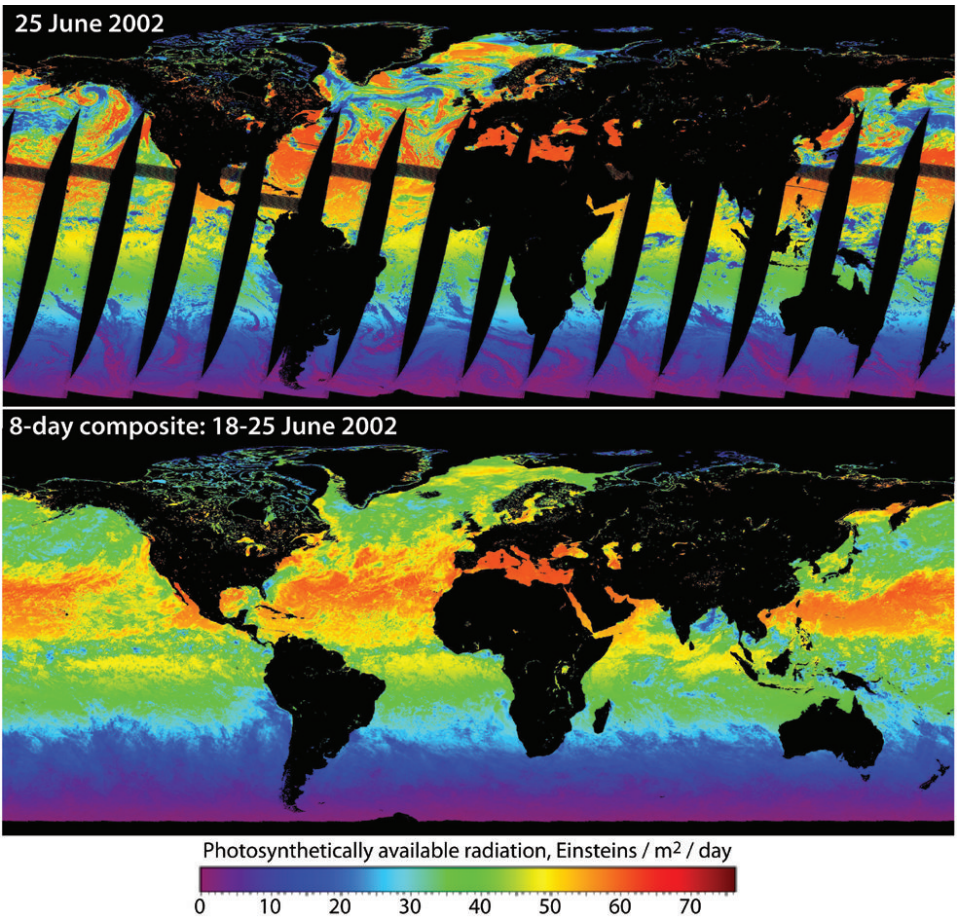


Figure 7.14. Examples of PAR distributions derived from SeaWiFS data. *Upper panel*: PAR for a single day, June 25, 2002. *Lower panel*: 8-day average for the period June 18–25, 2002.

Since 1983, the International Satellite Cloud Climatology Project (ISCCP) (Schiffer and Rossow, 1983) has routinely combined observations from multiple spaceborne sensors to monitor globally the occurrence of clouds and their optical properties and hence to produce a global radiance dataset (Schiffer and Rossow, 1985). From this the variability of total solar irradiance fields and PAR over the ocean can be determined at various timescales, from days to years (Bishop and Rossow, 1991; Pinker and Laszlo, 1992).

Global maps of PAR reaching the sea surface were generated daily from SeaWiFS data. The method assumes decoupling the effect of clouds on light reaching the ground from the effect of a clear atmosphere (Frouin and Cherlock, 1992; Frouin *et al.*, 2003). Details of the retrieval method are also described in an online document (Frouin *et al.*, 2001). An example of typical global maps of PAR is shown in Figure 7.14. The upper panel is from a single day and shows very clearly how dependent

PAR is on synoptic weather and cloud distribution, which generate strong variability at short lengthscales that must significantly impact primary production day by day. The lower panel shows an 8-day composite which, while smoother, still shows considerable weather dependence. This demonstrates clearly how important it is when modeling primary production to use actual PAR rather than an idealized or climatological value. Methods are also available for deriving PAR from multiple sensors providing more comprehensive atmospheric information (Bouvet *et al.*, 2002).

7.3.4 Measurements of primary production

Following the approaches outlined in previous sections, estimates have been made of global primary production and its geographical distribution. Using the biogeographical province method, Platt *et al.* (1995) presented the first monthly maps of production over the North Atlantic Basin, for 1979, based on CZCS data. They estimated that annual production for this basin, which accounts for about 15% of the surface area of the World Ocean, was 10.5 Gt C/yr. They had no way of confirming their results, since after using all the available *in situ* measurements to determine the characteristics of 18 different provinces, there were no independent measurements remaining for validation. Their error analysis concluded that the most serious uncertainties in production came from errors in specifying photosynthesis parameters, compared with which errors from getting the vertical distribution of chlorophyll wrong were much less significant.

An early calculation of global productivity (Antoine *et al.*, 1996) based on the ψ^* approach provided an alternative assessment. This generated monthly productivity values based on monthly global composite biomass maps from CZCS. Figure 7.15 shows the geographical distribution of resulting annually integrated production. The global total was estimated to be between 36.5 Gt C/yr and 45.6 Gt C/yr. A map such as this and those in Platt *et al.* (1995) represented at the time an exciting achievement, the culmination of much careful work. Even more than global total production, it is confirmation of the geographical distribution that has most to offer marine biologists. For example, it is possible to estimate relative contributions to global production from open-ocean oligotrophic and mesotrophic zones, and coastal eutrophic zones (Antoine *et al.*, 1996) (as shown in Table 7.1). The table confirms that, although the rates of production per unit area are very much higher in eutrophic waters, they still make a relatively small contribution to the total. As with all pioneering work, these first estimates of production need to be treated with great caution, particularly since they were based on CZCS data.

Given the availability of SeaWiFS data, and more recently additional information from the MODIS and MERIS sensors, more reliable estimates of global and regional production are to be expected. On the other hand, the differences between estimates by different models (Campbell *et al.*, 2002; Carr *et al.*, 2006; Friedrichs *et al.*, 2009) seems to have made those working in the field more cautious about publishing definitive maps of global primary production, since they are more aware of the uncertainties that still remain.

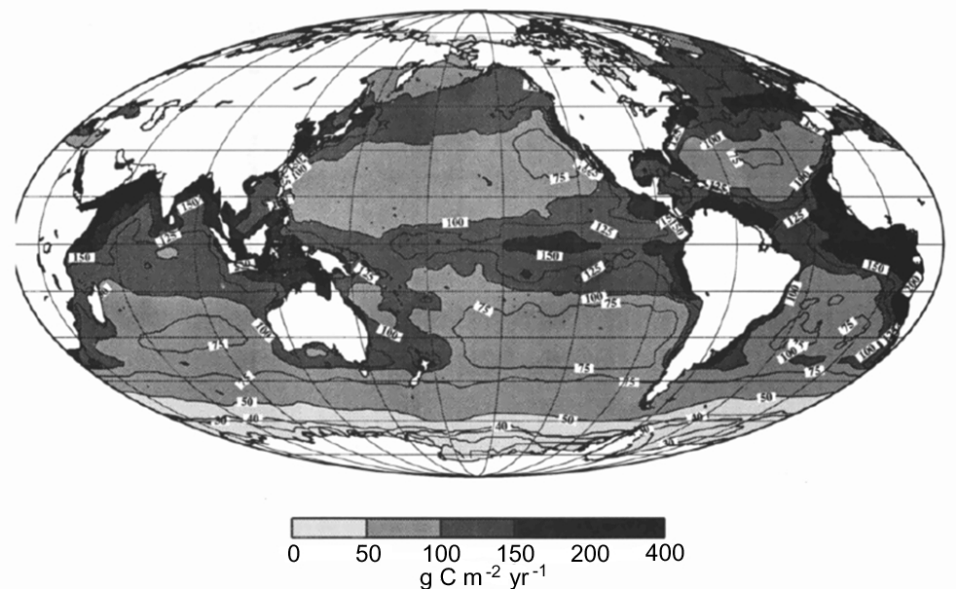


Figure 7.15. Annual primary production within the World Ocean, based on application of the ψ^* approach giving a global annual production of 36.5 Gt C/yr (Antoine *et al.*, 1996).

Table 7.1. Relative contribution of different ocean zones to total primary production (Antoine, *et al.*, 1996).

Zone defined by chlorophyll concentration ($mg\,m^{-3}$)	*Area* (%)	*Total production*		*Production per unit area* ($g\,C\,m^{-2}\,yr^{-1}$)
		(%)	($Gt\,C\,yr^{-1}$)	
Oligotrophic, Chl ≤ 0.1	55.8	44.0	14.5	91.0
Mesotrophic, $0.1 <$ Chl ≤ 1	41.8	47.5	15.7	131.5
Eutrophic, Chl > 1	2.4	8.5	2.8	422.0
Total	*100*	*100*	*33.0*	

However, at the time of writing in mid-2008, monthly global primary production gridded datasets are available from the Ocean Productivity website[6] of Oregon State University. This not only offers output from three different models using data from either SeaWiFS or MODIS but also gives access to model code and ancillary datasets. Readers wanting to learn more about the modeling of primary production using satellite ocean color data are recommended to explore this website. Figure 7.16a shows an example of their "standard" product, generated by the Vertically

[6] *http://web.science.oregonstate.edu/ocean.productivity/index.php*

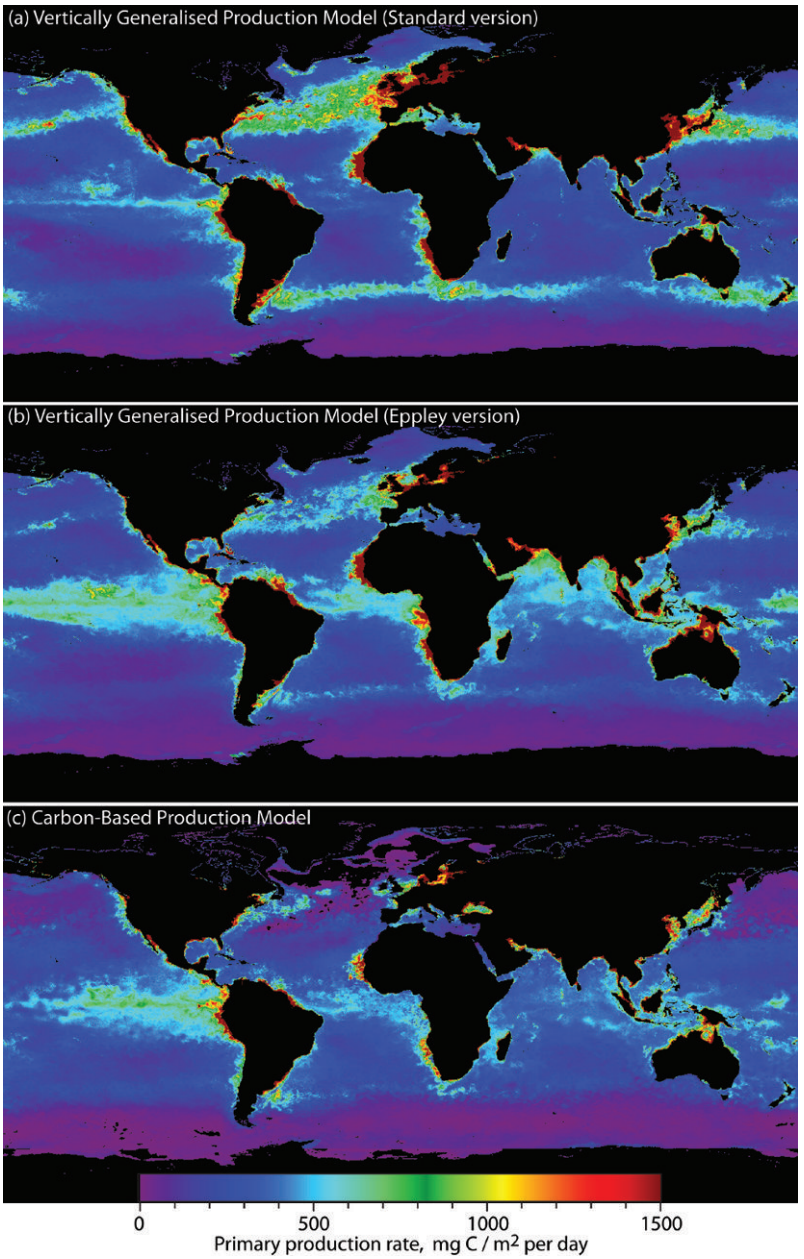


Figure 7.16. Net primary production in April 2007 estimated by (a) the Vertically Generalized Production Model (standard version), (b) the Vertically Generalized Production Model (Eppley version), and (c) the Carbon-Based Production Model, all of the Ocean Productivity Group at Oregon State University (odel output acquired as gridded data from *http://web.science.oregonstate.edu/ocean.productivity/index.php*).

Generalized Production Model (VGPM) (Behrenfeld and Falkowski, 1997b) using MODIS surface chlorophyll concentrations (Chl_{sat}), MODIS SST data, and SeaWiFS cloud-corrected PAR. Euphotic depths are calculated from Chl_{sat} following Morel and Berthon (1989). The model produces data on a grid with a cell size of $1/6^{\circ}$ or about 18 km at the Equator, providing a lot of local detail to which the small-scale global map printed in Figure 7.16 does not do justice. For comparison, Figure 7.16b shows an alternative product. This is again based on VGPM but uses a different formulation to represent quantum yield, based on Eppley (1972) and Morel (1991) and is referred to as the Eppley–VGPM model. The third alternative, Figure 7.16c, is the carbon-based model (Behrenfeld *et al.*, 2005) mentioned in Section 7.3.2.

In addition to these global maps of primary production, many attempts to use satellite data in estimating primary production on a regional or local basis have been published. The regions covered include, for example, Antarctic coastal waters (Dierssen *et al.*, 2000), the Celtic Sea on the northwest European shelf (Joint and Groom, 2000), the North Pacific Central Gyre (Leonard *et al.*, 2001), the northwest Indian Ocean (Watts *et al.*, 1999), the Japan Sea (Yamada *et al.*, 2005; Ishizaka *et al.*, 2007), and the China Sea (Saichun and Guangyu, 2006).

7.4 FISHERIES

7.4.1 General considerations

The potential connection between remote sensing and fisheries has been appreciated since the first use of satellites and aircraft to observe the sea. By the early 1980s the fishing industry was making regular use of satellite data products (Montgomery, 1981; Montgomery *et al.*, 1986) and specific types of application were clearly defined (Laurs and Brucks, 1985). Today the operational importance of satellite data to fisheries continues to influence policy in the U.S.A. where SeaWiFS data in near-real time have been provided commercially, and in Japan where the deployment of combined color and thermal-imaging sensors such as OCTS and GLI has partially been driven by the needs of ocean fishing fleets. Reviews (Santos, 2000; IOCCG, 2008, chapter 6) show continuing operational use of satellite data throughout the world's fisheries, although relatively little ongoing research in the field.

Fishermen are primarily concerned with finding and following shoals of fish in sufficiently large accumulations to make their catch profitable. Operational use of remote sensing can be categorized in three classes. The first is direct location of the fish themselves. This is achieved from aircraft, using a variety of techniques including visual sightings of surface disturbances caused by a school of fish, airborne radar which provides a wider area coverage for the same purpose, the detection of bioluminescence at night, and the use of LIDAR systems tuned to discriminate the signal reflected by fish. However, none of these techniques has yet been

applied effectively from space, and so further discussion of them is beyond the scope of this book.

The second category of remote sensing use is to identify environmental circumstances favorable for catching fish. This approach can be applied effectively from space, is normally specific to certain species, and will be discussed in Section 7.4.3. The third category is the more general, but economically just as important, operational use of satellite data to provide sea state and marine meteorological information to assist in navigation, safety at sea, and planning fishing strategy. This subject is addressed more generally in Chapter 14. This indirect use of satellite data probably yields the greatest benefit since it impacts all types of fisheries, improving safety, reducing the number of days spent at sea under unfavorable conditions, and leading to more economical operation.

However, the challenge to fisheries today lies not so much in how to catch more fish for less cost, but in how to manage fish stocks and to make sure that more efficient fishing methods do not overexploit the natural resource. Thus good fisheries management is needed, based on sound scientific understanding matched with accurate knowledge of both the marine environment and the fish population. Remote sensing can support this endeavor (as discussed in Section 7.4.2). Finally it is worth noting that about 15% of the world's total production of fish and shellfish now comes from aquaculture in which remote sensing has a role to play (as discussed in Section 7.4.4).

7.4.2 Fisheries management and research

Marine fish stocks and catches vary seasonally and interannually. Understanding the links between these fluctuations and the space-time variability of the marine environment is an important element of effective fisheries management. Regularly updated information about the state of the marine environment is needed, initially as input to research about the influence of environmental conditions on fish behavior, and eventually in applying that knowledge to the task of estimating fish stocks and predicting behavior.

An environmental parameter of high importance to many fisheries is water temperature, and especially the spatial patterns of its distribution. This provides a useful indicator of ocean processes important for various fisheries, such as oceanic and shelf sea fronts, coastal upwelling, mesoscale eddies, coastal currents, etc. Many fish are physiologically capable of detecting temperature changes and have adapted their behavioral response to it. The fisherman with knowledge of the temperature structure therefore has an extra aid in predicting the behavior of the fish. It is because temperature, or at least surface temperature, is readily observed from space that fisheries started to benefit very soon after the first thermal sensors were in orbit. Not only is this the primary information used operationally by commercial fisheries but it has also assisted research leading to better overall management and regulation of particular fisheries (Njoku *et al.*, 1985; Fiedler and Bernard, 1987; Myers and Hick, 1990).

Another ocean variable, possibly even more important for fisheries than

temperature, is phytoplankton biomass, which is now readily estimated using satellite ocean color sensors. This provides information about the distribution of primary production which ultimately supplies the food for fish. For some fisheries, such as anchovies and sardines which graze on phytoplankton at points in their life cycle, the connection to chlorophyll retrieved from ocean color is direct. For most other fisheries the connection is more convoluted. Nonetheless, without primary production there would be no higher trophic levels. In the open ocean it is estimated that primary production required to support the fish catch is about 2% of the total, whereas in coastal fisheries the figure is greater than 25% (Pauly and Christensen, 1995).

Very often it is the combination of SST and chlorophyll images which is most helpful for understanding fish behavior. An obvious example of this is the El Niño situation, when the suppression of upwelling off the Peruvian coast cuts off the nutrient supply and reduces primary production with disastrous results for the anchovy fishery. The phenomenon is clearly detectable from SST and color images (see Chapter 11). The absence of operational ocean color sensors for much of the 1980s and 1990s meant that little was written directly about the use of ocean color data for fisheries. While operational, data from SeaWiFS and OCTS were supplied routinely, in near-real time, to a number of fishing fleets, and the willingness of commercial fishing companies to pay for such data implies that it is considered useful.

A good example of how satellite ocean color data has facilitated research leading to better understanding of fisheries is in the area of fish stock recruitment, identifying what determines the number of new individuals joining the fish stock each year. It has been difficult from conventional *in situ* observations to confirm the hypothesis that recruitment is dependent on relative timing of spawning and seasonal phytoplankton bloom (Cushing, 1990). Regularly updated maps of chlorophyll from SeaWiFS, MODIS, and MERIS (as shown in Section 7.2) have made it possible to identify the times of the onset and peak of the spring bloom, to map these times spatially, and to characterize their variability from year to year. Figure 7.17 illustrates some research which exploited this new information in order to confirm the Cushing hypothesis in the case of the haddock fishery in the northwest Atlantic (Platt *et al.*, 2003). A climatology based on SeaWiFS was used to characterize the timing of the spring bloom (as shown in the left-hand panel of the figure). Then using both CZCS (1979–1981) and SeaWiFS (1998–2001) data, the actual timing of the bloom was identified year by year for a particular location off Nova Scotia and recorded as an anomaly relative to climatology. The graph in the right-hand panel of Figure 7.17 shows the survival index of larval haddock, derived from regular field sampling in the same area, plotted against the bloom timing anomaly. The correlation is obvious. This work clearly demonstrates the potential for satellite data to illuminate a scientific problem by providing additional information (in this case about bloom timing) that is not otherwise practically available.

Another research area that is actively using ocean color data is the study of higher predators in the North Pacific. Here the locations of tagged loggerhead turtles and albacore tuna locations deduced from fisheries data have been related to the

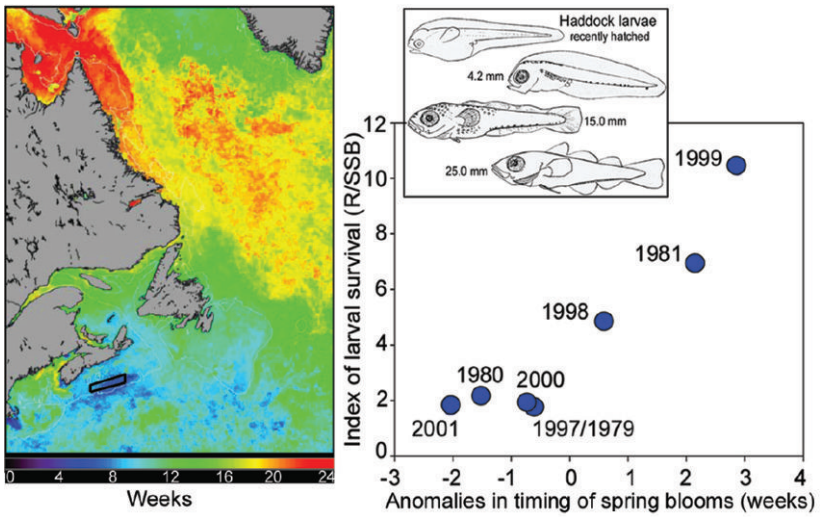


Figure 7.17. *Left panel*: Timing of maximum phytoplankton biomass in the Northwest Atlantic from February to July, derived from SeaWiFS climatology (1998–2001). Units in weeks, changing from blue (indicating early-spring bloom, in March) to red (indicating late-spring bloom, in July). *Right panel*: Relationship between larval haddock (shown in inset) survival index, normalized to recruitment, and local anomalies in bloom timing. Data from the continental shelf east of southern Nova Scotia (see black rectangle on map) for the periods 1979–1981 and 1997–2001 (reprinted with permission from IOCCG, 2003, as adapted from Platt *et al.*, 2003).

position of what is called the Transition Zone Chlorophyll Front (TZCF) between the low-chlorophyll Subtropical Gyre and the high-chlorophyll Subarctic Gyre, which is readily monitored by satellite data (Polovina *et al.*, 2000, 2001). Access to satellite data has thus facilitated monitoring of the marine environment at a large scale to provide a new research context for more conventional marine biology such as studying the migration habits of turtles (Polovina *et al.*, 2004). It has also provided a means of exploring the impact on fisheries of climate change in the ocean, as revealed by ocean color data (Polovina, 2005; Polovina *et al.*, 2008).

7.4.3 Operational applications to specific fisheries

In the next decade, the expected emergence of operational ocean forecasting and monitoring systems based on routine input of satellite data (as discussed in Chapter 14) should start to provide comprehensive knowledge of the oceanographic conditions needed for fisheries. However, it appears that satellite data have already long been used in support of fisheries management by many agencies (see, e.g., Fiedler *et al.*, 1984; Richards *et al.*, 1989; Santos and Fiúza, 1992; Tameishi *et al.*, 1992).

As mentioned above, some fisheries find satellite data sufficiently useful that they are prepared to pay for the routine supply of data. Although some attempts (dis-

cussed by Santos, 2000) have been made to provide fishing fleets with distilled fishery information derived from model assimilation of satellite data, it appears that many of those working at sea prefer to be sent SST maps and to work out their own interpretation based on experience. The criteria for using SST and ocean color maps vary with the type of fishery. Quite a lot is now known about behavior in relation to ocean thermal structures of albacore tuna (Laurs *et al*., 1984; Laurs and Lynn, 1991; Chen *et al*., 2005), bluefin tuna (Maul *et al*., 1984), swordfish (Podestá *et al*., 1993), butterfish (Herron *et al*., 1989), and anchovy (Lasker *et al*., 1981; Fiedler, 1983). In the case of some species the fish seem to congregate at predictable locations within the thermal structure (e.g., on one side of a front). Apparently it is not so much temperature itself that dictates the location, as the behavioral mechanisms of the fish related with feeding activity and the concentration of their prey species. More recent research has also identified the benefit of using color in addition to SST (Zainuddin *et al*., 2004). It is when these patterns of behavior are strong and predictable that fisheries benefit particularly from access to SST and ocean color data in near-real time.

In recent years India has developed its own systems for using satellite data to locate potential fishing zones (PFZs), based on satellite SST data (Narain *et al*., 1990), its own ocean color sensor (Nayak *et al*., 2003), and integration of color and SST (Dwivedi *et al*., 2005). It draws from the research referred to in the previous subsection and issues PFZ advisories three times a week. These may take the form of chlorophyll maps overlaid with SST contours, to show the juxtaposition of SST and ocean color structures (Solanki *et al*., 2003) or just ocean color by itself. Produced by a government agency, they are widely broadcast by all types of media to achieve maximum saturation. The outcomes are monitored and it is reckoned that the system leads to a more efficient fishery operation. However, in the interests of conservation, advisory notices are not issued during June to September which is the peak breeding season.

7.4.4 Aquaculture

Aquaculture may be thought of as the "taming" of fishing. By impounding fish populations, the effort and expense of finding and catching the fish is all but eliminated, but is replaced by the need to nurture and protect the stock. Thus the location chosen for fish farms is crucial, and the monitoring, even prediction, of marine environmental conditions is an important part of the operation. Measuring the ocean at a particular coastal location is normally best done by *in situ* sensors, and most inherent benefits of remote sensing do not apply in this context. It might therefore be supposed that satellite data have little relevance to aquaculture.

However, there are some aspects of aquaculture management in which remote sensing does offer benefits, and has the potential to be used operationally. These are concerned with providing warning of marine environmental hazards that come from the coastal sea adjacent to a sheltered bay or estuary where a fish farm is located. This is the circumstance where information supplied from satellites about the wider geographical context is useful. Physical hazards such as storms or anomalous wave

conditions are best predicted through routine meteorological forecasting, and do not benefit from specific remote-sensing input other than that already assimilated in wind and wave forecasts (see Chapter 8). However, the hazard of harmful algal blooms, which can be catastrophic for fish stocks, is one problem in which remote sensing can play a role if circumstances are appropriate. Sometimes an algal bloom originates from some distance away, and it may be just chance circumstances of wind and tide which bring it towards the aquaculture site. Such blooms can be monitored from space (Yin *et al.*, 1999) using a combination of ocean color and SST sensors.

As discussed in Chapter 13, remote sensing alone cannot be relied upon to discriminate between harmful and benign algal blooms. However, it can give advance warning of the development of a bloom in a potentially hazardous location. If the expense is justified, detection of an algal bloom in ocean color images could then trigger action to physically sample the bloom *in situ* and determine its nature. If a hazard is identified then steps can be taken to protect the stock, using the extra time won by the use of remote sensing.

7.5 HABITATS IN SHALLOW TROPICAL SEAS

Another quite distinct field of marine biology which can benefit from remote sensing is the study of sea bed habitats in water that is both shallow and sufficiently clear for the color and texture of the sea bed to be clearly detected from above. Thus the approach is limited to water depths shallower than about 10 m and is applicable to only a tiny fraction of the total ocean area. Examples are found mainly in tropical coastal seas and lagoons, where ocean color images are used to identify and map the extent of coral reefs and macroalgae, sea grass, and mangroves (Green *et al.*, 2000b). The approach derives from conventional land cover–mapping methods used in terrestrial remote sensing (see, e.g., Lillesand and Kiefer, 1999; Mather, 1999). This section discusses briefly how, and with what application in mind, those methods are adapted for identifying shallow-water habitats. A wider view of satellite oceanography techniques used to study coral reefs is given in Section 7.6.

The mapping of submerged vegetation in coastal lagoons requires multispectral observations of reflected sunlight in the visible waveband to be obtained at the finest spatial resolution available. The idealized objective is to determine the vegetation type in each element of a rectangular grid covering the study area. Since ground surveys of sea bed vegetation may use a grid spacing smaller than 1 m this might seem to be the ideal pixel size for remote sensing. It is possible to achieve such small pixels using airborne sensors, of which there is a wide variety (Green, 2000), but the drawback of low-altitude sensors is their limited area coverage. Within the context of this book, we consider only satellite sensors. Their drawback for detailed vegetation mapping, until very recently, has been that their smallest spatial resolution was measured in tens of meters, although they benefit from a coverage of many tens of kilometers in a single image.

The sensors available for this purpose, primarily the Landsat Thematic Mapper (TM) and the SPOT High Resolution Visible (HRV) radiometer are very different

from typical color sensors used for oceanographic applications in the rest of this chapter. Their spatial resolution is much finer and the swath width much narrower than sensors such as SeaWiFS or MERIS. Table 7.2 lists their spatial, spectral, and temporal sampling characteristics, and those of a number of other fine-resolution visible sensors intended primarily for land mapping. The table includes the CHRIS pointable imaging spectrometer which sacrifices swath width for spectral resolution, but apart from this these fine spatial–resolution sensors have fairly broad spectral bands. While they are capable of differentiating between types of sea bed vegetation that have strong color differences, they do not have the spectral or radiometric resolution to distinguish subtle changes of color used to measure the colored content of water itself. Thus they cannot retrieve a measure of phytoplankton chlorophyll. However, as technology advances this will not always be the case. A number of commercial land-mapping satellite sensors can now provide panchromatic views at resolution down to 1 m, while the WorldView-2 system launched in late 2009 carries a 50 cm resolution sensor that has eight visible and near-IR bands selected not only for commercial mapping but also for monitoring water content, including a 400 nm to 450 nm channel.

When vegetation covering the sea bed is being identified from high-resolution photographs, both color and the spatial texture are used to distinguish between different types of plants. However, until very recently most image datasets obtained from the sensors detailed in Table 7.2 have lacked the spatial resolution to be able to resolve this texture and must rely primarily on spectral variability to identify what is on the sea bed. Thus multispectral classification methods (Mather, 1999) are most successful. In this approach the relationship between reflectance in two or more different wavebands is used to identify clusters of pixels that have similar relationships. ''Unsupervised'' classifiers simply use the results of cluster analysis to partition a scene into different regions, and then attempt to assign sea bed types to different regions, using general knowledge of the region, water depth, etc., but normally with only limited success. However, if *in situ* observations of the sea bed habitat are available for a date close to the acquisition of the satellite image, it is possible to be more objective in assigning particular multispectral clusters to the type of sea floor cover. This approach is known as ''supervised'' classification. If there are sufficient independent observations of sea floor type, it becomes possible to determine which types of sea bed cover are distinguishable and which are not. It may also be possible to identify pixels which are a mixture of two or more types, as well as to assign detection accuracies, and hence achieve confidence estimates when a scene is classified.

The most significant difference between the classification of land cover and of the sea bed is caused by the effect of the water column through which sunlight must pass twice if it is to be reflected into the sensor field of view. In all but the clearest of waters, the absorption and scattering of light by sea water contents prevents the technique from working at all. Yet even when there are extremely few particulates or colored dissolved substances in the sea, the water itself scatters and absorbs light as shown in figure 6.17 of *MTOFS* (Robinson, 2004). The red end of the spectrum is preferentially attenuated, and the blue least affected. The effect on the typical

Table 7.2. High-resolution visible and near-infrared sensors with potential for shallow seabed vegetation mapping, showing only the visible and near-infrared waveband channels. Status checked up to November 2009.

Sensor	*Satellite/Agency*	*Pixel size* (m)	*Spectral bands (VNIR)* (nm)	*Revisit interval* (days)	*Swath width* (km)	*Period of operation*
MSS Multi-spectral scanner	Landsats 1, 2, 3/ NASA	76	500–600 600–700 700–800 800–1,100	18	185	L1: 1972–1978 L2: 1975–1982 L3: 1978–1983
TM Thematic mapper	Landsats 4 & 5/ NASA	30	450–520 520–600 630–690 760–900	16	185	L4: 1982–1984 L5: 1984–now
ETM Extended thematic mapper	Landsat 7/NASA	15[a] 30	520–900[a] 450–515 525–605 630–690 750–900	16	185	1999–now
HRV High-resolution visible	SPOTs 1, 2, 3 and 4/CNES	10[a] 20	490–690[a] 500–590 610–680 790–890	26	120	S1: 1986–1990 S2: 1990–now S3: 1993–1996 S4: 1998–now
HRG High-resolution geometrical	SPOT 5/CNES	5[a] 10	490–690[a] 500–590 610–680 790–890	26	120	2002–now
LISS-I, -II, -III Linear imaging self-scanning sensor	I: IRS-1A II: IRS-1B III: IRS-1C Indian Space Agency	I: 73 II: 36.5 III: 23.5	450–520(I,II only) 520–590 620–680 770–860	I & II: 22 III: 24	I: 148 II: 146 III: 142	1988–1992 1991—? 1995—?

AVNIR Advanced visible and near-IR radiometer	ADEOS/NASDA	8[a] 16	520–690[a] 420–500 520–600 610–690 760–890	41	80	
IKONOS	Space Imaging Inc.	(a) 1m	(a) 450–900[a] (b) 445–516 506–595 632–698 757–853	Pointable 3-5 off nadir 144 true-nadir	11	2000–now
CHRIS Compact high-resolution imaging spectrometer	PROBA / ESA	(a) 36 (b) 18	(a) 63 bands between 400 and 1,050 nm (b) Only 18 bands	Variable: pointable sensor	14	2001–
Quickbird	DigitalGlobe, Inc.	(a) 60 & 70 cm (b) 2.4 & 2.8 m	(a) 450–900[a] (b) 450–520 520–600 630–690 760–900	Pointable along-track and cross-track 3–7	16.5	2002–now
WorldView-1	DigitalGlobe, Inc.	50 cm	450-900[a]	Pointable 1.7 possible	17.6	2007–now
WorldView-2	DigitalGlobe, Inc.	50 cm	450–800[a] 400–450 450–510 510–580 585–625 630–690 705–745 770–895 860–1,040	Pointable 1.1 typical, or <1 in conjunction with WorldView-1	16.4	Oct 2009–now

[a] Denotes a broadband panchromatic mode achieving finer spatial resolution.

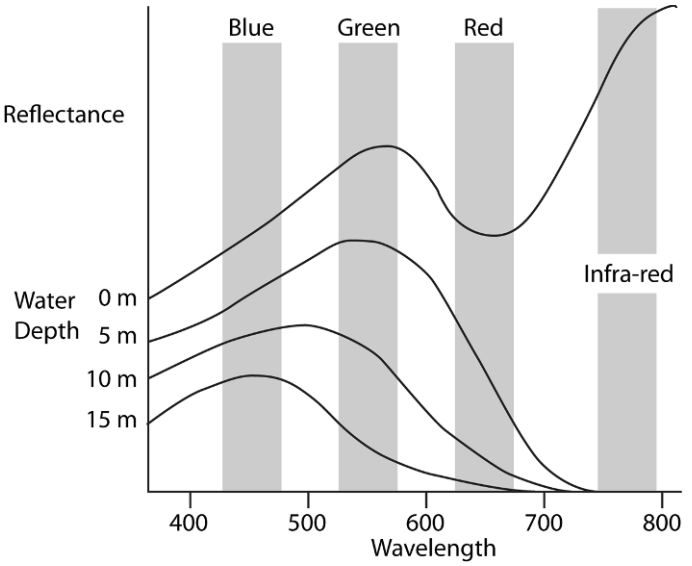


Figure 7.18. Schematic showing how the apparent spectral reflectance of the sea bed appears to vary when viewed through different depths of seawater. The gray bars show typical wavebands of high-resolution multi-spectral imagers (after Mumby and Edwards, 2000).

spectral reflectance signature of a sea floor surface when viewed through clear water of different depths is shown in Figure 7.18. The greatest impact occurs in the infrared part of the spectrum, which is effectively eliminated from the reflected signal by just a few centimeters of water, thus preventing the use of "red edge" vegetation detection methods used extensively in land remote sensing.

Differential attenuation, increasing with wavelength from blue to green, causes the reflectance spectral signature of a particular sea bed type to vary considerably with water depth, to the extent that it can be mistaken for a different bottom type at a different water depth. For example, white sand at a depth of 20 m may appear to have a spectral signature that is similar to sea grass at a depth of 2 m. It is therefore essential to apply a correction for this depth effect before applying the classification algorithm to the data, or else to adapt the classification technique to allow for the depth effect. To do this from first principles would require a good knowledge of the bathymetry of the study area and the attenuation characteristics of the water. However, a method developed by Lyzenga (1981) allows relative attenuation effects between pairs of spectral channels to be determined if sufficient pixels in the scene can be identified as having the same bottom substrate or vegetation cover. Practical implementation of this method, and the simple atmospheric correction needed before applying it, is illustrated by Mumby and Edwards (2000). Wavebands must be between 400 nm and 650 nm to be capable of seeing the sea bed, and those at the longer (red) end of this range are usable only in the shallowest waters.

Green *et al.* (2000a) provide a clear explanation of the methodology of ground cover classification applied in the underwater context, based on a comprehensive study of the application of remote sensing to the coasts and reefs of the Turks and

Caicos Islands at the eastern end of the Bahamas chain in the Caribbean Sea. They conclude that satellite remote sensing, using either Landsat or SPOT data, offers a cost-effective option for broad mapping of sea bed habitats over regions having dimensions of several tens of kilometers. Because of the fairly coarse spatial and spectral resolution of these sensors, the approach is limited to distinguishing four classes of sea bed: sea grass, macroalgae, coral reef, and sand (plus unclassified sea bed and deep water where no sea bed is visible). Figure 7.19 shows an example of a satellite-derived habitat map for the Caicos Bank region (Mumby and Green, 2000), based on a SPOT multispectral image. The overall accuracy of classification is estimated to be 73% (i.e., 27% of the pixels are likely to be misclassified). To increase the number of detectable classes, and also improve accuracy, requires use of an airborne sensor. Using the Compact Airborne Spectral Imager (CASI) with 16 spectral bands, a nine-class map of a subarea of Figure 7.19 was achieved with an accuracy of 81%. These and similar studies (Zainal *et al.*, 1993; Purkiss *et al.*, 2002) confirm that satellite sensors offer the tropical ecosystem biologist a means of approximately classifying the sea bed of large reef and lagoon areas. For more detailed analysis, airborne or *in situ* surveys have proved to be necessary, although recent use of 4 m resolution data from the Ikonos satellite, with a similar spectral capability to the Landsat TM, has allowed eight substrate classes to be resolved with an overall accuracy of 69% down to a water depth of 6 m (Purkiss, 2005).

In addition to distinguishing between different habitats, it is also possible to use remotely sensed data to quantify the relative proportions of two contrasting sea bed types. In the case of pixels consisting of mixtures of sand and sea grass, this provides the basis for estimating the standing crop of sea grass, measured as biomass per unit area. Semiempirical models have been developed to achieve this (Armstrong, 1993; Mumby, *et al.*, 1997). While airborne scanners provided spatial detail at the small scales (<10 m) needed to study the dynamics of sea grass, satellite-based estimates of biomass are still useful for mapping on the large scale (Dekker, *et al.*, 2006). Moreover, because the Landsat and SPOT archives span many years, there is the possibility of being able to make retrospective studies of large-scale changes in sea grass coverage. For example, the spatial extent of mass mortality of sea grass in Florida Bay was detected by comparing SPOT images before and after the event (Robblee, *et al.*, 1991). Another important application of satellite data to tropical coastal ecosystems is the monitoring and measurement of mangrove forests (Green and Mumby, 2000), although since these are detected above the water surface the remote-sensing methods are strictly beyond the scope of this book.

It seems fair to say, from the perspective of marine biological research, that the role of satellite remote sensing in the study of sea bed habitats in tropical seas is still rather limited by coarse spatial resolution. However, this should not be allowed to detract from the significant contribution that satellite monitoring and mapping can make to the management of shallow tropical seas, reefs, and lagoons (Green *et al.*, 1996; Mumby *et al.*, 1999). The capacity to monitor and regularly remap the sea bed over fairly large areas is perceived by many coastal managers as providing a good background for management planning and for detecting changes in habitats over time. Relatively cheap overviews from satellites can provide the basis for planning

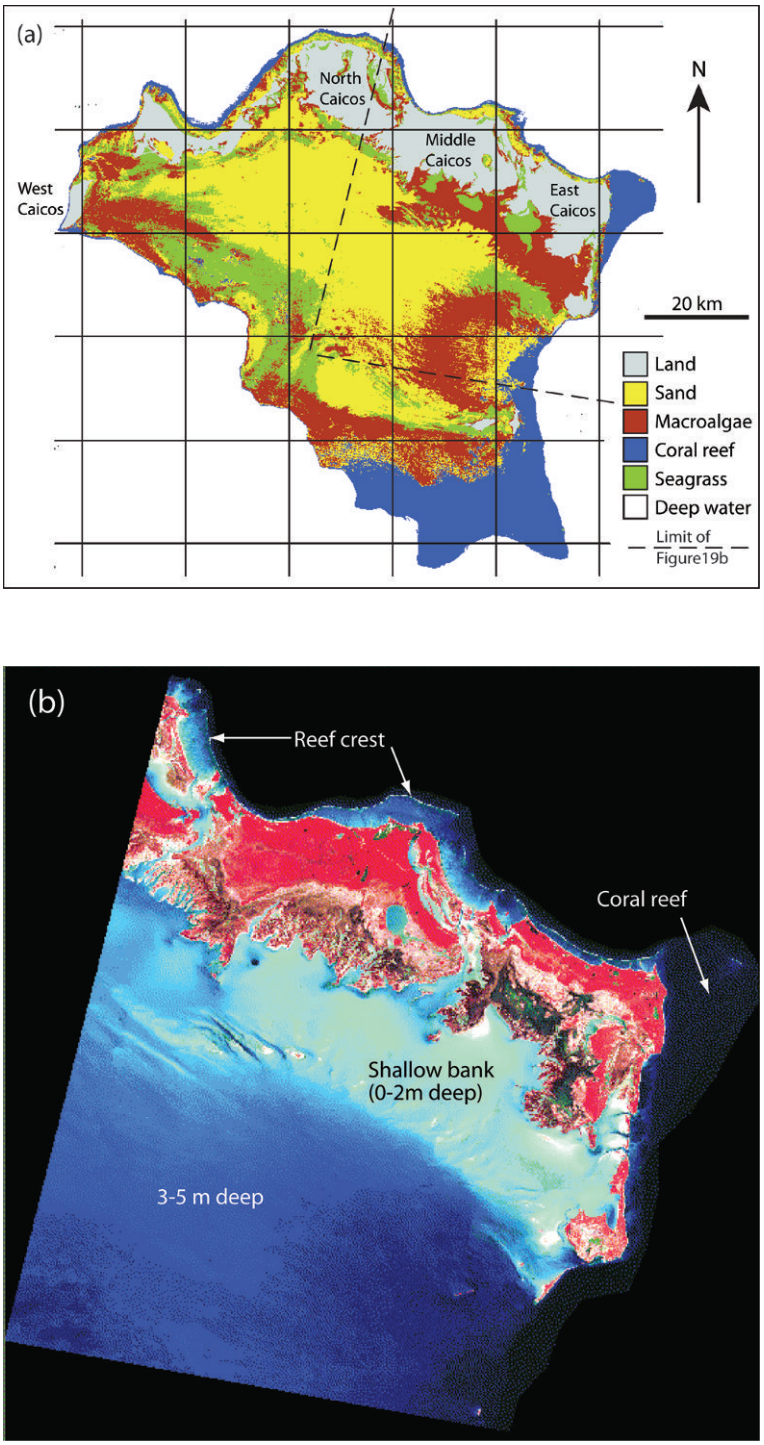


Figure 7.19. (a) Broad-scale habitat map of the Caicos Bank derived by supervised classification (four classes only) from a SPOT XS dataset. (b) A false-color composite image of SPOT multispectral image data from which (a) was derived, with band XS1 in blue, XS2 in green, and XS3 in red. This example comes from Mumby and Green (2000) and is a simplified version of Plate 12 of Green *et al.* (2000b) (provided courtesy of Alasdair Edwards, University of Newcastle).

monitoring strategies, and help to identify what density of *in situ* sampling is required to ensure representativeness. High-resolution satellite data also serve as a valuable tool to help managers meet their responsibilities in respect of establishing conservation criteria, locating fishing areas, and planning recreational usage of tropical coastal waters. We may expect coastal marine environmental managers in future to make increasing use of very high–resolution multispectral sensors like WorldView-2, coupled with geographical information systems for organizing *in situ* measurements.

7.6 CORAL REEFS—A WIDER ROLE FOR SATELLITE DATA

It is clear from the previous section that the remote-sensing methods needed for detailed study of coastal and reef aquatic ecosystems are essentially an extension of land-based, high-resolution mapping. Marine biologists working in this particular field might not expect to derive much benefit from the main methods of ocean remote sensing as defined in *MTOFS* and as applied to most of the other topics discussed in this book. That is mainly because the lengthscales of variability encountered on reefs and in benthic ecosystems is so much shorter than most other marine phenomena described in other chapters. However, there is one aspect of reef biology in which the wider overview provided by satellite oceanography techniques has become essential, and important enough to require this subsection to itself. This is the issue of coral bleaching, and the role that satellite monitoring of sea surface temperature (SST) plays in identifying regions where reefs are at risk of bleaching.

Corals are underwater animals that attach themselves to stony substrates. The order of corals known as stony corals, or *scleractinians*, are found as large colonies of individual coral polyps, each of which produces limestone deposits. Over the years these deposits have created the large reef systems found in shallow tropical and temperate seas, which provide a unique habitat for rich and complex ecosystems (see, e.g., pp. 117–141 in Barnes and Hughes, 1999). Corals thrive by hosting within their cells symbiotic algae called *Zooxanthellae*, which provide the coral with oxygen and organic compounds resulting from photosynthesis, while themselves obtaining from the coral carbon dioxide and other chemical compounds needed for photosynthesis. The algae give coral reefs their rich coloration and the symbiotic relationship is essential for the health of the whole reef ecosystem.

Coral bleaching is the name given to the situation when corals are subject to physiological stress and respond by ejecting the zooxanthellae. The departure of the algae is visually evident because corals lose the pigments that give them their yellow or brown coloration. In this case the white limestone substrate that the corals have deposited shows through the translucent cells of the polyps which then appear pale or even white. If the stress is quickly removed the algae return within a few weeks and the corals recover, but if the stress is prolonged for many weeks the corals will die and continue to appear stark white. The loss of live corals eventually causes damage to the whole reef ecosystem. Consequently coral-bleaching events pose a serious threat that is taken seriously by marine environmental managers.

The causes of coral bleaching include factors such as disease and changes in salinity or increased sedimentation, but the most consistent factor associated with coral-bleaching events are when water temperature is elevated above its typical maximum for that location; high surface solar irradiance in the ultraviolet may also be implicated (Glynn, 1996; Hoegh-Guldberg, 1999). These latter two variables are capable of routine monitoring by satellites. Because anomalous behavior of water temperature typically occurs over lengthscales of several tens of kilometers or greater, there is no need for very localized measurements. Composite SST products with 50 km resolution, derived from medium-resolution sensors (~1 km pixels) are satisfactory. Moreover, although locally it is water temperature at the depth of the corals that matters, the detection of anomalies relative to climatology can be based on surface temperature (i.e., the SST monitored by satellites).

Over the last decade, NOAA have developed their Coral Reef Watch[7] (CRW) service within the Center for Satellite Applications and Research (STAR). SST observations derived from AVHRR are used to note when and where temperature rises above a "bleaching threshold" temperature. This is defined as 1°C above the mean temperature for the warmest month of the year at a particular location (Glynn and D'Croz, 1990). It is therefore not a uniform threshold across all reef sites but is based on the assumption that corals are already acclimatized to maximum temperatures typically encountered at a particular location over the previous several years. Thus the bleaching threshold can be obtained quite easily from analysis of the 12 global monthly climatologies of SST. Using this as a reference, "hotspot" maps are published twice per week by the U.S. agency NOAA spanning the globe between 45°N and 45°S, and in a number of regional submaps at finer resolution. These identify locations where average SST during the last 4 days of satellite retrievals exceeds the maximum monthly climatology. Figure 7.20 is an example from the Carribean. Mauve–blue colors are used where the excess SST is between 0°C and 1°C. The brighter range of colors (yellow to red), representing excess SST greater than 1°C, show where the bleaching threshold has been crossed and therefore indicates where, and by how much, reefs are in danger of damage due to thermal stress.

The longer the period of time over which excess temperatures are encountered, the greater the risk of coral bleaching, so that in addition to the hotspot maps it is important for reef management agencies to know about accumulated thermal stress. This information is contained in twice weekly produced maps of "degree heating weeks" (DHWs) such as that in Figure 7.21. This integrates excess SST above the bleaching threshold per week over the previous 12 weeks. Thus, for example, a value of 10 DHW could be the result of having an excess of 1°C for 10 of the previous 12 weeks, or an excess of 2°C for 5 of the previous 12 weeks, or any other combination adding up to 10 DHW. Note that if SST falls below the bleaching threshold the contribution to the DHW index that week is zero; it does not become negative. In practice the DHW index is updated twice per week. Note that the DHW index will remain nonzero for a further 12 weeks from the last time the threshold was exceeded,

[7] *http://coralreefwatch.noaa.gov/satellite/index.html*

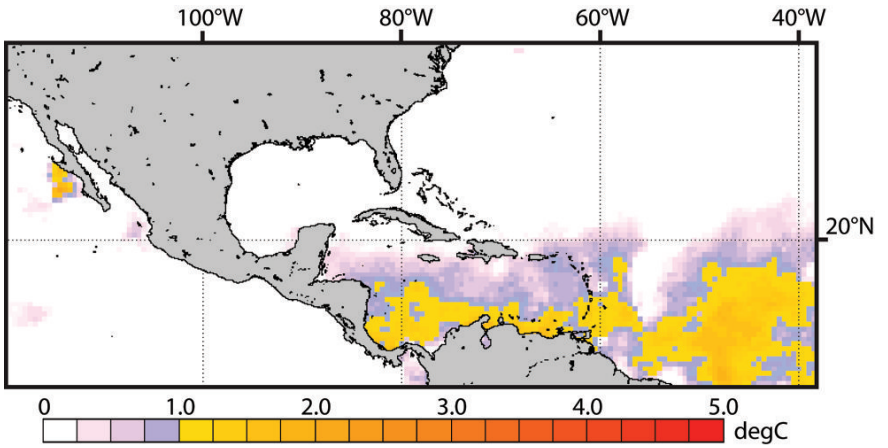


Figure 7.20. Example of a NOAA Coral Reef Watch map of bleaching hotspots, corresponding to conditions on October 16, 2008. The scale represents SST in degrees Celsius above the maximum monthly mean. The bleaching threshold is exceeded when values are above 1.0°C (image downloaded and adapted from NOAA Coral Reef Watch website).

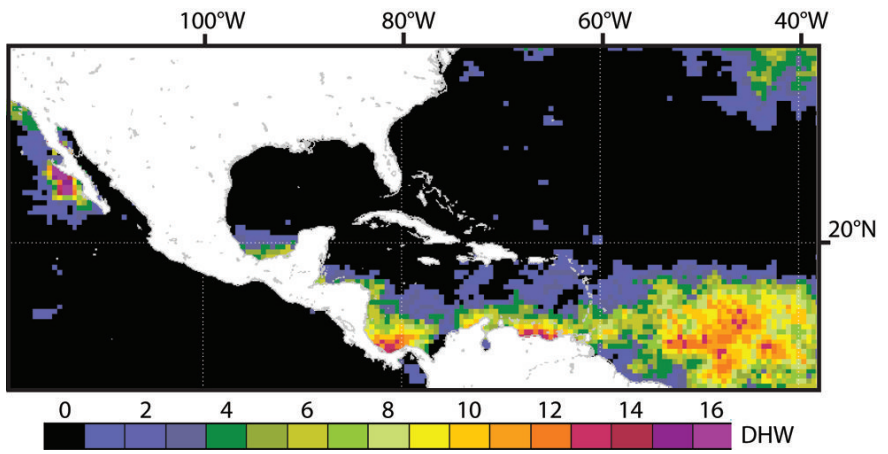


Figure 7.21. Map of degree heating week (DHW) index produced by NOAA Coral Reef Watch corresponding to conditions on November 13, 2008 (image downloaded and adapted from NOAA Coral Reef Watch website).

representing the period that the reef remains at risk while it takes time to recover after the direct thermal stress has been removed.

CRW has been producing maps of hotspots and the DHW index since 1998 and thus retain an archive of the thermal stress history of the tropical ocean. It is worth noting that these data are provided on a global scale, since they derive from global satellite SST data products, even though users of such data are likely to be locally

based marine conservation and environmental-monitoring officers. These are the scientists who compare the occurrence of actual coral bleaching in their local region with the recorded risk as represented by the DHW index. Because the DHW index is common to all reefs, it encourages comparison of how different reefs react to similar thermal stresses, providing a basis for research to identify the additional factors that may cause certain reefs to be more prone to coral bleaching than others. It is worth noting in passing that this is another benefit of satellite remote sensing: to stimulate scientific collaboration between geographically separated marine scientists focused on local ecosystems and who might remain completely independent from each other if they did not share a common global data source from satellites.

NOAA is presently trying out an experimental light stress product which seeks to represent the combined risk from both thermal stress and the strength of photosynthetically available radiation (PAR) (discussed in Section 7.3.3). NOAA also provides a subscription service which will automatically send warnings to subscribers about excess temperatures in specific reef areas.

7.7 MARINE BIOLOGY IN THE FUTURE

This chapter has presented a mixture of diverse ways in which satellite oceanography has become closely engaged with some aspects of marine biology. How is this partnership likely to progress? In the near future it is to be expected that the benefits of extending global time series of chlorophyll from SeaWiFS, MODIS, and MERIS datasets will be demonstrated in further papers exploring the natural variability of phytoplankton biomass distribution in the ocean, its seasonal variability, and its longer term variations. Where long-term change is detected in satellite observations of phytoplankton, we should expect to see it related to physical and geochemical causes in the ocean, and analysis will be needed to determine the extent to which climate change associated with global warming is implicated. It will also be important if possible to link long-term variability in the satellite ocean color record to variability detected by conventional marine biological analysis at established sampling stations for long time series, such as Bermuda. The same should be the case with improved estimates of primary production. The methodology for doing this is steadily improving and it is just a matter of time to accumulate and process the data before we learn about long fluctuations or trends in production and its geographic distribution that may correlate with other biological indicators.

Experience with MODIS and MERIS has prompted research with experimental algorithms to exploit their improved spectral resolution. We may expect to see the development of improved methods for estimating other biogeochemical variables than simply chlorophyll, which will contribute to the study of ocean biogeochemical fluxes. We should also expect to see improvements in the algorithms and methodology for analyzing ocean color data in Case 2 waters that will enable more reliable measurements of biomass and primary production in coastal waters and shelf seas,

although there is still a lot of hard work to be done in this field (Robinson *et al.*, 2008).

The global coverage of high-quality data viewing parts of the Earth hitherto unstudied is also likely to be a fruitful research field for the partnership between marine biology and satellite oceanography. Already studies of Antarctic waters prompted by the availability of SeaWiFS data have shown some interesting different optical characteristics, with high chlorophyll content bluer than in other seas, which may require modification to chlorophyll algorithms (Dierssen and Smith, 2000). Once this is resolved there is a lot more to be learned about the bursts of productivity that occur in the marginal ice zone as the ice melts. As the Arctic Ocean is expected to lose more of its ice in summer, remote sensing should be able to tell a story about changes in primary production at the ice edge there too.

In a 5 to 10-year timeframe, as well as the serendipitous discoveries that constant exposure to a stream of new data is bound to bring to some of the next generation of researchers reading these pages, the expected major development that will affect all marine scientists will be the emergence of operational oceanography. This theme is picked up in more detail in Chapter 14. What is envisaged is a merging and blending of all ocean measurements, both from space and *in situ*, using assimilation into high-resolution mathematical models that will maximize the information available about the sea at any given time and location, to provide the background environmental knowledge against which specific marine biological research is performed. These models should be capable not only of predicting ocean circulation dynamics and the physics and thermodynamics of the sea but also its biogeochemistry. This is the likely direction in which improvements in the estimation of primary production will be made.

Improved models constrained by satellite data also promise the opportunity to examine hypotheses concerning the way the biota of the ocean, particularly phytoplankton, may exert a greater measure of control over their physical environment than was previously recognized, through changing optical scattering and absorption of solar energy (Miller *et al.*, 2003; Manizza *et al.*, 2005; Subrahmanyam *et al.*, 2008), and perhaps even through the effects of natural surface films on the air–sea interface. Satellite remote sensing has a catalytic role in this approach, in that it provides a steady stream of spatially dense, regularly repeated measurements against which to refine, test, and validate ocean-forecasting models that may become the research tools of a new generation of satellite marine biologists!

7.8 REFERENCES

Aiken, J., J. R. Fishwick, S. J. Lavender, R. Barlow, G. Moore, and H. Sessions (2007), Validation of MERIS reflectance and chlorophyll during the BENCAL cruise October, 2002: Preliminary validation of new products for phytoplankton functional types and photosynthetic parameters. *Int. J. Remote Sensing*, *28*, 497–516.

Alvain, S., C. Moulin, Y. Dandonneau, and F. M. Breon (2005), Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep-Sea Res. I*, **52**, 1989–2004.

Antoine, D., and A. Morel (1996), Oceanic primary production, 1: Adaptation of a spectral light-photosynthesis model in view of application to satellite chlorophyll observations. *Global Biogeochemical Cycles*, **10**(1), 43–55.

Antoine, D., J.-M. André, and A. Morel (1996), Oceanic primary production, 2: Estimation at global scale from satellite (coastal zone color scanner) chlorophyll. *Global Biogeochemical Cycles*, **10**(1), 57–69.

Armstrong, R. A. (1993), Remote sensing of submerged vegetation canopies for biomass estimation. *Int. J. Remote Sensing*, **14**, 10–16.

Balch, W. M., P. M. Holligan, S. G. Ackleson, and K. J. Voss (1991), Biological and optical properties of mesoscale coccolithophore blooms in the Gulf of Maine. *Limnology and Oceanography*, **36**, 629–643.

Barale, V., J.-M. Jaquet, and M. Ndiaye (2008), Algal blooming patterns and anomalies in the Mediterranean Sea as derived from the SeaWiFS data set (1998–2003). *Rem. Sens. Environ.*, **112**, 3300–3313.

Barnes, R. S. K., and R. N. Hughes (1999), *An Introduction to Marine Ecology* (Third Edition, 286 pp.). Blackwell Science Ltd., Oxford, U.K.

Behrenfeld, M. J., and P. G. Falkowski (1997a), A consumer's guide to phytoplankton primary production models. *Limnology and Oceanography*, **42**, 1479–1491.

Behrenfeld, M. J., and P. G. Falkowski (1997b), Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*, **42**, 1–20.

Behrenfeld, M., E. Boss, D. A. Siegel, and D. M. Shea (2005), Carbon-based ocean productivity and phytoplankton physiology from space. *Global Biogeochemical Cycles*, **19**(GB1006), doi: 10.1029/2004GB002299.

Bishop, J. K. B., and W. B. Rossow (1991), Spatial and temporal variability of global surface solar irradiance. *J. Geophys. Res.*, **96**, 16839–16858.

Bouvet, M., N. Hoepffner, and M. D. Dowell (2002), Parameterization of a spectral solar irradiance model for the global ocean using multiple satellite sensors. *J. Geophys. Res.*, **107**(C12), 3215, doi: 10.1029/2001JC001126.

Brock, J. C., S. Sathyendranath, and T. Platt (1998), Biohydro-optical classification of the northwestern Indian Ocean. *Mar. Ecol. Prog. Ser.*, **165**, 1–15.

Broerse, A. T. C., T. Tyrell, J. R. Young, A. J. Poulton, A. Merico, and W. M. Balch (2003), The cause of bright waters in the Bering Sea in winter. *Cont. Shelf Res.*, **23**, 1579–1596.

Brown, C., and G. P. Podestá (1997), Remote sensing of coccolithophore blooms in the Western South Atlantic Ocean. *Rem. Sens. Environ.*, **60**, 83–91.

Brown, C. W., and J. A. Yoder (1994), Coccolithophorid blooms in the global ocean. *J. Geophys. Res.*, **99**, 7467–7482.

Campbell, J. (1995), The lognormal distribution as a model for bio-optical variability in the sea. *J. Geophys. Res*, **100**(C7), 13237–13254.

Campbell, J., D. Antoine, R. Armstrong, K. Arrigo, W. Balch, R. Barber, M. Behrenfeld, R. Bidigare, J. Bishop, M.-E. Carr *et al.* (2002), Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance. *Global Biogeochemical Cycles*, **16**(3), 74.

Carr, M.-E., M. A. Friedrichs, M. Schmeltz, M. N. Aita, D. Antoine, K. R. Arrigo, I. Asanuma, O. Aumont, R. Barber, M. Behrenfeld *et al.* (2006), A comparison of global estimates of marine primary production from ocean color. *Deep-Sea Res. II*, **53**, 741–770.

Chen, I.-C., P.-F. Lee, and W.-N. Tzeng (2005), Distribution of albacore (*Thunnus alalunga*) in the Indian Ocean and its relation to environmental factors. *Fish. Oceanogr.*, **14**, 71–80.

Cushing, D. H. (1990), Plankton production and year-class strength in fish populations: An update of the match mismatch hypothesis. *Adv. Mar. Biol.*, **26**, 249–294.

Dekker, A., V. Brando, J. Anstee, S. Fyfe, T. Malthus, and E. Karpouzli (2006), Remote sensing of seagrass ecosystems: Use of spaceborne and airborne sensors. In A. Larkum, R. Orth, and C. Duarte (Eds.), *Seagrasses: Biology, Ecology and Conservation* (pp. 347–359). Springer-Verlag, New York.

Dierssen, H. M., and R. C. Smith (2000), Bio-optical properties and remote sensing ocean color algorithms for Antarctic Peninsula waters. *J. Geophys. Res.*, **105**(C11), 26301–26312.

Dierssen, H. M., M. Vernet, and R. C. Smith (2000), Optimizing models for remotely estimating primary production in Antarctic coastal waters. *Antarctic Science*, **12**(1), 20–32.

Dugdale, R. C., C. O. Davis, and F. P. Wilkerson (1997), Assessment of new production at the upwelling center at Point Conception, California, using nitrate estimated from remotely sensed sea surface temperature. *J. Geophys. Res.*, **102**, 8573–8585.

Dwivedi, R. M., H. U. Solanki, S. Nayak, D. Gulati, and V. S. Sonvanshi (2005), Exploration of fishery resources through integration of ocean colour and sea surface temperature. *Ind. J. Mar. Sci.*, **34**(4), 430–440.

Edwards, M., and A. J. Richardson (2004), Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature*, **430**, 881–884.

Ellingson, R. G., and Y. Fouquart (1991), The intercomparison of radiation codes in climate models: An overview. *J. Geophys. Res.*, **96**, 8925.

Elskens, M., L. Goeyens, F. Dehairs, A. Rees, I. Joint, and W. Baeyens (1999), Improved estimation of f-ratio in natural phytoplankton assemblages. *Deep-Sea Research*, **46**, 1793–1808.

Eppley, R. W. (1972), Temperature and phytoplankton growth in the sea. *Fishery Bulletin*, **70**, 1063–1085.

Erickson, D. J., and B. E. Eaton (1993), Global biogeochemical cycling estimates with CZCS satellite data and general-circulation models. *Geophys. Res. Lett.*, **20**(8), 683–686.

Fiedler, P. C. (1983), Satellite remote sensing of the habitat of spawning anchovy in the southern California Bight. *CalCOFI Rep.*, **24**, 202–209.

Fiedler, P. C., and H. J. Bernard (1987), Tuna aggregation and feeding near fronts observed in satellite imagery. *Cont. Shelf Res.*, **7**(8), 871–881.

Fiedler, P. C., G. B. Smith, and R. M. Laurs (1984), Fisheries applications of satellite data in the eastern North Pacific. *Mar. Fish. Rev.*, **46**(3), 1–13.

Friedrichs, M. A. M., M.-E. Carr, R. Barber, M. Scardi, D. Antoine, R. A. Armstrong, I. Asanuma, M. J. Behrenfeld, E. T. Buitenhuis, F. Chai *et al.* (2009), Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean. *J. Marine Systems*, **76**(1/2), 113–133.

Frouin, R., and B. Cherlock (1992), A technique for global monitoring of net solar irradiance at the ocean surface, Part I: Model. *J. Appl. Meteorol.*, **31**, 1056–1066.

Frouin, R., C. Gautier, K. Katsaros, and R. Lind (1988), A comparison of satellite and empirical formula techniques for estimating insolation over the oceans. *J. Appl. Meteorol.*, **27**, 1016.

Frouin, R., D. W. Lingner, C. Gautier, K. S. Baker, and R. C. Smith (1989), A simple analytical formula to compute clear sky total and photosynthetically available solar irradiance at the ocean surface. *J. Geophys. Res.*, **94**, 9731–9742.

Frouin, R., B. Franz, and M. Wang (2001), *Algorithm to Estimate PAR from SeaWiFS Data, Version 1.2: Documentation*. NASA GSFC, available at *http://oceancolor.gsfc.nasa.gov/DOCS/seawifs_par_wfigs.pdf* (last accessed August 3, 2008).

Frouin, R., B. Franz, and P. J. Werdell (2003), The SeaWiFS PAR product. In: S. B. Hooker and E. R. Firestone (Eds.), *Algorithm Updates for the Fourth SeaWiFS Data Reprocessing* (NASA Technical Memorandum, 2000-206892, vol. 22). NASA-GSFC, Greenbelt, MD.

Garcia-Gorriz, E., and M. E. Carr (1999), The climatological annual cycle of satellite-derived phytoplankton pigments in the Alboran Sea. *Geophys. Res. Lett.*, **26**(19), 2985–2988.

Gautier, C. (1995), Remote sensing of surface solar radiation flux and PAR over the ocean from satellite observations. In: M. Ikeda and F. W. Dobson (Eds.), *Oceanographic Applications of Remote Sensing* (pp. 271–290). CRC Press, Boca Raton, FL.

Gautier, C., G. Diak, and S. Masse (1980), A simple physical model to estimate incident solar radiation at the surface from GOES satellite data. *J. Appl. Meteorol.*, **19**, 1005.

Glynn, P. W. (1996), Coral reef bleaching: Facts, hypotheses and implications. *Global Change Biology*, **2**, 495–509.

Glynn, P. W., and L. D'Croz (1990), Experimental evidence for high temperature stress as the cause of El Niño coincident coral mortality. *Coral Reefs*, **8**, 181–191.

Gomes, H. R., J. I. Goes, and T. Saino (2000). Influence of physical processes and freshwater discharge on the seasonality of phytoplankton regime in the Bay of Bengal. *Cont. Shelf Res.*, **20**(3), 313–330.

Gonzalez, N. M., F. E. Muller-Karger, S. C. Estrada, R. P. de los Reyes, I. V. del Rio, P. C. Perez, and I. M. Arenal (2000), Near-surface phytoplankton distribution in the western Intra-Americas Sea: The influence of El Niño and weather events. *J. Geophys. Res.*, **105**(C6), 14029–14043.

Gordon, H. R., G. C. Boynton, W. M. Balch, S. B. Groom, D. S. Harbour, and T. J. Smyth (2001), Retrieval of coccolithophore calcite concentration from SeaWiFS imagery. *Geophys. Res. Lett.*, **28**(8), 1587–1590.

Green, E. P. (2000). Satellite and airborne sensors useful in coastal applications. In: E. P. Green, P. J. Mumby, A. J. Edwards, and C. D. Clark (Eds.), *Remote Sensing Handbook for Tropical Coastal Management* (pp. 41-56) UNESCO, Paris.

Green, E. P., and P. J. Mumby (2000), Mapping mangroves. In: E. P. Green, P. J. Mumby, A. J. Edwards, and C. D. Clark (Eds.), *Remote Sensing Handbook for Tropical Coastal Management* (pp. 183–198). UNESCO, Paris.

Green, E. P., P. J. Mumby, A. J. Edwards, and C. D. Clark (1996), A review of remote senisng for the assessment and management of tropical coastal resources. *Coastal Management*, **24**, 1–40.

Green, E. P., C. D. Clark, and A. J. Edwards (2000a), Image classification and habitat mapping. In: E. P. Green, P. J. Mumby, A. J. Edwards, and C. D. Clark (Eds.), *Remote Sensing Handbook for Tropical Coastal Management* (pp. 141–154). UNESCO, Paris.

Green, E. P., P. J. Mumby, A. J. Edwards, and C. D. Clark (Eds.) (2000b), *Remote Sensing Handbook for Tropical Coastal Management* (edited by A. J. Edwards, 316 pp.). Coastal Management Sourcebooks, UNESCO, Paris.

Gregg, W. W., and K. L. Carder (1990), A simple spectral solar irradiance model for cloudless maritime atmospheres. *Limnology and Oceanography*, **35**, 1657–1675.

Hardman-Mountford, N. J., T. Hirata, K. A. Richardson, and J. Aiken (2008). An objective methodology for the classification of ecological pattern into biomes and provinces for the pelagic ocean. *Rem. Sens. Environ.*, **112**(208), 3341–3352. Hemmings, J. C. P., R. M. Barciela, and M. J. Bell (2008), Ocean colour data assimilation with material

conservation for improving model estimates of air–sea CO_2 flux. *J. Marine Res.*, **66**, 87–126.

Hemmings, J. C. P., M. A. Srokosz, P. Challenor, and M. J. R. Fasham (2003). Assimilating satellite ocean colour observations into oceanic ecosystem models. *Phil. Trans. Roy. Soc. Lond. A*, **361**(1802), 33–39.

Hemmings, J. C. P., M. A. Srokosz, P. Challenor, and M. J. R. Fasham (2004). Split-domain calibration of an ecosystem model using satellite ocean colour data. *J. Marine Sys.*, **50**(3/4), 141–179.

Henson, S. A., R. Sanders, J. T. Allen, I. S. Robinson, and L. Brown (2003), Seasonal constraints on the estimation of new production from space using temperature–nitrate relationships. *Geophys. Res. Letters*, **30**(17), 1912, doi: 10.1029/2003GL017982.

Henson, S. A., I. S. Robinson, J. T. Allen, and J. J. Waniek (2006), Effect of meteorological conditions on interanual variability in timing and magnitude of the spring bloom in the Irminger Basin, North Atlantic. *Deep-Sea Res.*, **53**, 1601–1615.

Herron, R. C., T. D. Leming, and J. Li (1989), Satellite-detected fronts and butterfish aggregations in the northeastern Gulf of Mexico. *Cont. Shelf Res.*, **9**(6), 569–588.

Hoegh-Guldberg, O. (1999) Climate change, coral bleaching and the future of the world's coral reefs. *Mar. Freshwater Res.*, **50**, 839–866.

Holligan, P. M., M. Viollier, C. Dupouy, and J. Aiken (1983), Satellite studies on the distributions of chlorophyll and dinoflagellate blooms in the western English Channel. *Cont. Shelf Res.*, **2**, 81–96.

Iglesias-Rodríguez, M. D., C. Brown, S. C. Doney, J. Kleypas, D. Kolber, and Z. Kolber (2002), Representing key phytoplankton functional groups in ocean carbon cycle models: Coccolithophorids. *Global Biogeochemical Cycles*, **16**(1100), doi: 10.1029/2001GB001454.

Iglesias-Rodríguez, M. D., P. Halloran, R. E. M. Rickaby, I. R. Hall, E. Colmenero-Hidalgo, J. R. Gittins, D. R. H. Green, T. Tyrell, S. J. Gibbs, P. von Dassow *et al.* (2008), Phytoplankton calcification in a high CO_2 world. *Science*, **320**, 336–340.

IOCCG (2006), Remote sensing of inherent optical properties: Fundamentals, tests of algorithms and applications. In: Z. P. Lee (Ed.), *Reports of the International Ocean-Colour Coordinating Group* (No. 5, 126 pp.). IOCCG, Dartmouth, Canada.

IOCCG (2008), Why ocean colour? The societal benefits of ocean-colour technology. In: T. Platt, N. Hoepffner, V. Stuart, and C. Brown (Eds.), *Reports of the International Ocean-Colour Coordinating Group* (No. 7, 141 pp.). IOCCG, Dartmouth, Canada.

Ishizaka, J., E. Siswanto, T. Itoh, H. Murakami, Y. Yamaguchi, N. Horimoto, T. Ishimaru, S. Hashimoto, and T. Saino (2007), Verification of vertically generalized production model and estimation of primary production in Sagami Bay. *Japan. J. Oceanography*, **63**(3), 517–524.

Joint, I., and S. B. Groom (2000), Estimation of phytoplankton production from space: current status and future potential of satellite remote sensing. *J. Exp. Mar. Biol. Ecol.*, **250**, 233–255.

Kahru, M., and B. G. Mitchell (2000), Influence of the 1997–98 El Niño on the surface chlorophyll in the California Current. *Geophys. Res. Lett.*, **27**(18), 2937–2940.

Kiefer, D. A., and B. G. Mitchell (1983), A simple steady-state description of phytoplankton growth based on absorption cross section and quantum efficiency. *Limnology and Oceanography*, **28**, 770–776.

Kirk, J. T. O. (1994), *Light and Photosynthesis in Aquatic Ecosystems* (Second Edition). Cambridge University Press, Cambridge, U.K.

Lasker, R., J. Peláez, and R. M. Laurs (1981), The use of satellite infrared imagery for describing ocean processes in relation to spawning of the northern anchovy (*Engraulis mordax*). *Rem. Sens. Environ.*, **11**, 439–453.

Laurs, R. M., and J. T. Brucks (1985), Living marine resources applications. In B. Saltzman (Ed.), *Satellite Oceanic Remote Sensing* (pp. 419–452). Academic Press, London.

Laurs, R. M., and R. J. Lynn (1991). North Pacific albacore ecology and oceanography. In: J. A. Wetherall (Ed.), *Biology, Oceanography and Fisheries of the North Pacific Transition Zone and Subarctic Frontal Zone* (NOAA Technical Report NMFS 105, pp. 69–87). National Oceanic and Atmospheric Administration, Silver Springs, MD.

Laurs, R. M., P. C. Fiedler, and D. R. Montgomery (1984), Albacore tuna catch distributions relative to environmental features observed from satellite. *Deep-Sea Res.*, **31**(9), 1085–1099.

Leonard, C. L., R. R. Bidigare, M. P. Seki, and J. J. Polovina (2001), Interannual mesoscale physical and biological variability in the North Pacific Central Gyre. *Progress in Oceanography*, **49**, 227–244.

Lévy, M., Y. Lehahn, J.-M. André, L. Mémery, H. Loisel, and E. Heifetz (2005), Production regimes in the northeast Atlantic: A study based on Sea-viewing Wide Field-of-view Sensor (SeaWiFS) chlorophyll and ocean general circulation model mixed layer depth. *J. Geophys. Res.*, **110**(C07S10), doi: 10.1029/2004JC002771.

Lillesand, T. M., and R. W. Kiefer (1999), *Remote Sensing and Image Interpretation* (Fourth Edition, 736 pp.). John Wiley & Sons, New York.

Longhurst, A. (1998), *Ecological Geography of the Sea*. Academic Press, San Diego, CA.

Longhurst, A., S. Sathyendranath, T. Platt, and C. M. Caverhill (1995), An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.*, **17**(6), 1245–1271.

Lyzenga, D. R. (1981), Remotre sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and Landsat data. *Int. J. Remote Sensing*, **2**, 71–82.

Manizza, M., C. LeQuéré, A. J. Watson, and E. T. Buitenhuis (2005), Bio-optical feedbacks among phytoplankton, upper ocean physics and sea-ice in a global model. *Geophys. Res. Lett.*, **32**(L05603), doi: 10.1029/2004GL020778.

Mather, P. M. (1999), *Computer Processing of Remotely-sensed Images: An Introduction* (Second Edition, 292 pp.). John Wiley & Sons, Chichester, U.K.

Maul, G. A., F. Williams, M. Roffer, and F. M. Souza (1984). Remotely sensed oceanographic patterns and variability of bluefin tuna catch in the Gulf of Mexico. *Oceanol. Acta*, **7**(4), 469–479.

McClain, C. R. (2009), A decade of satellite ocean color observations. *Annu. Rev. Mar. Sci.*, **1**, 19–42.

Miller, A. J., M. A. Alexander, G. J. Boer, F. Chai, K. Denman, D. J. Erickson III, R. Frouin, A. J. Gabric, E. A. Laws, M. R. Lewis *et al.* (2003), Potential feedbacks between Pacific Ocean ecosystems and interdecadal climate variations. *Bull. Am. Meteorol. Soc.*, **84**(5), 617–633.

Mitchell, B. G. (1994) Coastal zone color scanner retrospective. *J. Geophys. Res.*, **99**, 7291–7292.

Montgomery, D. R. (1981) Commercial applications of satellite oceanography. *Oceanus*, **24**(3), 56–65.

Montgomery, D. R., R. E. Wittenberg-Fay, and R. W. Austin (1986), The applications of satellite-derived ocean color products to commercial fishing operations. *Mar. Tech. Soc. J.*, **20**(2), 72–86.

Moore, J. K., M. R. Abbott, J. G. Richman, W. O. Smith, T. J. Cowles, K. H. Coale, W. D. Gardner, and R. T. Barber (1999), SeaWiFS satellite ocean color data from the Southern Ocean. *Geophys. Res. Lett.*, **26**(10), 1465–1468.

Morel, A. (1991), Light and marine photosynthesis: A spectral model with geochemical and climatological implications. *Progress in Oceanography*, **26**, 263–306.

Morel, A., and J.-F. Berthon (1989), Surface pigments, algal biomass profiles and potential production of the euphotic layer: Relationships reinvestigated in view of remote sensing applications. *Limnology and Oceanography*, **34**, 1545–1562.

Mumby, P. J., and A. J. Edwards (2000), Water column correction techniques. In: E. P. Green, P. J. Mumby, A. J. Edwards, and C. D. Clark (Eds.), *Remote Sensing Handbook for Tropical Coastal Management* (pp. 121–128). UNESCO, Paris.

Mumby, P. J., and E. P. Green (2000), Mapping coral reefs and macroalgae. In: E. P. Green, P. J. Mumby, A. J. Edwards, and C. D. Clark (Eds.), *Remote Sensing Handbook for Tropical Coastal Management* (Chap. 14, p. 155). UNESCO, Paris.

Mumby, P. J., E. P. Green, A. J. Edwards, and C. D. Clark (1997), Measurement of sea-grass standing crop using satellite and digital airborne remote sensing. *Mar. Ecol. Prog. Ser.*, **159**, 51–60.

Mumby, P. J., E. P. Green, A. J. Edwards, and C. D. Clark (1999), The cost-effectiveness of remote sensing for tropical coastal resources assessment and management. *J. Environmental Management*, **55**, 157–166.

Murtugudde, R. G., R. S. Signorini, J. R. Christian, A. J. Busalacchi, C. R. McClain, and J. Picaut (1999), Ocean color variability of the tropical Indo-Pacific basin observed by SeaWiFS during 1997–98. *J. Geophys. Res.*, **104**, 18351–18366.

Myers, D. G., and P. T. Hick (1990), An application of satellite-derived sea surface temperature data to the Australian fishing industry in near-real time. *Int. J. Remote Sensing*, **11**(11), 2103–2112.

Nair, A., S. Sathyendranath, T. Platt, J. Morales, V. Stuart, M.-H. Forget, E. Devred, and H. Boumain (2008), Remote sensing of phytoplankton functional types. *Rem. Sens. Environ.*, **112**, 3366–3375.

Narain, A., R. M. Dwivedi, H. U. Solanki, B. Kumari, and N. Chaturvedi (1990), The use of NOAA-AVHRR data in fisheries exploration in the Indian EEZ. Paper presented at *Proc. Sem. Remote Sensing for Marine Fisheries Studies, Beijing, China*. Economic and Social Commission for Asia and the Pacific/United Nations Development Program (ESCAP/UNDP), pp. 226–232.

Nayak, S., H. U. Solanki, and R. M. Dwivedi (2003), Utilization of IRS P4 ocean colour data for potential fishing zone: A cost benefit analysis. *Ind. J. Mar. Sci.*, **32**, 244–248.

Njoku, E. G., T. P. Barnett, R. M. Laurs, and A. C. Vastano (1985), Advances in satellite sea surface temperature measurement and oceanographic applications. *J. Geophys. Res.*, **90**(C6), 11573–11586.

Pauly, D., and V. Christensen (1995), Primary production required to sustain global fisheries. *Nature*, **374**, 255–257.

Pinker, R. T., and I. Laszlo (1992), Global distribution of photosynthetically active radiation as observed from satellites. *J. Climate*, **5**, 56–65.

Platt, T., and S. Sathyendranath (1988). Oceanic primary production: Estimation by remote sensing at local and regional scales. *Science*, **241**, 1613–1620.

Platt, T., and S. Sathyendranath (1993), Estimators of primary production for interpretation of remotely sensed data on ocean color. *J. Geophys. Res.*, **98**(C8), 14561–14576.

Platt, T. and S. Sathyendranath (1999), Spatial structure of pelagic ecosystem processes in the global ocean. *Ecosystems*, **2**, 384–394.

Platt, T., S. Sathyendranath, C. M. Caverhill, and M. R. Lewis (1988), Ocean primary production and available light: Further algorithms for remote sensing. *Deep-Sea Research*, **35**(6), 855–879.

Platt, T., P. Jahauri, and S. Sathyendranath (1992), The importance and measurement of new production. In: P. G. Falkowski and A. D. Woodhead (Eds.), *Primary Productivity and Biogeochemical Cycles in the Sea* (pp. 273–284). Plenum Press, New York.

Platt, T., S. Sathyendranath, and A. Longhurst (1995), Remote sensing of primary production in the ocean: Promise and fulfilment. *Phil. Trans. Roy. Soc. Lond. B*, **348**, 191–202.

Platt, T., C. Fuentes-Yaco, and K. T. Frank (2003), Spring algal bloom and larval fish survival. *Nature*, **423**, 398–399.

Podestá, G. P., J. A. Browder, and J. J. Hoey (1993), Exploring the association between swordfish catch rates and thermal fronts on US longline grounds in the western North Atlantic. *Cont. Shelf Res.*, **13**(2/3), 253–277.

Polovina, J. J. (2005), Climate variation, regime shifts, and implications for sustainable fisheries. *Bull. Mar. Sci.*, **76**, 233–244.

Polovina, J. J., D. R. Kobayashi, D. M. Parker, M. P. Seki, and G. H. Balazs (2000), Turtles on the edge: Movement of loggerhead turtles (*Caretta caretta*) along oceanic fronts, spanning longline fishing grounds in the central North Pacific, 1997–1998. *Fish. Oceanogr.*, **9**(1), 71–82.

Polovina, J. J., E. Howell, D. R. Kobayashi, and M. P. Seki (2001), The transition zone chlorophyll front, a dynamic global feature defining migration and forage habitat for marine resources. *Progress in Oceanography*, **49**, 469–483.

Polovina, J. J., G. H. Balazs, E. A. Howell, D. M. Parker, M. P. Seki, and P. H. Dutton (2004). Forage and migration habitat of loggerhead (*Caretta caretta*) and olive ridley (*Lepidochelys olivacea*) sea turtles in the central North Pacific Ocean. *Fish. Oceanogr.*, **13**, 36–51.

Polovina, J. J., E. A. Howell, and M. Abecassis (2008), Ocean's least productive waters are expanding. *Geophys. Res. Lett.*, **35**(L03618), doi: 10.1029/2007GL031745.

Purkiss, S. (2005) A "reef-up" approach to classifying coral habitats from IKONOS imagery. *IEEE Trans. Geoscience Rem. Sens.*, **43**(6), 1375–1390.

Purkiss, S., J. A. M. Kenter, E. K. Oikonomou, and I. S. Robinson (2002), High resolution ground verification, cluster analysis and optical model of reef substrate coverage from Landsat TM imagery (Red Sea, Egypt). *Int. J. Remote Sensing*, **23**(8), 1677–1698.

Richards, W. J., T. D. Leming, M. F. McGowan, J. T. Lamkin, and S. Kelley-Fraga (1989), Distribution of fish larvae in relation to hydrographic features of the Loop Current boundary in the Gulf of Mexico. *Rapp. P.-v. Réun. Cons. Int. Explor. Mer.*, **191**, 169–176.

Richardson, A. J., and D. S. Schoeman (2004), Climate impact on plankton ecosystems in the Northeast Atlantic. *Science*, **305**, 1609–1212.

Riebesell, U., I. Zondervan, B. Rost, P. D. Tortell, R. E. Zeebe, and F. M. M. Morel (2000), Reduced calcification of marine plankton in response to increased atmospheric CO_2. *Nature*, **407**(September 21), 364–367.

Robblee, M. B., T. R. Barber, P. R. Carlson, Jr., M. J. Durako, J. W. Fourqurean, L. K. Muehlstein, D. Porter, L. A. Yarbro, R. T. Zieman, and J. C. Zieman (1991), Mass mortality of the tropical seagrass *Thalassia testudinum* in Florida Bay (USA). *Marine Ecol. Prog. Ser.*, **71**, 297–299.

Robinson, I. S. (2004) *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Robinson, I. S., D. Antoine, M. Darecki, P. Gorringe, L. Pettersson, K. Ruddick, R. Santoleri, H. Siegel, P. Vincent, M. R. Wernand *et al.* (2008), *Remote Sensing of Shelf Sea Ecosystems: State of the Art and Perspectives* (edited by N. Connolly, Marine Board Position Paper No. 12., 60 pp.). European Science Foundation Marine Board, Ostend, Belgium.

Saichun, T., and S. Guangyu (2006), Satellite-derived primary productivity and its spatial and temporal variability in the China seas. *J. Geograph. Sci.*, **16**(4), 447–457.

Santos, A. M. P. (2000) Fisheries oceanography using satellite and airborne remote sensing methods: A review. *Fisheries Research*, **49**, 1–20.

Santos, A. M. P., and A. F. G. Fiúza (1992), Supporting the Portuguese fisheries with satellites. Paper presented at *Proc. European International Space Year Conference 1992 on "Space in the Service of the Changing Earth", Munich, Germany* (ESA SP-341, Part 2, pp. 663–668). European Space Agency, Noordwijk, The Netherlands.

Saraceno, M., C. Provost, A. R. Piola, J. Bava, and A. Gagliardini (2004), Brazil Malvinas Frontal System as seen from 9 years of advanced very high resolution radiometer data. *J. Geophys. Res.*, **109**(C05027), doi: 10.1029/2003JC002127.

Saraceno, M., C. Provost, and A. R. Piola (2005), On the relationship between satellite-retrieved surface temperature fronts and chlorophyll a in the western South Atlantic. *J. Geophys. Res.*, **110**(C11016), doi: 10.1029/2004JC002736.

Sathyendranath, S., T. Platt, C. M. Caverhill, R. E. Warnock, and M. R. Lewis (1989), Remote sensing of oceanic primary production: Computations using a spectral model. *Deep-Sea Research*, **36**(3), 431–453.

Sathyendranath, S., T. Platt, E. P. W. Horne, W. G. Harrison, O. Ulloa, R. Outerbridge, and N. Hoepffner (1991). Estimation of new production in the ocean by compound remote sensing. *Nature*, **353**, 129–133.

Sathyendranath, S., A. Longhurst, C. M. Caverhill, and T. Platt (1995), Regionally and seasonally differentiated primary production in the North Atlantic. *Deep-Sea Research*, **42**(10), 1773–1802.

Sathyendranath, S., L. Watts, E. Devred, T. Platt, C. Caverhill, and H. Maass (2004), Discrimination of diatoms from other phytoplankton using ocean-colour data. *Mar. Ecol. Prog. Ser.*, **272**, 59–68.

Schiffer, R. A., and W. B. Rossow (1983), The International Satellite Cloud Climatology Project (ISCCP): The first project of the World Climate Research Program. *Bull. Am. Meteorol. Soc.*, **64**, 779–784.

Schiffer, R. A., and W. B. Rossow (1985), ISCCP global radiance data set : A new resource for climate research. *Bull. Am. Meteorol. Soc.*, **66**, 1498–1505.

Signorini, S. R., R. G. Murtugudde, C. R. McClain, J. R. Christian, J. Picaut, and A. J. Busalacchi (1999). Biological and physical signatures in the tropical and subtropical Atlantic. *J. Geophys. Res.*, **104**(C8), 18367–18382.

Solanki, H. U., R. M. Dwivedi, S. Nayak, V. S. Somvanshi, D. K. Gulati, and S. K. Pattnayak (2003), Fishery forecast using OCM chlorophyll concentration and AVHRR SST: Validation results off Gujarat coast, India. *Int. J. Remote Sensing*, **24**, 3691–3699.

Subrahmanyam, B., K. Ueyoshi, and J. M. Morrison (2008), Sensitivity of the Indian Ocean circulation to phytoplankton forcing using an ocean model. *Rem. Sens. Environ.*, **112**, 1488–1496.

Subramaniam, A., C. W. Brown, R. R. Hood, E. J. Carpenter, and D. G. Capone (2002), Detecting *Trichodesmium* blooms in SeaWiFS imagery. *Deep-Sea Res. II*, **49**, 107–121.

Tameishi, H., O. Honda, T. Kohguti, S. Fujita, and K. Saitoh (1992), Application of satellite imageries data to fisheries in Japan. Paper presented at *Proc. European International*

Space Year Conference 1992 on "Space in the Service of the Changing Earth", Munich, Germany (ESA SP-341, Part 2, pp. 669–674). European Space Agency, Noordwijk, The Netherlands.

Thomas, A., and P. T. Strub (2001), Cross-shelf phytoplankton pigment variability in the California Current. *Cont. Shelf Res.*, **21**, 1157–1190.

Tyrell, T., P. M. Holligan, and C. D. Mobley (1999), Optical impacts of oceanic coccolithophore blooms. *J. Geophys. Res.*, **104**, 3223–3241.

Ueyama, R., and B. C. Monger (2005), Wind-induced modulation of seasonal phytoplankton blooms in the North Atlantic derived from satellite observations. *Limnology and Oceanography*, **50**(6), 1820–1829.

Watts, L., S. Sathyendranath, C. Caverhill, H. Maass, T. Platt, and N. J. P. Owens (1999), Modelling new production in the northwest Indian Ocean region. *Mar. Ecol. Prog. Ser.*, **183**, 1–12.

Weeks, S., B. Currie, and A. Bakun (2002), Massive emissions of toxic gas in the Atlantic. *Nature*, **415**, 493–494.

Wilson, C. (2003), Late summer chlorophyll blooms in the oligotrophic North Pacific Subtropical Gyre. *Geophys. Res. Lett.*, **30**(18), 1942, doi: 10.1029/2003GL017770.

Yamada, K., J. Ishizaka, and H. Nagata (2005), Spatial and temporal variability of satellite primary production in the Japan Sea from 1998 to 2002. *J. Oceanography*, **61**(5), 857–869.

Yin, K. D., P. J. Harrison, J. Chen, W. Huang, and P. Y. Qian (1999), Red tides during spring 1998 in Hong Kong: Is El Nino responsible? *Mar. Ecol. Prog. Ser.*, **187**, 289–294.

Yoder, J. A., J. E. O'Reilly, A. H. Barnard, T. S. Moore, and C. M. Ruhsam (2001), Variability in coastal zone color scanner (CZCS) chlorophyll imagery of ocean margin waters off the US East Coast. *Cont. Shelf Res.*, **21**, 1191–1218.

Zainal, A. J. M., D. H. Dalby, and I. S. Robinson (1993), Monitoring marine ecological changes on the east coast of Bahrain with Landsat TM. *Photogram. Eng. and Remote Sensing*, **59**, 415–421.

Zainuddin, M., S.-I. Saitoh, and K. Saitoh (2004), Detection of potential fishing ground for albacore tuna using synoptic measurements of ocean color and thermal remote sensing in the northwestern North Pacific. *Geophys. Res. Lett.*, **31**(20), L20311.

8

Ocean surface waves

8.1 INTRODUCTION

This is the first of three chapters that investigate what satellite remote-sensing data can tell us about dynamical processes and phenomena occurring at the air–sea interface. The subject here is the high-frequency oscillatory motion of the sea surface which we experience as ocean surface waves. Subsequently Chapter 9 shows how the distribution of the wind field over the ocean is observed by remote-sensing measurements of sea surface roughness, and Chapter 10 examines what we can learn from satellite data about the fluxes of momentum, heat, and gases between the liquid ocean and the gaseous atmosphere.

Nearly all the diverse methods of ocean remote sensing summarized in Chapter 2 are affected in one way or another by the shape, position, or movement of the sea surface. Sea state and surface roughness are therefore important factors that need to be known and taken into account when interpreting satellite ocean data for many of the other applications described elsewhere in this book. However, the perspective of this chapter is to focus on measurements of ocean waves that can be directly obtained from satellite sensors, and then to review the oceanographic applications of satellite-derived wave data.

The term "sea surface waves" is loosely used in ocean science to refer to phenomena spanning a wide range of lengthscales and timescales, from the short capillary ripples a few millimeters long that appear when a light breath of wind disturbs a calm sea, to the swell with a wavelength of several hundred meters and a height of several meters that can propagate across an ocean from a distant storm. This chapter will concentrate on gravity waves longer than tens of meters, having a period of several seconds up to 20 seconds for the longest swell. Although waves shorter than this are of special interest to the satellite oceanographer because of what they reveal about the phenomena which modulate them (see chapters 9 and 10 of *MTOFS*), they are not so important to most users of the sea. For mariners

responsible for the safety and stability of a vessel, for engineers constructing an offshore platform, or for the coastal protection manager concerned with beach erosion, the waves that matter most are those that have a vertical amplitude of several meters and transport sufficient energy to cause damage. These are also the waves of most importance for oceanographic scientists because they transfer energy from the wind into the ocean where it can have diverse and important consequences, such as promoting vertical mixing, causing resuspension of shelf sea sediments, or disrupting the ecology of normally sheltered embayments. Ultimately, the value of satellite data as a source of information about ocean waves must be judged by the impact made on operational and scientific users of wave data. That is the intended scope of this chapter.

Section 8.2 introduces the tools and techniques of the subject, beginning with an explanation of the parameters used to characterize and quantify waves, followed by an outline of the methods of satellite oceanography that are used to measure those parameters. While nearly all the remote-sensing techniques summarized in Chapter 2 are influenced by waves on the sea surface, most sensors do not record a signal that can be analyzed to retrieve practically important wave parameters. However, there are two types of instrument, the radar altimeter and the synthetic aperture radar (SAR), that are routinely used to obtain ocean wave information from satellites, and a third remote-sensing device, the wave spectrometer, which has potential but has not yet been proven as a satellite system. Section 8.3 details actual sensors and systems used to measure wave parameters, and the data products that are routinely delivered by these satellite instruments.

The rest of the chapter explores the various applications of wave measurements provided by satellites, considered generally in Section 8.4 and then specifically in relation to their use in ocean wave forecast models in Section 8.5 and in improving wave statistics and climatology in Section 8.6.

8.2 MEASURING OCEAN WAVES—PRINCIPLES

8.2.1 Characterizing ocean waves in terms of measurable parameters

The shape and motion of the sea surface at the lengthscales and timescales of interest respond to forcing by the wind which destabilizes and tilts the air–sea interface, and are constrained by the restoring force of gravity which always acts to pull a sloping sea surface towards the horizontal. On a beach or in a boat we experience sea waves as an almost random motion, although there is also an underlying regularity. How can we characterize such a phenomenon scientifically and define it mathematically? After nearly a century of scientific study hydrodynamicists and engineers have answered these and many other questions about ocean waves (see, e.g., LeBlond, 2002; Holthuijsen, 2007) although we should always be on the lookout for fresh insights from the novel perspective provided by satellite-derived wave data.

In the absence of wind forcing, a perturbation of the sea surface, $\zeta(x, y, t)$, propagates in space and time in a regular, predictable, wave-like manner. In deep

water a disturbance which oscillates in time with a *period* of T seconds is found to have a *wavelength* of λ *meters*, where

$$\lambda = \frac{gT^2}{2\pi} \tag{8.1}$$

and g is acceleration due to gravity (about $9.81\,\mathrm{m/s^2}$). An alternative way of writing this dispersion relation, which connects length and time variability, is

$$\omega^2 = gk \tag{8.2}$$

in terms of frequency $\omega = 2\pi/T$ radians/s and wavenumber $k = 2\pi/\lambda$ radians/m. The speed at which the wave profile propagates (the apparent speed of troughs and crests) is called the *phase speed* and is given by

$$V_{\mathrm{ph}} = \lambda/T = \omega/k = g/\omega. \tag{8.3}$$

Note that the speed depends on the wavelength or frequency of a particular train of waves. Such wave propagation is described as being *dispersive*. Longer period (lower frequency) waves have longer wavelengths and travel faster than short-period waves. It is important for understanding wave forecasting to note that the energy associated with a particular wavelength or frequency propagates at the *group velocity*, V_{gr}, which for surface waves is exactly half the phase speed. In shallow seas and over beaches, where the water depth, h, is less than half the wavelength, wave propagation is influenced by depth. Where $h < \lambda/20$ the wave speed is controlled entirely by depth and becomes nondispersive, so $V_{\mathrm{ph}} = V_{\mathrm{gr}} = \sqrt{(gh)}$. See section 9.3.3 of *MTOFS* for further discussion of surface wave theory as it impacts remote sensing, texts such as LeBlond and Mysak (1978) for a full treatment of wave theory, and the chapter on waves in Stewart (2008) for a clear introduction.

Equation (8.1) characterizes swell, which consists of a train of almost regular parallel waves of sinusoidal profile which have propagated a long way from the storm that generated them. Swell at a particular location is almost monochromatic, having waves with a single, dominant period that reduces gradually over time since shorter waves, being slower, take longer to travel from the source region.

However, in most parts of the sea, as well as any swell that may be present, the wind is continually creating waves locally and surface height and slope at any location is neither precisely regular nor predictable. To completely describe the wave field would require the height to be known at all places at all times—an impossibility given the random nature of surface displacement. Instead the surface wave field can be characterized by a statistical property of surface height or slope known as the *wave spectrum*, S. Another parameter known as *significant wave height*, H_S, is also used to characterize the height between crests and troughs of larger waves.

8.2.2 Wave energy and spectra

The one-dimensional frequency wave spectrum, $S(\omega)$, represents physically how energy per unit frequency is distributed across the whole range of frequencies found in the wave field. Thus the integral of $S(\omega)$ over all frequencies equals the

total energy of the wave field. Note that the energy of a single, monochromatic wave field is proportional to the square of the wave amplitude. Thus although the wave spectrum cannot tell us precisely what the surface height will be at a particular location at a particular time, it does contain very useful information about energy in the total wave field, and how it is distributed between lower and higher frequency waves. If the energy of waves is further partitioned according to the horizontal propagation direction, θ, then $S(\omega, \theta)$ is the directional frequency spectrum. If instead the directionality of waves is expressed by representing the distribution of their energy in wavenumber space, where the wavenumber, $\mathbf{k}$, is a vector (k_x, k_y) in two-dimensional space, then the result is the directional wavenumber spectrum $S(\mathbf{k})$.

The advantage of using the spectrum for parameterizing waves is that the knowledge it provides about wave energy is important for many applications. Moreover if the directional spectrum is known the flow of energy to other parts of the ocean can be predicted into the future for wave-forecasting purposes.

Typical *in situ* measurements of waves are derived from waverider buoys (hereafter wave buoys) which record how the height or acceleration of the buoy varies with time. Analysis of these data allows the simple frequency spectrum, $S(\omega)$, to be evaluated. If the pitch and roll of the buoy are also recorded then the directional properties of the spectrum $S(\omega, \theta)$ can be estimated. There are no simple buoy measurements that can presently be used to obtain the wavenumber spectrum, other than estimating it by applying the dispersion relation (8.2) to the directional frequency spectrum.

Wave buoys are able to monitor the evolution of wave spectra over time as the sea state grows in response to a storm and then decays afterwards, but each buoy is limited to sampling at a single location. Therefore the spatial variability of the wave spectrum cannot be clearly observed, and isolated buoy measurements cannot be considered to provide samples representative of conditions over a wider area. A single buoy may miss the highest waves, or may by chance be located where waves are focused and have a higher amplitude than typical for the region. This is the kind of circumstance where satellite sensors have, in principle, the capacity for spatially detailed snapshot sampling over quite a wide region. If the spatial variability of directional wave spectra could be observed this would bring a valuable new perspective to the understanding of wave growth, propagation, and decay processes, and to operational prediction of ocean waves. At the very least, knowledge of the spatial variability of the amplitude and directionality of the wave field would qualify the extent to which isolated buoy samples can be considered to be representative.

8.2.3 Significant wave height

Significant wave height was originally a subjective measure of the trough-to-peak height of waves observed by mariners from the bridge of a ship. It is now objectively defined in relation to a measured time series of sea surface height at a point. The height between turning points in the time series (i.e., the vertical distance from a trough to the next crest) are recorded and ranked; SWH is defined as the average of the highest one-third of these values. For this reason it is sometimes given the symbol

$H_{1/3}$. The choice of 1/3 recognizes the fact that the old subjective measure was influenced mostly by the larger and longer period waves that control the ship's motion and dominate the senses, and did not take into account small-amplitude, high-frequency waves riding on top of them. However, to avoid confusion with much of the remote-sensing literature, the symbol H_S is used here.

Despite being a very empirically based definition, H_S is widely used by oceanographers, ship scientists, and coastal engineers, and has become a standard engineering unit for reporting and predicting waves. Ships and offshore structures are designed for predicted extreme values of H_S. It is therefore important to be able to measure this as well as the wave spectrum. It does have an approximate relationship with a more precise statistical measure of average height which is σ_ζ, the standard deviation of sea surface elevation (Cartwright and Longuet-Higgins, 1956). The relationship is usually taken to be

$$H_S \approx 4\sigma_\zeta \tag{8.4}$$

although the constant depends on the shape of the wave spectrum and can reduce from 4 to about 3 for a very broad-band spectrum.

Without knowing the spectrum, a measurement of H_S conveys little about the character of waves other than their height. It is therefore useful to have supplementary information such as the period of dominant waves contributing to the sea state. In practice there are several different ways of defining this, including the peak period, T_P, the mean period, T_m, and the zero-crossing period, T_z. Further explanation of how these practical definitions are prescribed from wave data records can be found in the book by Tucker (1991).

Although simple descriptions of water waves assume that displacement of the surface above or below the mean level is symmetrical, in practice this is not quite so, especially for steeper waves. These tend to have peakier crests and flatter troughs. This effect is quantified by skewness, which is the third-order moment of ocean wave elevation distribution and indicates the nonlinearity of the waves in a particular situation.

Another property relevant to understanding how a radar altimeter pulse reflects from a wave field is *wave age*, the ratio of wave phase speed, as defined by (8.3), to wind friction velocity, u^*, explained in Chapter 10. Wave age is perhaps a misleading name for a dimensionless number which characterizes the stage of development of the wind–wave part of the total wave field, and has implications for energy interactions between the wind and the wave field (Jones and Toba, 2001). If V_{ph}/u^* is about 30, the sea is considered to be in equilibrium with the wind; less than this implies the wind is still forcing energy into the waves; greater than this and the reverse is likely to be happening. Wave age can also be used as an indicator of the relative contributions of swell and local wind-driven waves to the overall field.

As for wave spectra, it is difficult to be sure that H_S measured as a single point sample from a buoy or ship is representative of the wider region. Consequently there is scope for satellites to make a unique contribution if they can deliver a nearly instantaneous detailed distribution of wave height over a region. The potential for

useful applications increases further if, in addition to wave height, wave period can be estimated from satellite data. Ocean radar altimeters offer just such a possibility.

8.2.4 Measuring ocean waves from an altimeter

As mentioned in Section 2.4.5, a satellite altimeter is a nadir-pointing radar which emits short radar pulses towards the sea surface and measures the reflected echo. Several different measurements of the ocean can be made by analyzing the return echo. In this chapter we are interested in the way that the sea state in the reflection zone affects the shape of the return echo, allowing significant wave height to be estimated. Figure 8.1 illustrates the operating principle. It shows microwave energy from the emission of a single radar pulse radiating within a thin, spherical shell away from the radar and encountering a rough sea surface about 3 ms after the pulse left the altimeter.

Figure 8.1a shows the radiating pulse at three positions. A is the location when the energy encounters the highest tops of waves and is first reflected. This is the first part of reflected energy to be detected by the radar some 3 ms later (as shown in Figure 8.1b) and marks the gradual appearance of the start of the echo. A few microseconds later the leading point of the pulse has just reached mean sea level

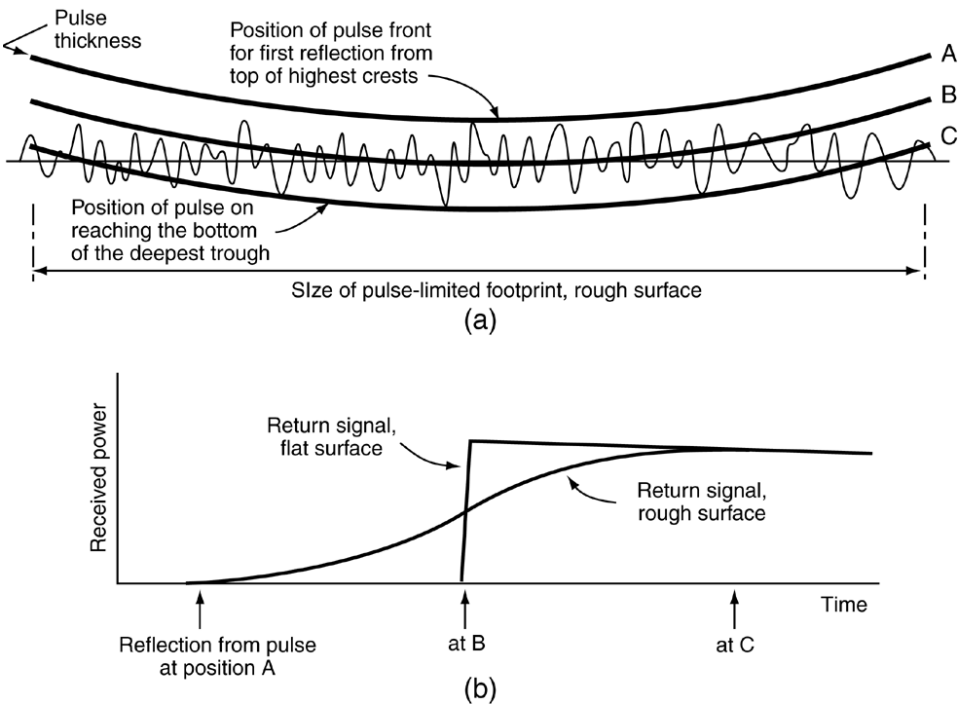


Figure 8.1. Interaction of an altimeter pulse with a rough sea surface. (a) Illuminated surface geometry. (b) The resulting shape of the reflected pulse.

(position B) and about half the energy in this part of the pulse has already encountered the sea surface and been reflected, causing the amplitude of the corresponding reflected echo to steadily increase with time (as shown in Figure 8.1b). However, at B some energy is still propagating down towards the wave troughs. Finally when the shell of emitted energy reaches position C all the energy in the part of the shell shown has been reflected. Correspondingly the echo at a time corresponding to the travel time from the sensor to C and back reaches its maximum level (as shown in Figure 8.1b). Thereafter, reflections continue from a ring-shaped region which spreads out from the point of first reflection, but which does not increase in area and so the magnitude of the echo levels out (see section 11.2.1 of *MTOFS* for a fuller explanation of the process).

The critical point in the context of wave detection is that for a sea where wave amplitude is larger, the radar pulse reaches position A earlier, and reaches C later, than for a sea where the waves are smaller. This results in an earlier start to the echo and a later time for it to reach its full magnitude in the case of larger sea waves. The timing of B is almost independent of wave amplitude and marks the mid-height of the rising echo. If the sea state decreases the reverse is true. In the extreme case of a completely flat sea, positions A, B, and C become the same and the pulse rises very sharply indeed (as also shown in Figure 8.1b). Consequently the shape (in particular the rise time) of the return echo is very sensitive to the sea state, allowing significant wave height to be estimated with a considerable degree of accuracy, irrespective of the distance between the satellite and mean sea level (the altimetry signal), and also of the amplitude of the full-height echo, which depends on short-scale roughness and hence the wind. Details of the model inversion procedure used for retrieving H_S from the average of a number of pulses is given in section 11.8.1 of *MTOFS*.

The resulting estimate of H_S corresponds to the average over the region swept out by the pulse-limited circle during the 1,000 or so pulses (depending on the system) that are averaged. The diameter of the pulse-limited circle is determined by the intersection of pulse position C (in Figure 8.1a) with mean sea level. It increases as wave height gets larger (equation 11.5 in *MTOFS*). Typically it is about 3 km for a calm sea increasing to 10 km for high seas, but also depends on the pulse width in time. For a typical altimeter system all echoes are averaged from pulses emitted during an interval of about 1 s, during which time the satellite travels about 6 km over the ground, so that cross-track and along-track footprint sizes are similar.

This approach to deriving H_S from altimeters is the method of satellite remote sensing that is presently most widely used by oceanographers requiring information about ocean waves.

As a refinement to the basic measurement, research studies have shown that it is possible to estimate some of the other wave field parameters mentioned in Section 8.2.3. Gommenginger *et al.* (2003) noted that the mean square slope of the surface is inversely proportional to the nadir-viewing radar backscatter cross-section, σ_0, measured from the amplitude of the altimeter echo, and so it can be obtained independently of wave height. Since mean slope depends on the ratio of wave height to wavelength, which (from Equation 8.1) depends on T^2, they reasoned

that T^4 should be proportional to $\sigma_0 H_S^2$. Therefore they tested empirically for a linear correlation between wave period as measured *in situ* by wave buoys and the variable $X = (\sigma_0 H_S^2)^{0.25}$ obtained entirely from a satellite altimeter. By matching satellite samples of σ_0 (from the Ku band of the TOPEX altimeter) and H_S (from the 1 Hz averaged TOPEX geophysical data record) to wave buoy data coincident within 100 km and 1 h, they confirmed the strong correlation of both T_m and T_z with X. The peak period, T_p, was less clearly correlated with X, probably because of the way it is derived from the discrete wave record whereas mean and zero-crossing periods are derived as integrated properties of the wave spectrum.

From these results, empirical algorithms were developed for retrieving T_z and T_m from altimeter data alone. The error for T_z retrieval, based on an independent set of validation matchups, was 0.8 s. Since this empirical model is based on matches to buoys dispersed throughout the World Ocean it is believed to have wide applicability although it showed some weakness in coping with seas dominated by swell where the altimeter underestimates the wave period. Quilfen *et al.* (2004) developed another wave period model. Subsequent work (Caires *et al.*, 2005) compared altimeter-derived wave periods with T_m and H_S from the ERA-40 reanalysis dataset of meteorological variables including ocean surface wind waves, and with buoy measurements independent of both. This demonstrated the global validity of the altimeter-derived wave period. After adjusting the algorithm to wind-dominated conditions (about 40% of all cases) global root-mean-square error was reduced to 0.5 s. It was also confirmed that in seas dominated by swell the altimeter still produces reliable wave period estimates useful to wave modelers, provided the wind is greater than 4 m/s. Recently a new algorithm has been produced for retrieving wave period from Ku-band altimeters, which approaches the theoretical limit of accuracy (Mackay *et al.*, 2008).

Global mapping of ocean wave skewness has also been achieved by reanalyzing echo waveforms from the Envisat RA-2 altimeter (Gomez-Enri *et al.*, 2007). When a nonlinear model is used, wave skewness is one of the outputs. As expected this shows increased skewness over the Southern Ocean and other regions at times when the waves are highest. The most direct application of this achievement is in improving the retrieval of wave height and sea surface height when the model used for retracking altimeter echo waveforms allows for skewness.

8.2.5 Observing waves with the synthetic aperture radar (SAR)

A very different approach to the measurement of ocean surface waves is possible if waves can be explicitly imaged by a high-resolution device that obtains a snapshot view of the sea surface. From the experience of looking by eye at waves on the ocean, from the deck of a ship or from a cliff top overlooking the sea, we are familiar with the spatial patterns of waves, particularly their slope, which can be clearly detected if illumination conditions are suitable. Unfortunately it is not practicable to extend this use of visible light reflection to the high vantage point and wide coverage of the viewpoint of a satellite in Earth orbit.

However, a comparable effect is possible using high-resolution imaging radar.

For the purpose of achieving the fine spatial resolution of a few tens of meters needed to detect even longer surface gravity waves and swell, the use of SARs is called for. If an imaging radar is to detect waves, there has to be a strong correlation between some aspect of the ocean surface which is tied to the phase of the wave (e.g., its height, its slope, or the currents at the surface) and the amount of microwave energy backscattered from different parts of the surface. This is the classic requirement for a radar imaging mechanism (see section 10.7 of *MTOFS*).

In fact three different imaging mechanisms have been identified by which SAR images can represent aspects of the ocean surface wave field (see section 9.3.6 of *MTOFS*). The *tilt modulation* process simply returns a stronger echo from those parts of the wave surface profile facing towards the radar than from those facing away. The *hydrodynamic modulation* process relies on the fact that small-scale (centimeter–decimeter wavelength) ripples control the magnitude of the radar echo through the Bragg mechanism when the radar is pointing obliquely at the surface. These short ripples are themselves modulated by convergent and divergent surface currents associated with longer waves. The stretching and compression of the sea surface as ocean waves pass along it generate patterns of smoother and rougher texture at Bragg wavelengths, and these patterns are locked to the phase of longer waves. The resulting patterns of light and dark on the radar image appear to correspond, in suitable circumstances, to troughs and crests of longer surface waves.

Both tilt and hydrodynamic modulation work only for ocean waves propagating with a component in the radar range direction (Alpers *et al*., 1981) (i.e., the direction in which the radar is pointing—see Figure 8.2). Waves that propagate largely in a direction cutting across the range direction (i.e., waves propagating parallel to the radar azimuth) cannot be detected at all by these mechanisms. Fortunately there is another process known as *velocity bunching* which can image these waves (Alpers and Rufenach, 1979; Alpers and Brüning, 1986). This is a byproduct of the aperture synthesis technique by which the processing of raw SAR data achieves fine resolution in the azimuth direction, although it is a nonlinear process and liable to strong distortion in high seas. Each of these mechanisms is described in more detail in section 10.9 of *MTOFS*.

Image spectra

When an observer is presented with a SAR ocean image such as Figure 8.3, apparently showing waves on the sea surface, it is tempting simply to treat the image pixel values as if they corresponded directly to some property of the wave field, such as surface elevation or slope. A two-dimensional Fourier transform of the image readily produces a two-dimensional image spectrum which describes the directional properties of apparent waves on the image. Fourier analysis of a simple image like Figure 8.3 leaves a 180° directional ambiguity, meaning that although wave orientation is determined it is not known whether the waves travel forward or backward. However, most SAR processors can generate a "complex" image, in the mathematical sense that it consists of real and imaginary parts. What this means when interpreted physically is that it defines the relative phase of backscatter at

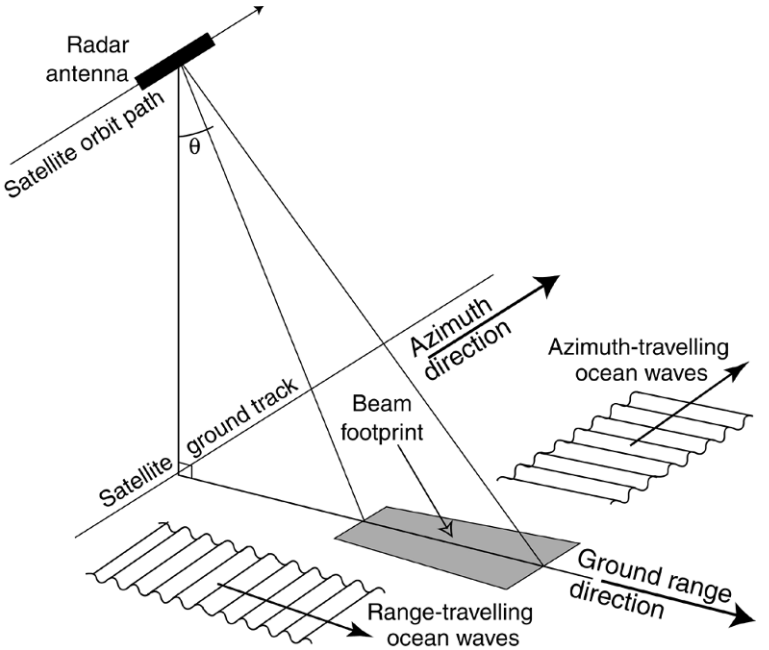


Figure 8.2. Schematic defining the range and azimuth directions for a SAR viewing ocean waves.

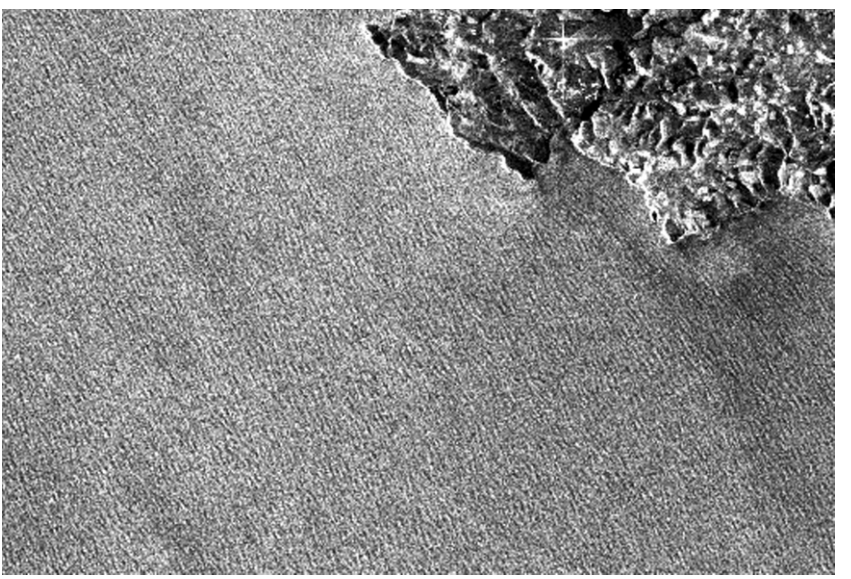


Figure 8.3. ERS-1 SAR image showing long surface waves off Prawle Point in the English Channel (3°43′W, 50°12′N). This image is 25 km wide and composed of 50 m pixels created by averaging blocks of 4×4 original pixels from the PRI image data file. An Atlantic swell appears to be propagating from west-southwest bringing the waves to the Devon coast in southwest England.

different parts of the image as well as its amplitude. This contains information about how σ_0 changes during the time over which the SAR builds up the image from the radar echoes of several hundred separate pulses. The availability of this additional knowledge can be used to remove the 180° directional ambiguity in retrieval of the image spectrum.

Relating image spectra to ocean wave spectra

But it would be wrong to treat this *image* spectrum as a representation of the two-dimensional ocean wave spectrum. The physical quantity represented by an image spectrum, based on calibration of the radar device, is σ_0, the normalized radar backscatter cross-section. Analysis of the imaging theory shows that the relation between σ_0 and surface height or slope is one which varies with wavenumber, with direction, and according to which of the imaging mechanisms is being considered. In practice the SAR image results from all three imaging processes occurring together in unknown relative proportions, plus any other processes not included in the theory such as the random addition of speckle noise across the image (see section 9.2.3 of *MTOFS*).

Separating out this complicated mix of processes to retrieve a true measure of the directional wave energy spectrum from the image spectrum is difficult. The approach adopted has been to define a *modulation transfer function* (MTF) which expresses the physics of the imaging process. If we already know the ocean wave field and its spectrum we can apply the MTF to predict what the SAR image and its spectrum would be. The challenge is to invert this procedure so that the wave spectrum can be derived from the measured image spectrum. Progress was made when an analytic expression for the nonlinear ocean-to-SAR spectral transform was derived (Hasselmann and Hasselmann, 1991), refined (Krogstad, 1992), and then adjusted for complex spectra which resolve directional ambiguity (Engen and Johnsen, 1995).

Inverting the transform to retrieve wave spectra from image spectra

Algorithms to perform the inverse transform were developed for application to ERS-1 data. These use an iterative approach requiring an initial guess of the wave spectrum from which the image spectrum is evaluated and compared with the measured SAR spectrum. The wave spectrum is then successively adjusted to reach convergence on the best agreement between the modeled and actual image spectrum. The effectiveness of the method was demonstrated in comparison with *in situ* wave measurements (Brüning *et al.*, 1994) before the algorithm was improved further (Hasselmann *et al.*, 1996). Developed at the Max-Planck-Institut für Meteorologie at Hamburg this is now often referred to as the MPI method. The initial guess is provided by wave model forecasts, although that can make it difficult to determine how much extra information is provided by the SAR that was not already in the *a priori* spectrum. We know that the SAR responds preferentially to longer wavelengths, and is not capable of fully resolving wavelengths shorter than about 100 m in the range direction or 200 m in the azimuth direction. Therefore we

might expect that little new knowledge is gained from SAR data about the short, high-frequency part of the wave spectrum at periods less than 10 s.

In a comparison between open-ocean directional wave buoy records in the South Atlantic and wave spectra retrievals from SAR following the MPI method it was found that the use of SAR added some extra information about long-period swell but tended to degrade the model forecast of short-period wind sea (Violante-Carvalho *et al.*, 2005). Nonetheless this approach was used for ERS SAR data and has been fairly effective in measuring spectra under swell-dominated sea conditions. In the absence of any other measurements it has been useful, even if its accuracy has been difficult to evaluate. Since the method requires a model prediction as a first guess it lends itself to incorporation into an assimilation scheme to constrain ocean wave models using SAR data. A further refinement of the MPI method was used to analyze wave mode data from ERS-2 (Schulz-Stellenfleth *et al.*, 2005).

An alternative inversion method, described as a semi-parametric retrieval algorithm (SPRA), was conceived (Mastenbroek and Valk, 2000) which does not use a first-guess spectrum predicted by a wave model, although it still requires some *a priori* knowledge. It uses the same nonlinear transfer function (Hasselmann and Hasselmann, 1991) as the MPI scheme, but bases its first wave spectrum guess on the equilibrium spectrum associated with wind speed and direction. For ERS-1 and 2, the SAR wave mode was sampled simultaneously with the scatterometer, so exactly coincident wind speed and direction could be guaranteed. The best fit to the SAR image spectrum is achieved by adjusting two parameters of the wind-dependent wave spectrum, those which define the stage of development of the wind sea and the propagation direction of the peak. The residual mismatch to the SAR image spectrum is assumed to be caused by the ocean swell spectrum. This can be translated from the SAR image spectral domain to the wave spectral domain using a local linearization of the MTF allowing an easy inversion to retrieve the swell spectrum. Thus the method distinguishes sea and swell, and also retrieves additional information along the way, such as wave age and direction of the wind sea.

This approach has recently been adapted for retrieving wave spectra from Envisat ASAR standard-mode and wide swath–mode images (Ardhuin *et al.*, 2004). In this case wind speed and direction must be obtained from local measurements or from a numerical weather prediction model, since there is no scatterometer on Envisat.

Alpers (2003) has provided a brief but helpful history of how radar oceanographers responded to the challenge of measuring ocean wave spectra from SAR, an interesting story that is still being written! Section 8.3.4 discusses the level 2 products (wave spectra) now available.

8.2.6 Wave spectrometry

There is a third method for obtaining wave information by radar remote sensing, which was demonstrated from aircraft two decades ago but has not yet been attempted from a satellite. Described generically as a wave spectrometer, it appears rather like a combination of an altimeter and a conically scanning real

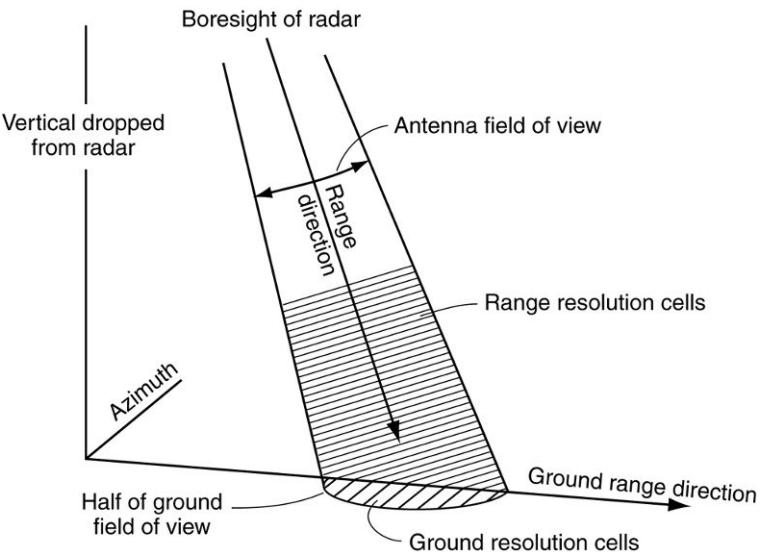


Figure 8.4. Schematic of radar spectrometer geometry showing how range resolution achieved by time-sampling the echo of a nearly normal radar beam achieves a coarser resolution in the ground-track direction.

aperture radar. Novel aspects of the technique compared with other established ocean radar systems are (i) to use a radar with a very low incidence angle (within 2° and 10° of nadir) to obtain a profile of backscatter in the ground range direction, and (ii) to sweep that view conically around the nadir position in order to sample that profile in all directions.

Figure 8.4 illustrates point (i). It shows a circular beam of aperture 2° pointed slightly obliquely on the ground. If this was mounted on a satellite at a height of 500 km the diameter of the almost circular illuminated region would be about 17 km. Because of the off-nadir tilt, the pulse illuminates the nearest range first and the farthest range last, so that by time-sampling the echo the measured backscatter can be resolved in the ground range direction across the field of view. No azimuth resolution is attempted and so the range-resolved signal corresponds to averaged backscatter within slightly curved strips cutting across the field of view as shown. If the echo is sampled at a sufficiently high frequency to achieve a range resolution of, say 1 m, this would correspond to a ground range resolution of 29 m at 2° incidence reducing to 6 m at 10°.

At nearly normal incidence, backscatter is linearly dependent on the tilt of the sea surface (i.e., the tilt about a horizontal axis parallel to the radar azimuth). Thus the spectrum of the time-varying echo from a single pulse can be interpreted as the spatial spectrum of surface tilt along the ground range direction. In other words each radar pulse can estimate a one-dimensional spectrum for the ground range component of wavenumber, with the capacity to resolve wavelengths down to about 60 m. The spectrum is the average over the 17 km instantaneous field of view of the antenna. Although such information is usable, by itself it does not provide a complete view of the spatial variations of the two-dimensional wave spectrum. To achieve this requires conical scanning.

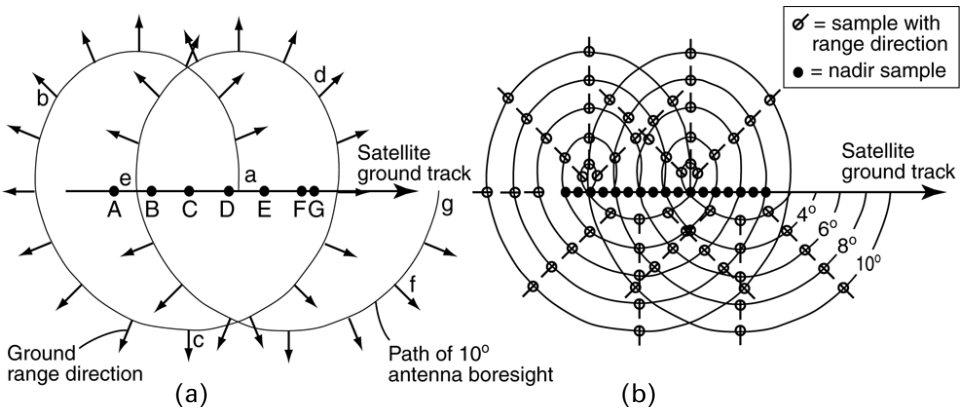


Figure 8.5. Schematic showing how a conical scanning wave spectrometer radar achieves views in all directions. (a) The locus of the center of the field of view of a single antenna pointing at 10° to the vertical, with arrows showing the local range direction at each point. Points a, b, . . . , g are where the radar is pointing when the platform is above positions A, B, . . . , G. (b) The same as (a) but with five separate scanning antenna pointing at 2°, 4°, 6°, 8°, 10°, and also a nadir-viewing radar altimeter.

The off-nadir-pointing radar is rotated about a vertical axis so that the beam describes a cone and its intersection with the ground would be almost circular if the platform was stationary. However, because it is moving, the ground path of the radar bore sight describes a cycloid curve on the ground (as shown in Figure 8.5). A single antenna is shown in Figure 8.5a, scanning around the vertical once in the time taken for the subsatellite point to travel a distance of about 70% of the radius of the radar scan on the ground. For a satellite at an altitude of 500 km this would correspond to an off-nadir tilt of 10° and a scan rotation period of 10 s, and the strip of sea illuminated by the radar would stretch approximately 90 km on either side of the satellite ground track. Within this strip it can be seen from Figure 8.5a that the radar would be able to measure one-dimensional wave spectra pointing in many different directions. If it could be assumed that the wave field is uniform over distances comparable with the radius of the scanning circle, then this arrangement could be used to estimate the average directional wave spectrum over the area. When used on aircraft, the radius of the circle is less than 1 km, and so that assumption is valid. This approach was successfully demonstrated with the radar ocean wave spectrometer (ROWS) deployed from aircraft (Jackson *et al.*, 1985a, b, 1987), and in conjunction with a scanning radar altimeter (Chapron *et al.*, 1994; Vandemark *et al.*, 1994).

In the suggested satellite case it is not reasonable to assume uniformity over a distance of up to 90 km. However, a design for a satellite wave-measuring mission called SWIMSAT (Hauser *et al.*, 2001) proposed that several antennas could be used together as shown in Figure 8.5b. There are five antennas mounted on the rotating assembly, pointing at 2°, 4°, 6°, 8°, and 10° from nadir and achieving a denser

coverage with more probability of different sensors observing the same patch of sea from different directions. Note that many more spectra could be sampled along the scan paths of each of the antenna than is shown in the sketch. This would allow average directional wave spectra to be evaluated over regions no larger than around 50×50 km. At the same time a nadir-pointing radar would measure significant wave height as described in Section 8.3.1. Since the paths of each scanning antenna cross the satellite subtrack there would thus be an opportunity to calibrate wave spectral estimates using the independently measured H_S.

Although this approach has been proved on aircraft, and the designs for a satellite instrument look promising, no agency has yet been prepared to support a trial satellite mission. One reason for this may be that the other available methods described above go a long way to meeting measurement requirements, and so they have reduced the pressure to try an alternative which may not necessarily be any better than what is available using altimetry and SARs.

8.3 MEASURING OCEAN WAVES—PRACTICAL SYSTEMS

8.3.1 Altimeters for measuring SWH

Because of the inherent simplicity of the way altimeters measure H_S and the correspondingly modest technological requirement of being able to track the first arrival of each echo and to sample its profile with sufficiently high frequency, this method has been operating successfully for more than two decades. Table 8.1 lists the satellites which have carried altimeters that successfully recorded wave height. The Seasat altimeter in 1978 first proved the success of the technique (Fedor and Brown, 1982) and satellite wave data started to be used operationally from Geosat onwards. Since then every new altimeter flown has been shown to be capable of measuring H_S to an accuracy at least as good as the buoy data against which it has been validated (Carter *et al.*, 1992; Cotton and Carter, 1994; Ebuchi and Kawamura, 1994; Gower, 1996). That is, a scatter plot of buoy-derived H_S against satellite measurements shows a spread from the equivalence line that is no greater than the uncertainty of the buoy data (see figure 11.32 in *MTOFS*). There have been three strands of altimeter deployment; the U.S. Navy flew Geosat and the subsequent Geosat follow-on (GFO) although data are distributed by NOAA; the European Space Agency (ESA) flew ERS-1, ERS-2, and Envisat; TOPEX/Poseidon was a joint program between NASA and CNES, followed by Jason-1 and Jason-2. Each program has sought individually to provide continuity of measurements with, if possible, an overlap between the launch of a new satellite and the shutdown of the sensor it is replacing, while maintaining identical orbits within the program. Although this was done primarily for the sake of altimetric measurement, it ensures continuity of sampling characteristics for wave data too.

Table 8.1 shows the orbit characteristics of each sensor, and from these spatial and temporal coverage can be deduced. The altimeter is not a scanning sensor but measures waves only along a line corresponding to the satellite ground track and so

Table 8.1. Altimeters providing measurements of wave height since 1978.

Altimeter	*Wave data record*		*Orbit repeat* (days)	*Orbit inclination* (max. lat.)	*Comments*
	Start date	*End date*			
Seasat	July 1978	Sep. 1978	3	108° (72°)	Proof-of-concept mission
Geosat	1986	1989	17 . . . 05	108° (72°)	Data from only the Exact Repeat Mission phase
ERS-1	1991	1999	3, 35, 180	98° (82°)	Changed orbit patterns for different mission phases
TOPEX/Poseidon	1992	Jan. 2006	9.9156	66° (66°)	In an orbit optimized for altimetry
ERS-2	1995		35	98° (82°)	Continuity from 35-day repeat phase of ERS-1
Geosat FO	1998	2008	17 . . . 05	108° (72°)	Continuity from Geosat
Jason-1	2001		9.9156	66° (66°)	Continuity from TOPEX/ Poseidon
Envisat RA-2	2002		35	98° (82°)	Continuity from ERS-2
Jason-2	Dec. 2008		9.9156	66° (66°)	Continuity from Jason-1

sampling is much less dense than from sensors that have imaging capability. The precise orbit repeat cycle determines the space-time resolution that is achieved. There are typically about 14 or 15 orbits per day, which are separated from each other in longitude by about 24°, but on subsequent days a different ground track is mapped out, and so on until eventually the orbit returns to its original track and repeats the cycle indefinitely. The extent to which the gaps in longitude are filled in depends on the orbit repeat period. If this is short, say 3 days, then only about 44 or 45 evenly spaced tracks are ever sampled, resulting in a spacing of about 8° of longitude which at the Equator is nearly 900 km. This dictates the cross-track spatial resolution, although at higher latitudes the tracks are more closely spaced than at the Equator. Note that the along-track sampling rate is very much higher, giving a resolution of order 10 km.

The sampling consequences of different orbit repeat periods is illustrated in Figure 8.6 which compares the coverage of tracks over North Atlantic between (a) 3-day and (b) 10-day repeat cycles. The thick lines correspond to tracks on one single day, and show that sampling capability within one day is very similar in both cases. However, over time the t10-day repeat orbit builds up measurements on a much finer grid. On the other hand, the longer orbit repeat cycle results in poorer time sampling at each location. It is important to recognize the restrictions

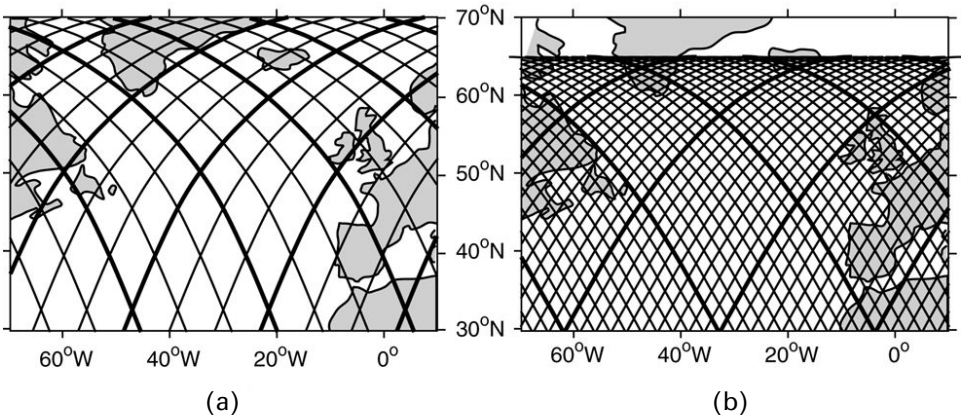


Figure 8.6. Spatial distribution of the ground track of an altimeter on a satellite with (a) a 3-day orbit repeat cycle and an inclination of 72° and (b) a 10-day repeat cycle and an inclination of 65°. The thick line shows tracks for a single day.

which sampling limitations of a single altimeter place on uses to which wave height measurements can be put (as discussed in Sections 8.4 and 8.5).

The inclination of the orbit determines the maximum latitude that is reached by the satellite (as shown in Figure 8.6). None of the three altimeter programs have selected orbits optimized for wave measurement. ERS and Envisat satellites carry other sensors that need a near-polar, Sun-synchronous orbit, while the TOPEX/Poseidon, Jason, and Geosat programs chose orbits to suit the primary needs of altimeter sea level measurements. Being driven by different requirements, the orbit types of the three independent programs differ considerably from each other, but this has ensured a degree of complementarity between data sampling, with ESA sensors reaching much higher latitudes.

8.3.2 SWH data products from altimeters

Wave data are acquired from all the altimeters mentioned in Table 8.1. At the time of writing, the current and most readily available data are those from the Poseidon-2 altimeter on the Jason-1 and 2 spacecraft and the RA-2 altimeter on Envisat. Some details of the data products from these sensors are presented in Tables 8.2 and 8.3. In each case, wave data records consist of measurements of H_S derived from averaged echoes every second, equivalent to spatial sampling approximately every 7 km along-track. The data are made available in digital form as part of the Geophysical Data Record (GDR) which also contains the sea surface height anomaly, the wind speed, and several other variables derived from the altimeter. The tables provide references to more detailed product description documents.

Note that wave height data are made available within 3 hours of an overpass. The preliminary dataset is based on the results of onboard processing and is intended

Table 8.2. Details of wave data products from Jason-1.

Sensor and satellite	Poseidon-2 on Jason-1
Agency providing data	NASA/JPL or CNES
Internet address for product access information	*http://nereids.jpl.nasa.gov/cgi-bin/ssh.cgi?show=overview*
Rapidly available product	Operational sensor data record (OSDR)
Time available after overpass	Within 3 hours
Interim data product	Interim geophysical data record (IGDR)
Final data product	Geophysical data record (GDR)
Data user's guide	Zanife *et al.* (2001)

Table 8.3. Details of wave data products from Envisat RA-2.

Sensor and satellite	RA-2 on Envisat
Agency providing data	ESA
Internet address for product access information	*http://earth.esa.int/dataproducts/* Then use the "Browse Products by" menu to select RA-2
Rapidly available product	Fast Delivery Geophysical Data Record. (Code: RA2_FGD_2P)
Time available after overpass	Within 3 hours
Interim data product	Intermediate geophysical data record (Code: RA2_IGD_2P)
Final data product	Geophysical data record (Code: RA2_GDR_2P)
Special Met product	Wind/wave product for Meteo users (Code: RA2_WWV_2P)
Data user's guide	Benveniste and Milagro (2000)

for near real–time operational applications rather than scientific analysis. After several days, more detailed ground-based processing is performed on downloaded waveforms to produce an interim data product that should be more reliable for scientific applications. It is termed an interim product because it is several weeks before a fully accurate satellite orbit is available and the final GDR is released as the most accurate product. Although knowledge of the orbit is crucial for the accuracy of sea surface height anomaly records, there is very little difference between wave heights in the IGDR and in the final GDR.

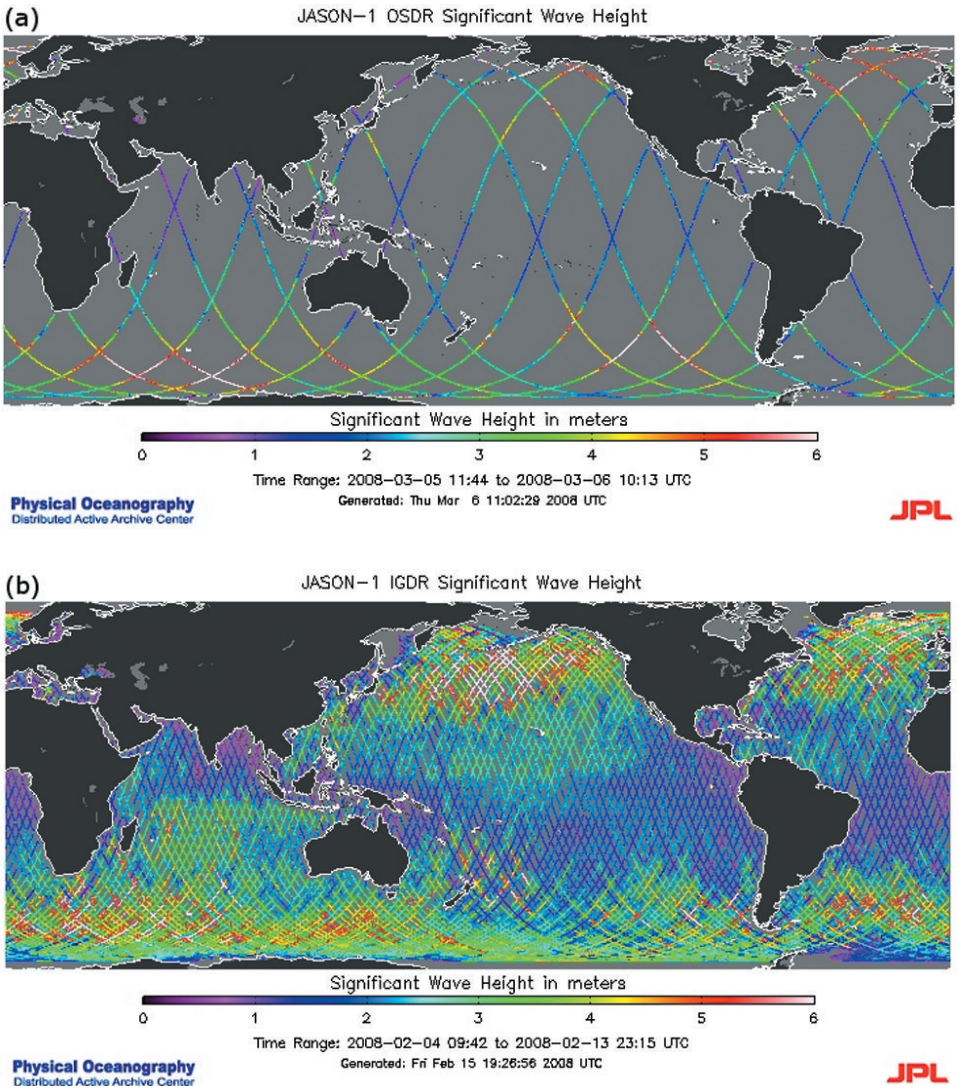


Figure 8.7. Significant wave height data products produced by NASA/JPL from the Poseidon altimeter on Jason-1. (a) The global operational sensor data record (OSDR) shows data from the previous 24 hours at the time this image was requested. (b) The interim geophysical data record (IGDR) shows all the tracks from the most recent complete orbit repeat cycle of Jason-1 at the time the data were requested (courtesy NASA/JPL-Caltech).

NASA/JPL provide globally mapped images of wave height data (as shown in Figure 8.7). The upper panel (a) presents tracks from a single day. Although sparse, this does provide a useful overview of the locations of major storms and high seas, but users must be aware that samples from a single day are likely to miss some

storms entirely. The lower panel shows the data acquired over a full 10-day repeat orbit cycle of Jason-1 and represents the finest across-track spatial resolution that is achievable from this platform. Note that data are presented as individual tracks and not averaged or otherwise transformed into a smoothed map of wave height. To do so could be misleading as it would tend to conceal extremes of high and low wave heights that remain in individual tracks. It is inappropriate to smooth the data because wave height can change a great deal over the full 10-day cycle. Even within a single day there is evidence of large differences between measurements from an ascending and a descending orbit at some of the crossing points in Figure 8.7a. Instead, the user presented with Figure 8.7b is immediately aware of where the highest seas were encountered during the 10-day period, and also of temporal variability. Although the time corresponding to each plotted point is not evident in the graphical plot, it is retrievable from the digital data record.

When there are several altimeters in operation at the same time, spatial coverage in a single day is considerably improved. With enough sensors in orbit it would be feasible to produce smoothed maps of wave height that do not disguise extremes. In practice, for operational applications wave models are used to nowcast or forecast wave height distributions (as discussed in Section 8.5). However, some seagoing users of wave forecasts consider it useful to obtain these altimeter snapshots of waves along individual tracks in near–real time. This allows mariners to compare satellite measurements with forecast waves and thus evaluate model performance in near–real time. Ultimately they make navigation decisions based on their own judgment, with reference to all the available information including both satellite observations and model forecasts.

8.3.3 Synthetic aperture radars

A review of the various SARs flown in space since the Seasat SAR proved the concept in 1978 can be found in section 10.5 of *MTOFS*. Currently the main source of freely available ocean wave spectral information comes from the Advanced SAR (ASAR) on ESA's Envisat satellite launched in March 2002 in a near-polar, Sun-synchronous orbit. The ASAR's heritage is that of the SARs on ERS-1 and ERS-2 and it uses the same radar frequency of 5.331 GHz (C-band) but it can operate in HH, VV, and cross-polarization states. It has several different imaging modes and all, except global-monitoring (low-resolution) mode, are able to reveal waves.

For operationally monitoring ocean surface waves, the most important mode is the ASAR (WV). This images a small region of the ocean, with a size between 10×5 km and 5×5 km, at regular intervals of 100 km along-track, and positioned within a standard *image mode* (IM) swath. Up to two positions in a single swath or in different swaths may be specified, with acquisitions alternating between one and the other. HH or VV polarization may be chosen. It thus continues the wave-imaging mode introduced with the ERS-1 SAR and continued with ERS-2, but with improved flexibility and higher performance. The rationale of wave mode is to acquire during each orbit over the ocean a few hundred small imagettes of the sea

surface from which ocean wave directional spectra can be derived. This is much less demanding on the satellite's power and data bandwidth resources than continuously operating the SAR in full imaging mode over the ocean. In principle, it offers effective along-track sampling of typical wave variability even though, like the altimeter, the spacing between orbits in a single day is too large to completely map the full ocean wave field.

Radarsat-2 was successfully launched on December 14, 2007 and carries a C-band synthetic aperture radar operating at 5.405 GHz. Canada's second Earth-observing satellite has a planned life of 7 years and, like its predecessor Radarsat, its primary mission is to provide environmental information for managing Canada's high-latitude land, lakes, sea, and ice. Beyond data for the Canadian government it operates as a commercial venture and so data for all other users must be purchased. Radarsat-2 will be used like ASAR to provide wave spectra although this does not seem to be a high priority in comparison with land, ice, and wind-mapping applications. At the time of writing it remains to be seen whether wave data will be produced routinely to complement ASAR products.

8.3.4 ASAR wave-related products

There is a single, main, level 1 data product from ASAR wave mode (ESA code: ASA_WVI_1P) which consists of a wave-mode, single-look complex (SLC) imagette (including real and imaginary parts) for each sampled location. Figure 8.8 shows an example of the amplitude of a single-look complex image. This is accompanied by the power spectrum of each imagette computed by the cross-spectral method. Note that the cross-spectrum eliminates directional ambiguity of a simple directional image spectrum. Typically up to 400 imagettes and matching spectra are provided in each dataset, corresponding to acquisitions from about one orbit of Envisat. Because these products require a much lower data rate than full SAR image modes, they are produced continuously and routinely span the whole ocean.

The ASAR standard, wave-mode, level 2 product (ESA code: ASA_WVW_2P) is the result of applying the inverse MTF transform procedure to the level 1 cross-spectrum (as discussed in Section 8.2.5). In its original form it used the MPI method which requires wave model forecasts as the *a priori* initial prediction of an iterative algorithm that minimizes, with respect to the wave spectrum, the mean square difference between the observed and the computed SAR image spectrum. Compared with wave-spectral products originally produced from ERS SAR, those from ASAR have eliminated directional ambiguity and deal better with the effects of radar speckle, but it is still unclear how much additional information about the wind wave field comes from the radar rather than from the original model forecast.

The SPRA method has also been developed for use with Envisat wave-mode data, although at the time of writing it is not yet clear whether ESA is formally issuing this as an alternative level 2 wave product or leaving users to apply it themselves to level 1, wave-mode data.

The SPRA method has also been applied to Envisat standard-mode and wide swath–mode images (Ardhuin *et al.*, 2004); the SAR image is divided into a grid of

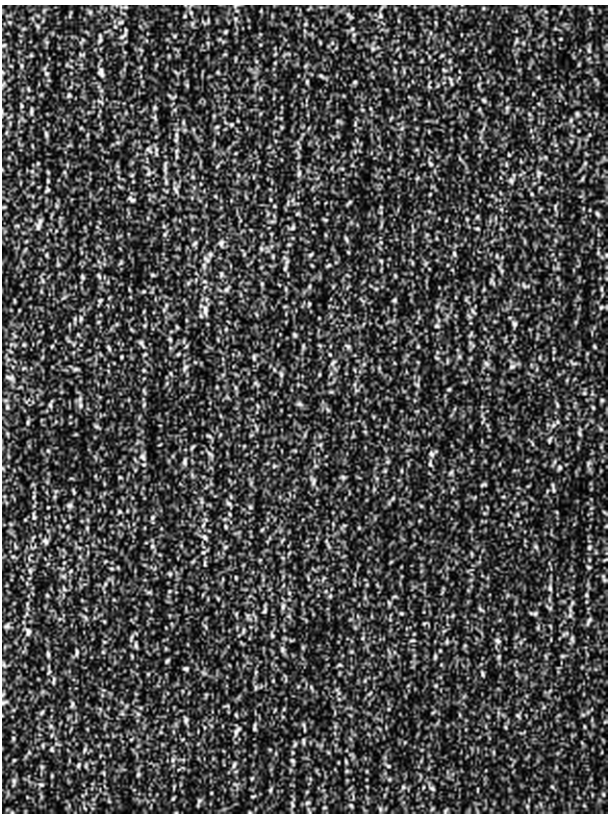


Figure 8.8. Example of an Envisat ASAR, wave-mode, level 1 product showing the amplitude image of a single-look complex radar backscatter cross-section for a wave mode imagette, 5 km wide (image obtained from the ESA online Envisat User Manual).

2.0×2.5 km imagettes and an individual spectrum retrieval is performed in each. The retrieved wavenumber spectrum at each grid cell is converted into a directional frequency spectrum, and can be presented graphically as a map showing the peak direction as an arrow and the wave height in color (as shown in Figure 8.9). This method is presently under consideration by ESA to be adopted as a new, standard wave product.

Figure 8.10 shows another example of an ASAR image analyzed to show the direction of the spectral peak and significant wave height. This is an image off the Californian coast, just south of Santa Barbara, showing the northern group of California's Channel Islands. In this case, the color-coded H_S and the directional arrows have been overplotted onto the basic σ_0 image so that surface slicks, wave-breaking at the coast, and other SAR ocean features are still visible if the viewer is able to zoom in to a higher resolution (as shown in Figure 8.10b).

In a contrasting approach to making new SAR products, it is interesting to note that a purely empirical technique has been developed which extracts wave information from radiometrically calibrated SAR images without any reference to either MPI or SPRA inversions (Schulz-Stellenfleth *et al.*, 2007). Instead of retrieving the wave spectrum the method estimates wave integral properties, such as H_S, mean

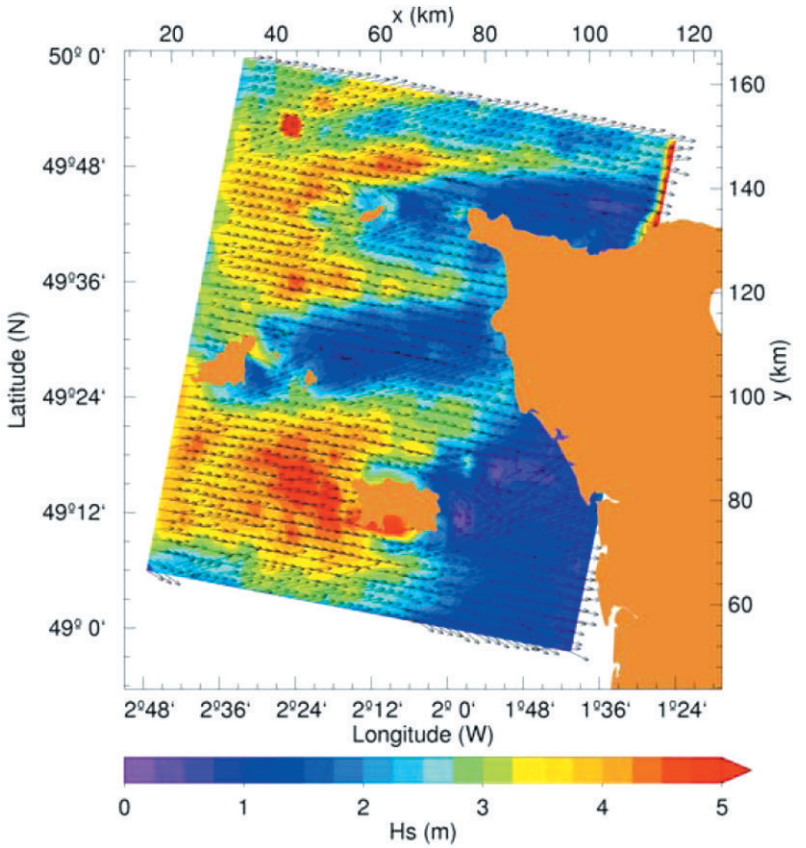
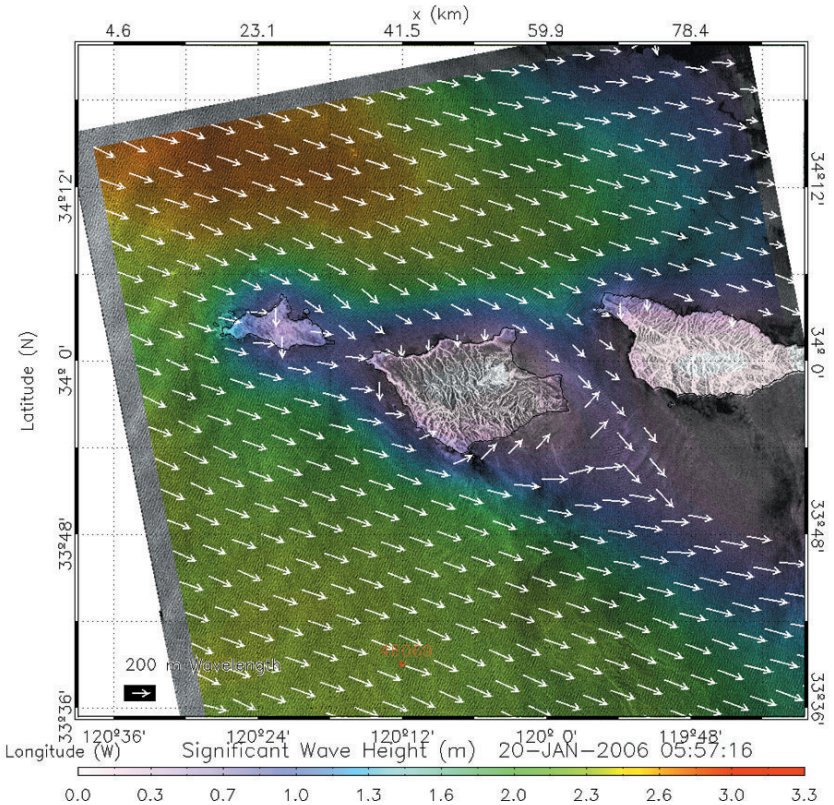


Figure 8.9. Example of wave heights (colored scale) and mean wave propagation direction, retrieved on a grid with spacing 2.5×2.0 km from an Envisat ASAR, single-look, complex image over the Gulf of St. Malo in the English Channel on March 9, 2003 at 10:22 UTC (from Ardhuin *et al.*, 2004).

wave period specified in different ways, and wave power and wave height associated with different spectral bands. It does this by characterizing each SAR wave-mode imagette in terms of mean σ_0, variance, and 20 other parameters derived from orthonormal functions fitted to the image spectrum. These parameters then form inputs to a quadratic model function that estimates wave integral properties. The model function is entirely empirical, its coefficients tuned to achieve a match for a training set of 3,000 SAR imagettes co-located with coincident samples from a WAM wave model which generated the wave integral properties. The advantage of an empirical approach like this is the speed and computational efficiency with which wave parameters can be delivered. However, the validity of the results depends on the scope of the training set (how representative the samples are of the range of possible wave conditions) and on the accuracy of the WAM model used to forecast integral wave properties for the training set.



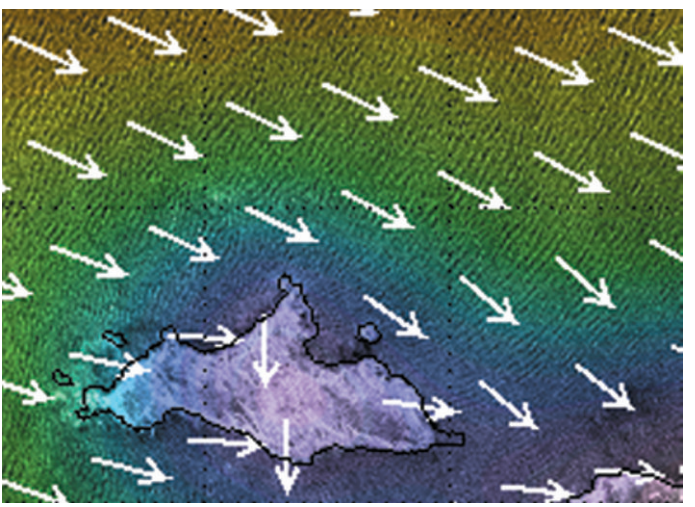


Figure 8.10.
(a) Envisat ASAR image of the Pacific Ocean south of Santa Barbara, showing the northern group of Channel Islands, on January 20, 2006, overlaid with color denoting significant wave height H_S and wave direction. (b) Closeup of a part of (a), revealing the signatures of the swell on the SAR image (images downloaded from the webpage of Boost Technologies).

8.4 APPLICATIONS OF WAVE DATA FROM SATELLITES

8.4.1 Applications of SWH

Applications of H_S measurements from altimetry can be classified into three broad categories. These are: operational uses in which data are required in real-time to support decision making for navigation and ship-routing, offshore engineering activities, and other diverse marine activities; scientific use of the data for learning more about the variability of ocean waves and their relationship to wind, currents, and bathymetry; and the use of satellite data to generate wave climate statistics (discussed in Section 8.6).

Extreme events at sea, when storm winds and waves damage and delay ships, claim many hundreds of lives and cost the marine insurance industry billions of dollars every year. The better able we are to forecast such events the more these losses can be reduced. The main tool presently available for improving operational safety at sea is the use of global numerical weather forecasting, coupled to the wave-forecasting models discussed in Section 8.5. But just as important as developing reliable ocean wave–forecasting models is to have a means of sampling the ocean sufficiently frequently in space and time to be able to map the current distribution of sea state. Satellites offer the possibility of measuring wave heights over the whole of the World Ocean. Reliable H_S measurements by altimeters have been available for more than 20 years and it is not surprising that a number of private companies and public agencies offer specialized services to collate and interpret wave information from satellites, in order to satisfy the requirements of their clients, including the transmission of near real–time data to mariners to assist in their navigation decisions. To be most effective this must be done in near real–time. The variety of specific application contexts in which near real–time data are used, in addition to general ship routing, include transport of unconventional loads, cable-laying operations, offshore oil and mineral prospecting and drilling operations, marine insurance, yacht races, naval exercises, high-speed ferries, and coastal defences (Krogstad and Barstow, 1999).

The main factor that continues to constrain the growth of the sector supplying satellite data to operational users is the relatively limited sampling capability of the altimeters currently in orbit. Although there have been periods during the last 17 years when as many as four instruments have been operating at the same time, there are presently plans for only two continuing altimeter series. These are the ESA RA-2 on Envisat to be succeeded by an altimeter on the planned Sentinel-3 satellite series, and the Jason series following from TOPEX/Poseidon. As Figures 8.6 and 8.7a show, the line of points sampled during a single day by an altimeter gives rather sparse coverage, with gaps between separate orbits of up to 2,500 km. For a high-sea-state warning system to be reliable, it must be capable of detecting all regions where very high waves occur. Since these high-risk zones for shipping may persist for no longer than one day, and may be confined spatially to a few hundred kilometers, it is quite possible that they are missed by satellite monitoring if there are only two or three sensors operating. It is therefore impossible at present to offer a credible,

operational, high-wave warning system based on satellite altimeter data alone; which is why reliance must be placed on wave model predictions. In this situation, satellite data are most effectively used to improve the performance of wave models.

8.4.2 Applications of SAR

There are several potential applications of SAR-derived ocean spectra (Heimbach and Hasselmann, 2000) but the operational use of SAR wave-mode data for monitoring the distribution of wave spectra around the World Ocean is hindered by the same sampling limitations as discussed above for H_S. Thus, like altimeter data, the best route at present for SAR wave-mode data to contribute to operational applications is through assimilation into ocean models (see Section 8.5) and the accumulation of wave statistics for defining climatologies of ocean waves. The more direct applications of SAR data to ocean wave research and application is through the use of SAR image mode data, both standard and wide swath.

The recently developed technique of wave spectral retrieval on a 2×2.5 km grid (Ardhuin *et al.*, 2004) (as shown in Figure 8.9) offers new, detailed information about how wave fields vary spatially. A phenomenon very evident on this image is the way in which shadows form behind islands and waves have a very much lower amplitude. It is not clear how much of the shadowing is due to wind reduction behind the island and how much is because the wave energy flux is diverted through reflection and refraction by the island. Also to be noted are the bending of wave vectors by refraction as they interact with shallow bathymetry and "hot-spots" of high amplitude (where $H_S > 4.5$ m). This is probably also caused by bathymetry, the wave energy flux converging as group velocity slows down when waves are constrained to go slower by shallowing water depth. For the first time, images like this show the spatially detailed variability of wave energy, and provide a way of testing the behavior of waves predicted by models. The SAR shows the real sea, responding to a combination of the local wind that controls local growth or damping of shorter waves, as well as the wave energy propagating into the region.

Although time sampling is inadequate to follow the evolution of a wave field at the necessary temporal resolution, there are isolated opportunities at orbit crossover locations to obtain a second overpass about 12 hours later (as shown in Figure 8.11) (Ardhuin *et al.*, 2004). Although there is only a limited area of overlap between the two images, there is some evidence that the wave field has changed slightly in that time.

Although these new SAR wave products are yet to be fully validated in order to confirm their quantitative accuracy, their qualitative value to maritime safety agencies and coastal engineers is already evident. Their release has been too recent to result in publications. However, there are some papers based on earlier SAR wave products which demonstrate the effectiveness of SAR in studying particular phenomena (e.g., in the study of how swell waves propagate within the marginal ice zone—Schulz-Stellenfleth and Lehner, 2002).

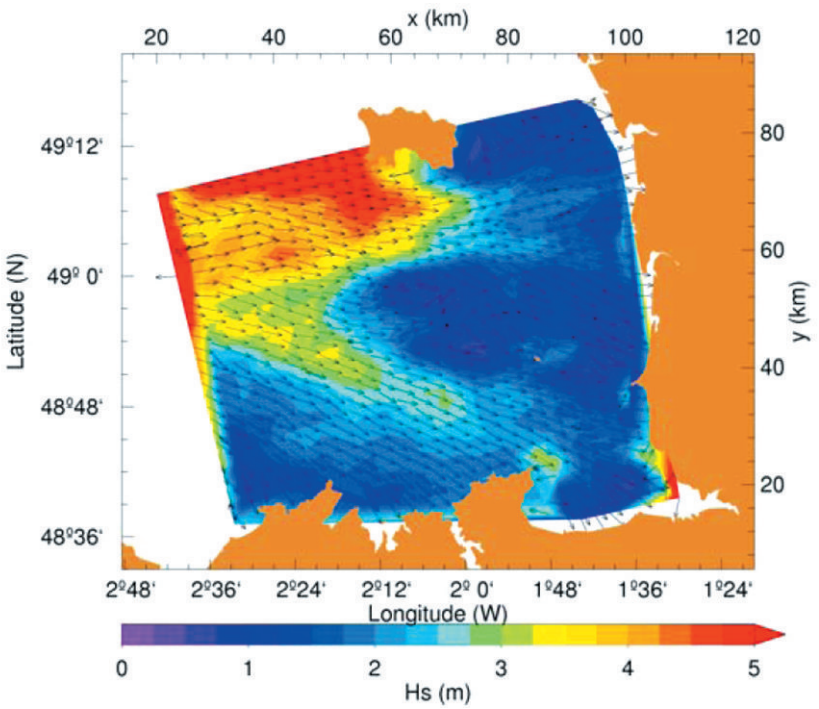


Figure 8.11. Example of wave heights (colored scale) and mean wave propagation direction, retrieved on a grid with spacing 2.5×2.0 km from an Envisat ASAR, single-look, complex image over the Gulf of St. Malo in the English Channel on March 9, 2003 at 21:44 UTC (from Ardhuin *et al.*, 2004).

8.5 USING SATELLITE DATA IN WAVE PREDICTION MODELS

8.5.1 Wave prediction models

In many respects the development of forecasting models for ocean waves during the last three decades of the 20th century (Komen *et al.*, 1994) was a successful demonstration of the application of mathematical analysis to a complex environmental phenomenon. The foundations were laid in the 1950s and 1960s when the concept was developed of representing a wave field by its spectral components, which are the amplitude and propagation direction of distinct frequency bands containing all the energy of wind sea and swell. Put simply, a spectral wave model is a set of equations, the solution of which specifies the spectral components at each grid point in two-dimensional space. The wave model must contain a source term, by which energy is fed into the wave system to increase the amplitude of each wave component depending on wind speed and direction, a propagation term that defines how wave energy in each spectral component travels and spreads across the ocean surface, and a dissipation term that characterizes the ways a wave system loses energy.

This was the basis of the *first generation* of wave models. They were able to use information about wind distribution over the ocean to forecast the effect this would have on the wave field. They could therefore exploit the improvement in ocean wind forecasting that grew out of the establishment of numerical weather prediction (NWP) models that assimilate satellite wind data. The first priority for any wave-modeling system is to have the most accurate inputs of wind for the source term. However, first-generation models treated each of the wave components independently and failed to take into account the complex nonlinear interactions by which energy is transferred between components (Hasselmann, 1962). *Second-generation* wave models introduced terms which allowed energy transfer to occur between spectral components, but in a simplified way that still left them inadequate to effectively represent what actually happens in the ocean in all possible conditions. With improving NWP these inadequacies have become more evident, and so *third-generation* models have been produced which attempt a more accurate representation of nonlinear wave–wave interactions (WAMDI_Group, 1988; Booij *et al.*, 1999; Tolman *et al.*, 2002).

8.5.2 Use of satellite data with wave models

Despite their relative success, challenges are still faced by the wave-modeling community (Cavaleri, 2006) as shortcomings still remain in third-generation models. What can the use of satellite-retrieved ocean wave data offer towards the improvement of ocean wave modeling and prediction? As in other branches of oceanography where numerical models are used, there are basically three ways of confronting numerical models with satellite-derived data: comparing model predictions with satellite observations for validation purposes; improving parameter estimation for wave models by tuning them to minimize the difference between model output and satellite data; and assimilation of satellite data directly into wave models.

Validation of wave models has generally been performed using *in situ* observations from wave buoys. However, the advantage of using satellite observations is their wider spatial sampling and global coverage, although their relatively infrequent revisit interval at a given location limits their effectiveness in evaluating how well the models predict the timing of peak waves. Ironically, the first comparisons between a new, satellite wave data product and wave models are usually made by the remote-sensing community seeking to validate their data products, and using wave forecasts as the best representation of the real ocean waves (Mastenbroek *et al.*, 1994). Sometimes archived satellite data such as the Seasat altimeter H_S record have been retrospectively validated in this way (Bauer *et al.*, 1992). However, once the broad reliability of satellite data is established, comparison becomes an exercise from which both modelers and remote sensors can learn (Romeiser, 1993). With careful analysis of the differences, lessons can be learned about the shortcomings of each data product. The relatively sparse sampling of satellite data is not a serious obstacle to this activity. Remote sensing can offer a key integral wave property with a 20-year heritage, H_S, as well as more recently developed wave properties like T_m or T_p that

can be used to compare with model output. As confidence grows in the validation of SAR-derived ocean wave spectra, these should begin to provide a more complete spectral test of wave model predictions. While there are some examples of this type of comparison being performed (Heimbach *et al.*, 1998; Chen *et al.*, 2002), it is far less common than might be expected. Clearly there is scope for improving the communication and exchange of data between the two communities and for developing methods to facilitate the co-location of satellite data and model outputs.

Using satellite data to assist in wave model parameter estimation is also a potentially important application. The substantial archive of H_S available from several different altimeters over more than 20 years has the potential to deliver many examples of particular types of wave behavior. These can produce test datasets as the basis for refining the parameterization of terms used in third-generation wave models. For example, Kalantzi *et al.* (2009) used TOPEX altimeter data co-located with WaveWatch III model predictions in the Indian Ocean to compare the effectiveness of two different wave dissipation parameterizations under different monsoon conditions.

8.5.3 Assimilating satellite data into models

In the 1990s the emergence of a reliable source of H_S from altimetry and wave spectra from SAR prompted a lot of work by wave modelers to develop methods for assimilating such data and to assess potential gains (e.g., Lionello *et al.*, 1992; Foreman *et al.*, 1994; Young and Glowacki, 1996; Holthuijsen *et al.*, 1997; Dunlap *et al.*, 1998; Greenslade, 2001). It is interesting to note the different approaches to assimilation that are used, according to whether it is H_S, wave spectra, or both being assimilated. Because of the nature of propagation of swell wave energy along great circle paths over the globe, adjusting the model to correspond to reliable observations can ensure that any improvements to the accuracy persist for several days and propagate into other parts of the model domain. When spectra need to be adjusted to match observations, the correction is not just carried forward but can be traced backwards to allow adjustment to the input of energy by a source at a previous time. In one system this was done explicitly through the use of a Green's function approach (Bauer *et al.*, 1996).

The conclusions of these assessments were mostly positive. Some improvements were found from assimilating H_S and some operational agencies started to assimilate altimeter data routinely (Breivik and Reistad, 1994). However, the impact of SAR was limited, and most agencies appear to have decided that the slight benefits in swell prediction did not justify changing the operational system. The reason why SAR made little difference was ascribed to directional ambiguity in ERS SAR spectra and the low number of available SAR observations (Breivik *et al.*, 1998). It must not be forgotten that a correct description of ocean surface winds, normally derived from NWP models, is the fundamental prerequisite for successful wave forecasting. Assimilation of wave data cannot remedy poor wind input or a poorly configured wave model and so was not considered to be a very significant factor by some operational forecasting agencies in the 1990s.

However, now that wave models and wind inputs have improved further, more reliable and unambiguous wave spectra are being retrieved from ASAR wave mode in greater numbers, and as there is growing commitment by agencies to maintain operational altimeters indefinitely, more agencies appear to be using assimilation of both H_S and SAR spectra into their operational wave-forecasting systems (Abdalla *et al.*, 2003). A recent study confirmed the benefits of assimilating satellite H_S into a wave model in the Indian Ocean, although not in the Mediterranean which is dominated by wind seas (Emmanouil *et al.*, 2007).

While forecasting of high wind seas has tended to be considered more important than swell for most marine operations, there are increasing numbers of applications where predicting swell is also critical, The use of floating rigs for offshore oil operations in deeper waters must take into account the effect of heave, even when wind sea is negligible. Moreover rigs may resonate at the frequency of long-period swell, and therefore more accurate forecasts of swell period and amplitude that can be achieved using satellite data assimilation are operationally very desirable.

8.6 WAVE CLIMATE

Wave climate is a statistical description of how wave parameters vary in space and time. To create the climatology of significant wave height, H_S, for example, requires a probability distribution of H_S from many samples at many locations. Then a parameter which partly describes that distribution, such as its mean, can be evaluated at each location and mapped globally. It can be broken down by month and a separate mean climatology can be mapped for each month of the year to show seasonal changes in the spatially mapped mean H_S. Interannual variability can also be detected. At locations where there has been a wave buoy moored for many decades and a large population of samples has accrued, it is possible to build up a clear picture not only of seasonal variability but interannual and decadal-scale changes. From many samples it is possible to identify trends, or relate interannual and decadal variability to other factors which may be linked to the wave field, including climatological indices. It is also possible to estimate the probability of extreme wave events at those locations.

Until satellite wave data became available, wave climate statistics were limited, either to locations where a wave gauge had been installed on a moored buoy, or to observer reports or instrumented records from ships. The former result in temporally detailed but spatially very sparse data, The latter have to be composited into rather large areas (hundreds to thousands of kilometers) if useful numbers of samples are to accrue. The locations of ship reports are generally constrained to the major shipping routes that fail to reach large parts of the ocean and are not well distributed even in the North Atlantic and northwest Pacific where shipping is more dense. Moreover the quality of observer reports is subjective (Gulev *et al.*, 2003) and instruments need to be calibrated.

The lack of a densely sampled and spatially distributed wave database left large gaps in both knowledge and understanding. For example if the wave climate at a

given buoy location was found to be anomalously low for a given month or year, there was no way of distinguishing whether this was an instrument malfunction, or a distinctive change of behavior characteristic of the whole surrounding region, or whether there had been a small positional shift in the local spatial pattern of wave height, but overall little regional change. In principle it is possible to populate climatologies by running ocean wave–forecasting models for a wide range of different forcing conditions, or for an extended period of actual forcing, but to establish the validity of the outcome requires observational data for comparison. Now that we have more than 20 years of altimeter wave heights, sampled all over the globe, it is possible for the first time to produce a spatially detailed, globally distributed, observation-based wave climatology.

Creating climatologies from satellite data

Because the altimeter is not a scanning sensor the spatial resolution of satellite-derived climatology must be relatively coarse, typically based on $2^\circ \times 2^\circ$ (latitude–longitude) elements of the sea surface. To create a multisensor wave climate it is first necessary to intercalibrate different altimeters and apply bias adjustments relative to a given sensor acting as the "standard". This ensures uniformity of data independently of which instrument produced them. Then each time an altimeter track passes through a particular cell, sampling H_S typically every 7 km along the track, the median value of all the samples is entered as a measure of the wave height for that cell for that time. For a global-scale, open-ocean climatology, 2° cells are acceptable since wind and wave systems typically vary on a lengthscale larger than 200 km. In coastal waters and shelf seas, however, wind shadowing by mountains or headlands and wave refraction by shallow bathymetry can introduce fine-scale variability into wave height (see, e.g., Figures 8.9, 8.10, and 8.11). Here smaller cells are desirable, but this requirement must be traded against the number of samples available. Figure 8.12 shows an example of mean wave height distribution with a resolution of about 110 km (Woolf *et al.*, 2003). It is better resolved than previous wave climatologies from conventional wave data and from the earlier Geosat altimeter (Challenor *et al.*, 1990; Carter *et al.*, 1991). It allows more detailed analysis to be performed, but is clearly still not well enough resolved for coastal areas within 100 km of the coast.

Once the data have been collated and assigned to a given cell, they are available for analysis. They can be accumulated into different time "bins" according to the particular questions being asked. For example, to produce Figure 8.12a, which is based on a 5-year span of altimeter data, all the samples from December, January, and February for each year have been averaged at each cell, to characterize winter values, and similarly for the other three seasons. The result clearly displays the seasonality of the mean wave field off northwest Europe. However, because the data were originally binned by month, preliminary analysis was able to confirm that December, January, and February were the three months with highest waves, which led to this particular allocation of months to seasons, which is different from those used for other purposes. It is worth noting that other variables could be

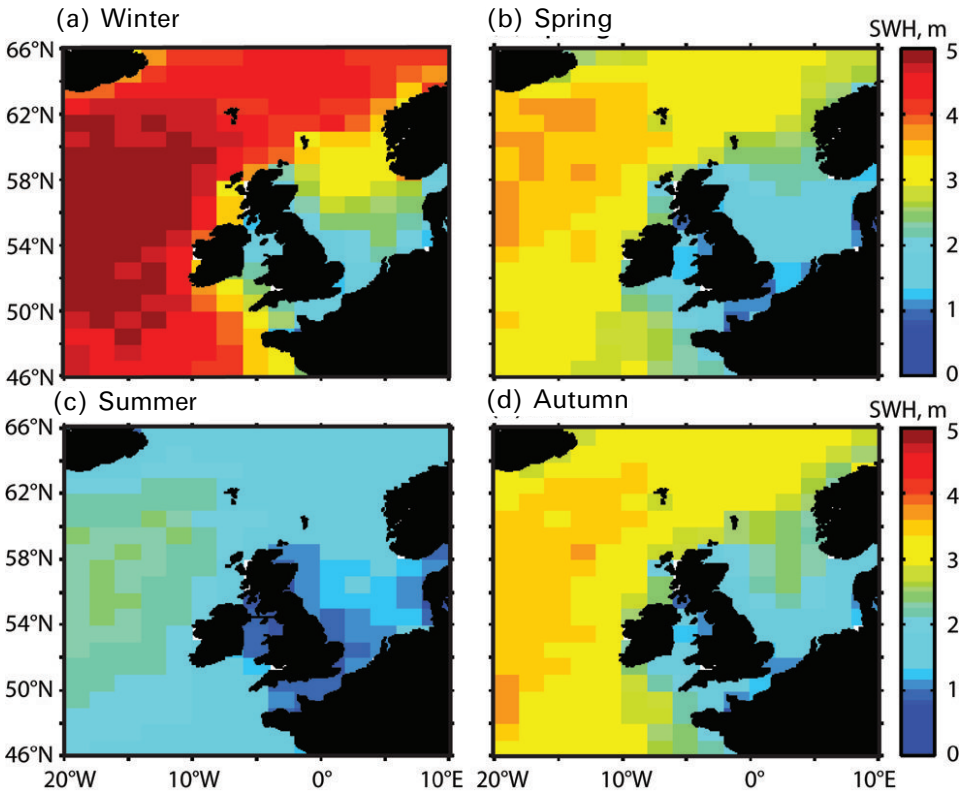


Figure 8.12. Seasonal climatology of mean significant wave height, H_S, over the northeast Atlantic, derived from 5 years of altimeter measurements between 1993–1997, binned monthly on a 1° latitude × 2° longitude grid. The color scale is in meters. (a) Dec, Jan, Feb. (b) Mar, Apr, May. (c) June, July, Aug. (d) Sep, Oct, Nov. (Woolf *et al.*, 2003).

plotted, such as maximum waves, or the mean of the highest 10% of samples, or some other property of the distribution of values contained in the climatological database. Furthermore, the database could contain quantities other than H_S, such as the dominant wave period or statistical properties like skewness and kurtosis of the sea surface height distribution that potentially are retrievable from altimeters (see Section 8.2.4).

Exploiting satellite-derived wave climatologies

There is a great opportunity to exploit the wave data being assembled from satellite altimeters, although little has been published so far. Many applications relate to planning and design of offshore engineering structures, or the specification for ships operating in particular ocean regions. There are several agencies selling their expert services and wave data to engineering companies that require reliable infor-

mation on wave climatology, although little of this work is published in the scientific literature.

One particular application is prediction of extreme wave heights (Cooper and Fornstall, 1997; Panchang *et al.*, 1999; Henrique *et al.*, 2000, 2003; Chen *et al.*, 2004). Extreme wave statistics are often expressed in terms of the ''100-year return'' wave height, the height of wave at a given location that is expected to be reached or exceeded only once per century (Muir and El-Sharaawi, 1986). Predicting this from a database acquired over only 20 years requires extrapolation that is based on the assumption of a particular mathematical model of wave height distribution, taking into account the particular characteristics of a satellite-based wave climatology (Wimmer *et al.*, 2006). In the cited example, the motivation was in the context of developing offshore wave power devices for renewable energy. This is a growing application area whose practitioners need to know both (1) mean wave conditions as a measure of potential power available and the proportion of time during which the wave height is low, when the power plant would be idle, and (2) the extent of occasions when wave height would be too great to allow safe operation. These are precisely the types of information that a wave climate database can provide. Having such data available in a geographically distributed form allows appropriate locations to be detected for different types of wave power devices.

A very interesting scientific use of new, altimeter-derived, wave climatologies has come from studies which attempt to explain wave distribution in terms of other physical forcing. For example, the records from wave buoys in the North Atlantic had shown strong variability at interannual and longer timescales (Bacon and Carter, 1991), at one stage appearing to increase significantly (relative to mean seasonal climatology) during a whole decade. The extended wave climatology gained by the use of altimetry from the mid-1980s and through the 1990s was able to confirm that this trend did reverse and was simply part of a longer oscillation (Woolf *et al.*, 2002, 2003). In fact analysis revealed a correlation between waves in the northeast Atlantic and the North Atlantic Oscillation (NAO). The NAO is a climatological index based on the difference in sea level atmospheric pressure between the Azores and Iceland. It therefore serves to represent in a gross sense the strength and track of winds over the northeast Atlantic and it is consistent to find that it correlates with monthly mean wave height in winter months when waves are greatest.

Because the altimeter provides a spatial distribution of wave climatology, the analysis could be taken a step further than was possible with the wave record from a few isolated buoys. By correlating the NAO with the time series of the H_S monthly mean anomaly (relative to the climatological seasonal cycle) at every grid cell it was possible to show the geographical extent over which correlation was significant. This means that in those areas where correlation is high (in some cases up to 0.85) we can account for much of interannual to decadal–scale variability of wave height. Thus the wave statistics for a particular month can be predicted if we know the NAO index during that month. One implication of this is that, if climate change–forecasting models are able to predict NAO conditions in a future warmer planet, we shall be able to draw conclusions about the likely wave climate also. Figure 8.13

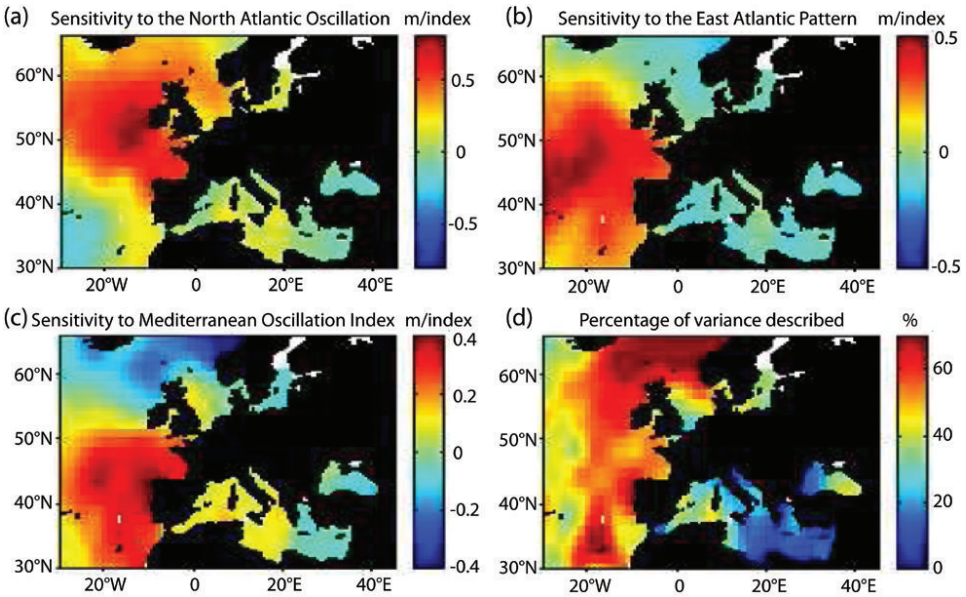


Figure 8.13. Sensitivity of wintertime significant wave height to (a) the North Atlantic Oscillation, (b) East Atlantic Pattern, and (c) Mediterranean Oscillation Index (in meters/index). (d) Total fraction of interannual variability described by a linear relationship to all three indices. Values are derived from a $1.5^\circ \times 1.5^\circ$ climatology of TOPEX for 1993–2002 (adapted from Woolf and Gommenginger, 2008).

illustrates the geographical distribution of the correlation of winter, monthly, wave height anomalies and climatological indices. These results (Woolf and Gommenginger, 2008) are based on an H_S anomaly relative to climatology for the winter months December to March sampled at $1.5^\circ \times 1.5^\circ$ from TOPEX data for the 10-year period 1993–2002. Correlation has been made against not only the NAO but also another routinely produced index, the East Atlantic Pattern (EAP), and a recently proposed new index, the Mediterranean Oscillation Index (MOI) defined by the pressure difference between mid-northern Atlantic and southeastern Mediterranean (Grbec *et al.*, 2003). It is evident that the different indices relate to waves in different regions, as should be expected since they serve to integrate the influence of winds over different parts of the ocean. The fourth panel in Figure 8.13 shows how much of the variability in wave height can be explained by a linear model based on all three indices.

8.7 ASSESSMENT AND FUTURE PERSPECTIVES

Ocean wave measurement from space is at an exciting point in its development. The ability of altimeters to measure H_S has been thoroughly confirmed and a number of

useful operational applications have been demonstrated, particularly in support of wave modeling and wave climatology. The future looks promising for an operationally reliable series of altimeters through the next two decades, on ESA's Sentinel 3 platforms in support of European operational oceanography (see Chapter 14) and also on the joint U.S./Europe Jason series. Both instruments are absolutely essential for their combined capacity to monitor mesoscale variability of the ocean, and equally for their contribution to ocean wave monitoring. However, the potential benefits of altimetry to wave forecasting and climatology will not be reached without significant takeup of the methodology by the ocean engineering community, and this would be jeopardized by any uncertainty about continuity of the altimeter program.

At the same time, it is important not to lose sight of the advantages of having a more dense altimetry-sampling capability which would allow wave-warning services to use direct observations of waves independently of wave model predictions. To do this effectively needs a constellation of altimeters, ideally on at least 12 platforms. A system has been proposed in which up to eight microsatellites, each carrying an altimeter and weighing less than 80 kg (about 1% of the weight of Envisat), could be launched by a single rocket. To increase the coverage and to ensure overlap between different satellite tracks for cross-calibration purposes, the satellites should be in more than one orbit plane. The system has been named GANDER (Global Altimeter Network Designed to Evaluate Risk). By using the constellation of microsatellite altimeters in parallel with a high-specification altimeter such as Jason, sea surface height as well as significant wave height could be retrieved to an acceptable accuracy for use by global ocean circulation models. This would make it more cost-effective, the total cost of a 24-microsatellite system requiring three rocket launches being estimated at around only 10% of the cost of Envisat (Allan, 2006). Such a system would also offer a capability in monitoring storm surges and tsunamis (as discussed briefly in Chapter 11).

Recent advances in wave spectra retrieval from SAR are also exciting. Here too there is the prospect of wave-forecasting agencies starting to assimilate data into their models and the coastal engineering community developing applications of what is a rich source of detailed information about sea state in coastal regions. In coming years this should be an exciting focus for both oceanographic research and applied coastal engineering. For those who have struggled for decades to tease reliable quantitative information from SAR images, the calibrated products illustrated in Figures 8.9 to 8.11 seem remarkable. Of course for each of these clear images there may be many others where the results of new processing techniques are less convincing, but nonetheless it is important that ESA should be able to process sufficient data to meet the needs of researchers who will develop applications of SAR data that could make a big difference to applied coastal oceanography.

There are also opportunities for more fundamental research to take satellite wave oceanography forward. This includes developing techniques to retrieve wave period, skewness, and other wave properties from the altimeter signal. The application of altimetric wave-monitoring techniques in coastal and shelf waters would benefit from the development of a swath altimeter, but until then there are gains

to be won from the current initiatives to get more out of existing altimeters near the coast.

> "Satellite altimetry can play an important role in establishing links between atmospheric variability and coastal vulnerability" (Woolf and Gommenginger, 2008).

There are other techniques, not mentioned so far in this chapter, which also could enrich the sources of remotely sensed wave data. These are the knife beam altimeter (Karaev *et al*., 2005) and the exploitation of bistatically reflected GPS signals to provide more information about sea state (Gleason *et al*., 2005). These are both techniques to watch for the future.

8.8 REFERENCES

Abdalla, S., P. Janssen, and J.-R. Bidlot (2003), Use of satellite data and enhanced physics to improve wave predictions. In: J. M. Smith (Ed.), *Coastal Engineering 2002* (Vol. 1, pp. 87–96). World Scientific, Singapore.

Allan, T. (2006), The story of GANDER. *Sensors*, **6**, 249–259.

Alpers, W. (2003), Ocean surface wave imaging from Seasat to Envisat. Paper presented at *Proc. Geoscience and Remote Sensing Symposium, IGARSS '03, Toulouse* (Vol. 1, pp. 35–37). IEEE International.

Alpers, W., and C. L. Rufenach (1979), The effect of orbital motions on synthetic aperture radar images of ocean waves. *IEEE Trans. Antennas Propagat*., **AP-27**, 685–690.

Alpers, W. R., and C. Brüning (1986), On the relative importance of motion-related contributions to the SAR imaging mechanism of ocean surface waves. *IEEE Trans. Geosc. Remote Sensing*., **GE-24**(6), 873–885.

Alpers, W., D. B. Ross, and C. L. Rufenach (1981), On the detectability of ocean surface waves by real and synthetic aperture radar. *J. Geophys. Res*., **86**(C), 6481–6498.

Ardhuin, F., F. Collard, and B. Chapron (2004), Wave spectra from ENVISAT's synthetic aperture radar in coastal areas. Paper presented at *Proc. 14th International Offshore and Polar Engineering Conference, Toulon, France* (pp. 221–225). International Society of Offshore and Polar Engineers, Cupertino, CA.

Bacon, S., and D. J. T. Carter (1991), Wave climate changes in the North Atlantic and North Sea. *Int. J. Climatol*., **11**, 545–558.

Bauer, E., S. Hasselmann, and K. Hasselmann (1992), Validation and assimilation of Seasat altimeter wave heights using the WAM wave model. *J. Geophys. Res*., **97**, 12671–12682.

Bauer, E., K. Hasselmann, I. R. Young, and S. Hasselmann (1996), Assimilation of wave data into the wave model WAM using an impulse response function method. *J. Geophys. Res*., **101**, 3801–3816.

Benveniste, J., and M. P. Milagro (2000), *Envisat RA-2 and MWR Products and Algorithms User Guide* (ESA Envisat Report, RA-TN-ESR-GS-0013, 13 pp.). ESA, Noordwijk, The Netherlands.

Booij, N., R. C. Ris, and L. H. Holthuijsen (1999), A third-generation wave model for coastal regions, 1: Model description and validation. *J. Geophys. Res*., **104**, 7649–7666.

Breivik, L. A., and M. Reistad (1994), Assimilation of ERS-1 altimeter wave heights in an operational numerical wave model. *Weather and Forecasting*, **9**(3).

Breivik, L. A., M. Reistad, H. Schyberg, J. Sunde, H. E. Krogstad, and H. Johnsen (1998), Assimilation of ERS SAR wave spectra in an operational wave model. *J. Geophys. Res.*, **103**(C4), 7887–7900.

Brüning, C., S. Hasselmann, and K. Hasselmann (1994), First evaluation of ERS-1 synthetic aperture radar wave mode data. *Global Atmos. Ocean. Syst.*, **2**, 61–98.

Caires, S., A. Sterl, and C. P. Gommenginger (2005), Global ocean mean wave period data: Validation and description. *J. Geophys. Res.*, **110**(C02003), doi: 10.1029/2004JC002631.

Carter, D. J. T., S. Foale, and D. J. Webb (1991), Variations in global wave climate throughout the tear. *Int. J. Remote Sensing*, **12**(8), 1687–1697.

Carter, D. J. T., P. G. Challenor, and M. A. Srokosz (1992), An assessment of Geosat wave height and wind speed measurements. *J. Geophys. Res.*, **99**, 25015–25024.

Cartwright, D. E., and M. S. Longuet-Higgins (1956), The statistical distribution of the maxima of a random function. *Proc. Roy. Soc. London A*, **237**, 212–232.

Cavaleri, L. (2006), Wave modeling: Where to go in the future. *Bull. Am. Meteorol. Soc.*, **87**(2), 207–214.

Challenor, P. G., S. Foale, and D. J. Webb (1990), Seasonal changes in the global wave climate measured by the Geosat altimeter. *Int. J. Remote Sensing*, **11**(12), 2205–2213.

Chapron, B., D. Vandemark, and F. C. Jackson (1994), Airborne measurements of the ocean's Ku-band radar cross section data at low incidence angles. *Atmosphere–Oceans*, **32**(1), 179–193.

Chen, G., S. W. Bi, and R. Ezraty (2004), Global structure of extreme wind and wave climate derived from TOPEX altimeter data. *Int. J. Remote Sensing*, **25**(5), 1005–1018.

Chen, G., B. Chapron, R. Ezraty, and D. Vandemark (2002), A global view of swell and wind sea climate in the ocean by satellite altimeter and scatterometer. *J. Atm. Ocean. Tech.*, **19**(11), 1849–1859.

Cooper, C. K., and G. Z. Fornstall (1997), The use of satellite altimeter data to estimate the extreme wave climate. *J. Atm. Ocean. Tech.*, **14**(2), 254–266.

Cotton, P. D., and D. J. T. Carter (1994), Cross calibration of TOPEX ERS-1 and Geosat wave heights. *J. Geophys. Res.*, **99**, 25025–25033.

Dunlap, E. M., R. B. Olsen, L. Wilson, S. D. Margerie, and R. Lalbeharry (1998), The effect of assimilating ERS-1 fast delivery wave data into the North Atlantic WAM model. *J. Geophys. Res.*, **103**(C4), 7901–7915.

Ebuchi, N., and H. Kawamura (1994), Validation of wind speeds and significant wave heights observed by the TOPEX altimeter around Japan. *J. Oceanogr.*, **50**, 479–487.

Emmanouil, G., G. Galanis, G. Kallos, L. A. Breivik, H. Heiberg, and M. Reistad (2007), Assimilation of radar altimeter data in numerical wave models: An impact study in two different wave climate regions. *Ann. Geophys.*, **25**, 581–595.

Engen, G., and H. Johnsen (1995), SAR–ocean wave inversion using image cross-spectra. *IEEE Trans. Geosc. Remote Sensing.*, **33**(4), 1047–1056.

Fedor, L. S., and G. S. Brown (1982), Waveheight and windspeed measurements from the Seasat radar altimeter. *J. Geophys. Res.*, **87**(C), 3254–3260.

Foreman, S. J., M. W. Holt, and S. Kelsall (1994), Preliminary assessment and use of ERS-1 altimeter wave data. *J. Atm. Ocean. Tech.*, **11**, 1370–1380.

Gleason, S., S. Hodgart, Y. Sun, C. P. Gommenginger, S. Mackin, M. Adjrad, and M. Unwin (2005), Detection and processing of bistatically reflected GPS signals from low Earth orbit for the purpose of ocean remote sensing. *IEEE Trans. Geosc. Remote Sensing*, **43**(6), 1228–1241.

Gomez-Enri, J., C. P. Gommenginger, M. A. Srokosz, P. G. Challenor, and J. Benveniste (2007), Measuring global ocean wave skewness by retracking RA-2 ENVISAT waveforms. *J. Atmos. Oceanic Tech.*, **24**(6), 1102–1116.

Gommenginger, C. P., M. A. Srokosz, P. G. Challenor, and P. D. Cotton (2003), Measuring ocean wave period with satellite altimeters: A simple empirical model. *Geophys. Res. Letters*, **30**(22), doi: 10.1029/2003GL017743.

Gower, J. F. R. (1996), Intercalibration of wave and winds data from TOPEX/Poseidon and moored buoys off the west coast of Canada. *J. Geophys. Res.*, **101**, 3817–3829.

Grbec, B., M. Morovic, and M. Zore-Armanda (2003), Mediterranean Oscillation Index and its relationship with salinity fluctuation in the Adriatic Sea. *Acta Adriatica*, **44**(1), 61–76.

Greenslade, D. J. M. (2001), The assimilation of ERS-2 significant wave height data in the Australian region. *J. Marine Systems*, **28**(1/2), 141–160.

Gulev, S. K., V. Grigorieva, A. Sterl, and D. K. Woolf (2003), Assessment of the reliability of wave observations from voluntary observing ships: Insights from the validation of a global wind wave climatology based on voluntary observing ship data. *J. Geophys. Res.*, **108**(C7), 3236, doi: 10.1029/2002JC001437.

Hasselmann, K. (1962), On the non-linear energy transfer in a gravity wave spectrum, Part 1: General theory. *J. Fluid Mech.*, **12**, 482–500.

Hasselmann, K., and S. Hasselmann (1991), On the non-linear mapping of an ocean wave spectrum into a synthetic aperture radar image spectrum and its inversion. *J. Geophys. Res.*, **96**(C6), 10713–10729.

Hasselmann, S., C. Brüning, K. Hasselmann, and P. Heimbach (1996), An improved algorithm for the retrieval of ocean wave spectra from SAR image spectra. *J. Geophys. Res.*, **101**(C), 16615–16629.

Hauser, D., E. Soussi, and L. R. Thouvenot (2001), SWIMSAT: A real aperture radar to measure directional spectra of ocean waves from space—Main characteristics and performance simulation. *J. Atmos. Oceanic Tech.*, **18**(3), 421–437.

Heimbach, P., and K. Hasselmann (2000), Development and application of satellite retrievals of ocean wave spectra. In: D. Halpern (Ed.), *Satellite Oceanography and Society* (Elsevier Oceanography Series, Vol. 63, pp. 5–33). Elsevier Science, Amsterdam.

Heimbach, P., S. Hasselmann, and K. Hasselmann (1998), Statistical analysis and intercomparison of WAM model data with global ERS-1 SAR wave mode spectral retrievals over 3 years. *J. Geophys. Res.*, **103**, 7931–7977.

Henrique, J., G. M. Alves, and I. R. Young (2000), Extreme significant wave heights from combined satellite altimeter data. In: B. L. Edge (Ed.), *Proc. 27th International Conference on Coastal Engineering Sidney, Australia, July 16–21* (pp. 1064–1077). American Society of Civil Engineers, Reston, VA.

Henrique, J., G. M. Alves, and I. R. Young (2003), On estimating extreme wave heights using combined Geosat, Topex/Poseidon and ERS-1 altimeter data. *Applied Ocean Research*, **25**(4), 167–186.

Holthuijsen, L. H. (2007), *Waves in Oceanic and Coastal Waters* (404 pp.). Cambridge University Press, Cambridge, U.K.

Holthuijsen, L. H., N. Booij, M. van Endt, S. Caires, and C. G. Soares (1997), Assimilation of buoy and satellite data in wave forecasts with integral control variables. *J. Marine Systems*, **13**(1), 21–31.

Jackson, F. C. (1987), The radar ocean–wave spectrometer. *Johns Hopkins APL Technical Digest*, **8**, 70–74.

Jackson, F. C., W. T. Walton, and C. Y. Peng (1985a), Aircraft and satellite measurement of ocean wave directional spectra using scanning-beam microwave radars. *J. Geophys. Res.*, **90**, 987–1004.

Jackson, F. C., W. T. Walton, and C. Y. Peng (1985b), A comparison of in situ and airborne radar measurements of ocean wave directionality. *J. Geophys. Res.*, **90**, 1005–1018.

Jones, I. S. F., and Y. Toba (Eds.) (2001), *Wind Stress over the Ocean* (326 pp.). Cambridge Uiniversity Press, Cambridge, U.K.

Kalantzi, G. D., C. P. Gommenginger, and M. A. Srokosz (2009), Assessing the performance of the dissipation parameterizations in WAVEWATCH III using collocated altimetry data. *J. Phys. Oceanogr.*, **39**(11), 2800–2819.

Karaev, V. Y., M. B. Kanevsky, G. N. Balandina, P. G. Challenor, C. P. Gommenginger, and M. A. Srokosz (2005), The concept of a microwave radar with an asymmetric knifelike beam for the remote sensing of ocean waves. *J. Atm. Ocean. Tech.*, **22**, 1809–1820.

Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. Janssen (1994), *Dynamics and Modelling of Ocean Waves* (532 pp.). Cambridge University Press, Cambridge, U.K.

Krogstad, H. E. (1992), A simple derivation of Hasselmann's nonlinear ocean-synthetic aperture radar transform. *J. Geophys. Res.*, **97**(C2), 2421–2425.

Krogstad, H. E., and S. F. Barstow (1999), Satellite wave measurements for coastal engineering applications. *Coastal Engineering*, **37**, 283–307.

LeBlond, P. H. (2002), Ocean waves: Half-a-century of discovery. *J. Oceanogr.*, **58**, 3–9.

LeBlond, P. H.. and L. A. Mysak (1978). *Waves in the Ocean* (Elsevier Oceanography Series, 602 pp.). Elsevier, Amsterdam.

Lionello, P., H. Günther, and P. Janssen (1992), Assimilation of altimeter data in a global third-generation wave model. *J. Geophys. Res.*, **97**(C9), 14453–14474.

Mackay, E. B. L., C. H. Retzler, P. G. Challenor, and C. P. Gommenginger (2008), A parametric model for ocean wave period from Ku band altimeter data. *J. Geophys. Res.*, **113**(C03029), doi: 10.1029/2007JC004438.

Mastenbroek, C., and C. F. d. Valk (2000), A semi-parametric algorithm to retrieve ocean wave spectra from SAR. *J. Geophys. Res.*, **105**(C2), 3497–3516.

Mastenbroek, C., V. K. Makin, A. C. Voorrips, and G. J. Komen (1994), Validation of ERS-1 altimeter wave height measurements and assimilation in a North Sea wave model. *Global Atmos. Ocean. Syst.*, **2**, 143–161.

Muir, L. R., and A. H. El-Sharaawi (1986), On the calculation of extreme wave heights: A review. *Ocean Engineering*, **13**(1), 93–118.

Panchang, V., L. Z. Zhao, and Z. Demirbilek (1999), Estimation of extreme wave heights using GEOSAT measurements. *Ocean Engineering*, **26**(3), 205–225.

Quilfen, Y., B. Chapron, and M. Serre (2004), Calibration/validation of an altimeter wave period model and application to TOPEX/Poseidon and Jason-1 altimeters. *Marine Geodesy*, **27**, 535–549.

Romeiser, R. (1993), Global validation of the wave model WAM over a one-year period using Geosat wave height data. *J. Geophys. Res.*, **98**(C3), 4713–4726.

Schulz-Stellenfleth, J., and S. Lehner (2002), Spaceborne synthetic aperture radar observations of ocean waves traveling into sea ice. *J. Geophys. Res.*, **107**(C8), 3106, doi: 10.1029/2001JC000837.

Schulz-Stellenfleth, J., S. Lehner, and D. Hoja (2005), A parametric scheme for the retrieval of two-dimensional ocean wave spectra from synthetic aperture radar look cross spectra. *J. Geophys. Res.*, **110**(C05004), doi: 10.1029/2004JC002822.

Schulz-Stellenfleth, J., T. König, and S. Lehner (2007), An empirical approach for the retrieval of integral ocean wave parameters from synthetic aperture radar data. *J. Geophys. Res.*, **112**(C03019), doi: 10.1029/2006JC003970.

Stewart, R. H. (2008), *Introduction to Physical Oceanography* (e-book). Texas A & M University, available at *http://oceanworld.tamu.edu/home/course_book.htm*

Tolman, H. L., B. Balasubramaniyan, L. D. Burroughs, D. V. Chalikov, Y. Y. Chao, H. S. Chen, and V. M. Gerald (2002), Development and implementation of wind generated ocean surface wave models at NCEP. *Weather and Forecasting*, **17**, 311–333.

Tucker, M. J. (1991), *Waves in Ocean Engineering: Measurement, Analysis, Interpretation* (431 pp.). Ellis Horwood, Chichester, U.K.

Vandemark, D., F. C. Jackson, E. J. Walsh, and B. Chapron (1994), Airborne radar measurements of ocean wave spectra and wind speed during the Grand Banks ERS-1 SAR Wave Experiment. *Atmosphere–Oceans*, **32**(1), 143–178.

Violante-Carvalho, N., I. S. Robinson, and J. Schulz-Stellenfleth (2005), Assessment of ERS synthetic aperture radar wave spectra retrieved from the Max-Planck-Institut (MPI) scheme through intercomparisons of 1 year of directional buoy measurements. *J. Geophys. Res.*, **110**(C07019), doi: 10.1029/2004JC002382.

WAMDI_Group (1988), The WAM model: A third generation ocean wave prediction mode. *J. Phys. Oceanogr.*, **18**, 1775–1810.

Wimmer, W., P. G. Challenor, and C. H. Retzler (2006), Extreme wave heights in the North Atlantic from altimeter data. *Renewable Energy*, **31**, 241–248.

Woolf, D. K., and C. P. Gommenginger (2008), Radar altimetry: Introduction and application to air–sea interaction. In: V. Barale and M. Gade (Eds.), *Remote Sensing of the European Seas* (pp. 283–294). Springer-Verlag, Berlin.

Woolf, D. K., P. G. Challenor, and P. D. Cotton (2002), The variability and predictability of North Atlantic wave climate. *J. Geophys. Res.*, **107**(C10), 3145, doi: 10.1029/2001JC001124.

Woolf, D. K., P. D. Cotton, and P. G. Challenor (2003), Measurements of the offshore wave climate around the British Isles by satellite altimeter. *Phil. Trans. Roy. Soc. London A.*, **361**, 27–31.

Young, I. R., and T. J. Glowacki (1996), Assimilation of altimeter wave height data into a spectral wave model using statistical interpolation. *Ocean Engineering*, **23**(8), 667–689.

Zanife, O. Z., J. P. Dumont, J. Stum, and T. Guinle (2001), *SSALTO Products Specifications, Volume 1: Jason-1 User Products* (CNES Report, SMM-ST-M-EA-10879-CN, 36 pp.). Centre National d'Études Spatiales, Toulouse, France

9

Wind over the sea

This is a book about satellite oceanography and so this chapter about satellite-based measurements of wind and phenomena related to the wind does not attempt to address the topic from the meteorologist's point of view. For that there are texts devoted to satellite meteorology (e.g., Kidder and Vonder Haar, 1995). Instead, the objective is to review briefly, from the oceanographer's perspective, the circumstances in which knowledge of wind above the ocean is important for marine science and to assess whether, and in which circumstances, satellite data can provide it more appropriately than other sources.

The first section (Section 9.1) outlines different sensors and methods used to measure the wind field over the ocean from satellites, summarizing the more detailed treatment of techniques given in the companion volume *MTOFS* (Robinson, 2004). Section 9.2 identifies the variety of situations where wind data are important for the study of oceanographic processes. Many of these are topics discussed in detail in other chapters, so the discussion here is about the extent to which satellite wind measurements are preferable to the use of alternative data sources. Section 9.3 is a more substantial section about hurricanes over the ocean, how they are monitored by remote sensing, and particularly how satellite ocean data from all types of sensor can be used to reveal the impact of a hurricane's passage on the structure of the water column. The final section (Section 9.4) explores the use of fine-resolution radar mapping of wind distribution and its potential for operational exploitation, especially in relation to the location and planning of offshore wind power installations.

9.1 MEASURING WIND OVER THE SEA FROM SATELLITES

Everyone has first-hand experience of the effect of the wind blowing over a water surface. Gentle winds produce small ripples, which become steeper and longer as the

wind increases. We learn to read the strength of the wind in the roughness of the sea surface. In a similar way, the remote sensing of winds over the sea is possible because microwave sensors can detect the mean square slope of short waves and thence quantify the wind speed.

It is the use of microwaves in C, X, and Ku bands, with frequencies between 5 GHz and 20 GHz that has proved to be most effective for measuring wind speed. Figure 2.20 shows how radar backscatter at these wavelengths (between 1 and 6 cm) varies with incidence angles at different wind speeds. For oblique-viewing radars the measured radar backscatter cross-section, σ_0, increases with the increasing amplitude of short surface waves at the Bragg wavelength, comparable with the radar wavelength, which grow with increasing wind (as discussed in chapter 9 of *MTOFS*). Wind can therefore be measured by three different types of active radar instrument and also by passive microwave radiometry (as described in the four following subsections). Liu and Xie (2006) provide a review of the subject.

There are other remote-sensing methods used by satellite meteorologists to measure wind speed at higher altitudes, such as tracking clouds and new techniques using lidar. However, these have little direct relevance to oceanography, for which it is surface wind that interacts with the ocean and is of particular interest.

9.1.1 Scatterometry

Scatterometers are the most effective sensors for mapping the distribution of wind speed and direction. Different instrument designs and how they work are explained in chapter 9 of *MTOFS*. Essentially they measure the oblique backscatter, σ_0, from the same patch of sea viewed from several directions. They then use an empirical model such as CMOD4, which specifies σ_0 for different wind speeds and viewing directions relative to the wind, to retrieve an estimate of wind speed and direction (Stoffelen and Anderson, 1997).

At the time of writing two scatterometers are in operation (see Table 2.9). NASA's SeaWinds Ku-band scatterometer on QuikScat was deployed in 1999 and has delivered high-quality data since then, based on mature Ku-band backscatter models (Donnelly *et al.*, 1999; Ebuchi *et al.*, 2002). Its conically scanning antennas sweep out a swath of about 1,400 km, providing about 90% global coverage each day, and two samples a day of the wind vector at many places, with a spatial resolution of 12.5 km. Figures 9.1 and 9.2 show examples of the wind field as detected on the same day by QuikScat from, respectively, its morning (ascending) passes and evening (descending) passes on August 9, 2008. The time of acquisitions is later towards the west. The main wind patterns are captured well, and comparison between morning and evening passes shows the changes that occur within about 12 hours. QuikScat data are capable of measuring high wind speeds effectively (Liu, 2002) and have made a significant impact on ocean weather prediction (Von Ahn *et al.*, 2006).

The C-band ASCAT on MetOp was launched in 2006 and declared operational on May 15, 2007 as part of Eumetsat's new polar system (Gelsthorpe *et al.*, 2000; Figa-Saldaña *et al.*, 2002). An evolution from the AMI scatterometer on ESA's

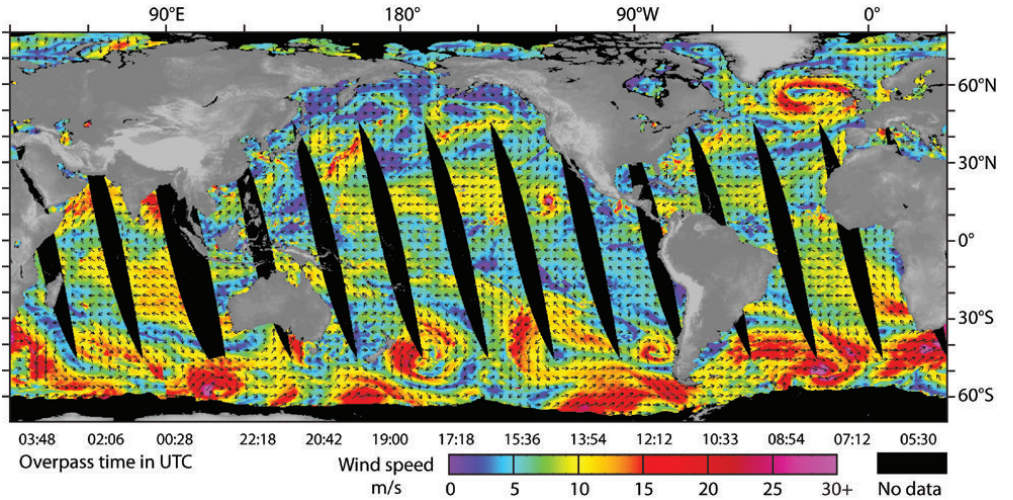


Figure 9.1. Example of daily coverage of ocean surface winds measured by ascending (morning) passes of QuikScat on August 9, 2008. Note that the Equator crossing time for each pass is shown on the lower x-axis. QuikScat data are produced by Remote Sensing Systems and sponsored by the NASA Ocean Vector Winds Science Team (mage adapted from a graphic image map acquired from *www.remss.com*).

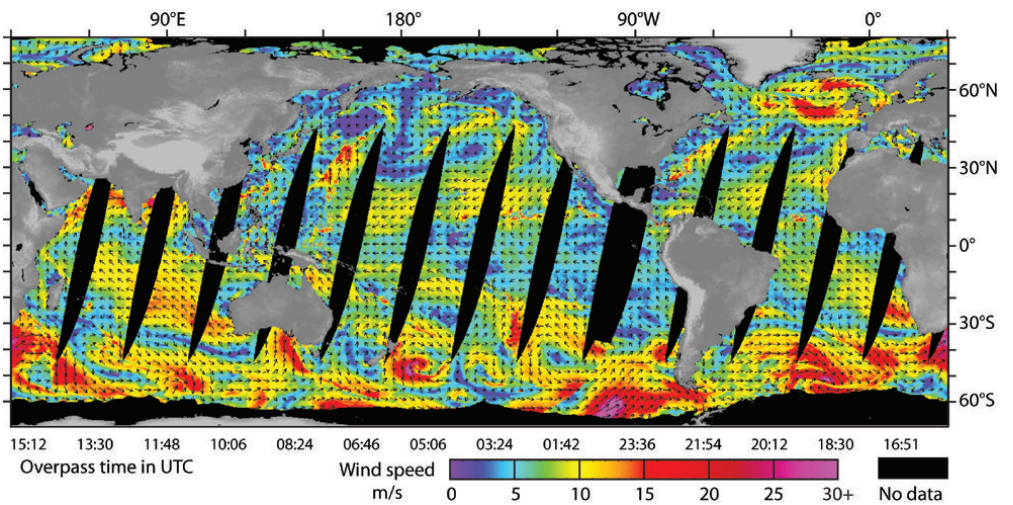


Figure 9.2. Example of daily coverage of ocean surface winds measured by descending (evening) passes of QuikScat on August 9, 2008. QuikScat data are produced by Remote Sensing Systems and sponsored by the NASA Ocean Vector Winds Science Team (image adapted from a graphic image map acquired from *www.remss.com*).

ERS-1 and ERS-2 satellites, the ASCAT has fixed antennas pointing on both sides of the scatterometer, each with a 500 km swath, to achieve almost global coverage every day at a resolution of 50 km. There is also an experimental data product with a resolution of 25 km. It uses an improved backscatter model, CMOD5 (Hersbach *et al.*, 2007). This is the first of a series of three European operational meteorological polar satellites to which Eumetsat and ESA are committed up to 2020.

9.1.2 Wind data from SAR

Since synthetic aperture radar (SAR) produces fine-resolution maps of σ_0, and wind is the primary agent that determines σ_0 before it has been modulated by ocean surface dynamical processes, SAR offers a direct means of mapping small-scale variability in the near-surface wind field. However, unlike scatterometry, the SAR does not view the sea surface from widely differing directions and therefore it is necessary to know the wind direction before scatterometer backscatter models such as CMOD4 can be applied. More information about SARs current at the time of writing is given in Table 2.8.

Section 10.8 of *MTOFS* identifies some of the short-lengthscale atmospheric phenomena that can be detected using SAR and discusses ways in which wind direction can be estimated using clues in the SAR image itself, such as roll vortices in the atmospheric or ocean surface boundary layer that produce high roughness streaks across the water surface that align within 5° or 10° of the wind direction. Using this, or external knowledge from forecast wind direction, allows wind speed and its variability to be retrieved from SAR image data. A number of processing systems have been developed to retrieve wind speeds from SAR (Vachon and Dobson, 1996; Kerbaol *et al.*, 1998; Korsbakken *et al.*, 1998; Lehner *et al.*, 1998; Fichaux and Ranchin, 2002).

In general, SAR is not an appropriate means for determining the wind speed for large-scale applications, for which a scatterometer is much more suitable. However, there are circumstances (as discussed in Section 9.4) where information about wind speeds and their distribution in the coastal zone offers unique benefits for certain applications.

9.1.3 Wind data from altimeters

Chapter 11 of *MTOFS* considers all aspects of altimetry over the ocean, one of which is the capacity of radar altimeters to measure wind speed. The altimeter, being a nadir-viewing (zero incidence angle) radar responds differently to wind than an oblique-viewing scatterometer or SAR. The maximum signal is returned when the sea is calm and increased scattering in high winds reduces the magnitude of the radar echo. This behavior is summarized in the low-incidence part of Figure 2.20. This simply means that an empirical algorithm to retrieve the wind has a different form, but it is still an effective way of measuring wind speed (Brown *et al.*, 1981; Chelton and Wentz, 1986; Freilich and Challenor, 1994). Wind direction cannot be measured from an altimeter. The accuracy of wind retrieval algorithms

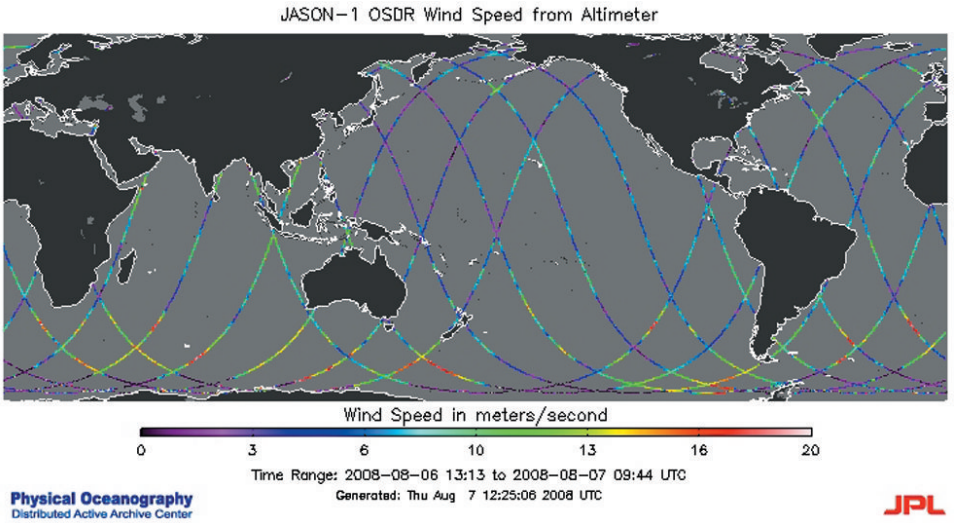


Figure 9.3. Example of wind speed data retrieved from the Jason-1 altimeter on August 6–7, 2008.

continues to be improved (e.g., by developing algorithms that relate the altimeter signal to coincident measurements from scatterometers at points widely distributed around the world—Gommenginger *et al.*, 2002), and use the altimeter wave height, H_S, as a second parameter in addition to σ_0, for driving wind speed or wind stress algorithms (Gourrion *et al.*, 2002).

The weakness of using altimetry for wind is sparse sampling from a point-sensing instrument compared with a scanning or swath-covering sensor like the scatterometer. Figure 9.3 shows the wind speed acquired by the Jason altimeter during a single day. Trying to fill in the gaps on subsequent days would be meaningless because the variability timescale for the wind field is no longer than a few hours, as illustrated in the differences between morning and evening scatterometer overpasses shown in Figure 9.1. For this reason the altimeter wind record is not often used as an operational means of retrieving winds unless there are no alternatives. Nonetheless it provides a valuable independent measurement for assessing the quality of other wind speed estimates, both from satellites and numerical weather prediction (NWP) analyses.

9.1.4 Microwave radiometry

As mentioned in Section 2.4.4 (and discussed in some depth by chapter 8 of *MTOFS*) the microwave radiation emitted by the sea surface in a particular direction depends not only on water temperature and its dielectric properties, but also on the orientation and shape of the sea surface. Consequently, wind speed is one of the parameters that can be retrieved from a multifrequency microwave radiometer using empirical

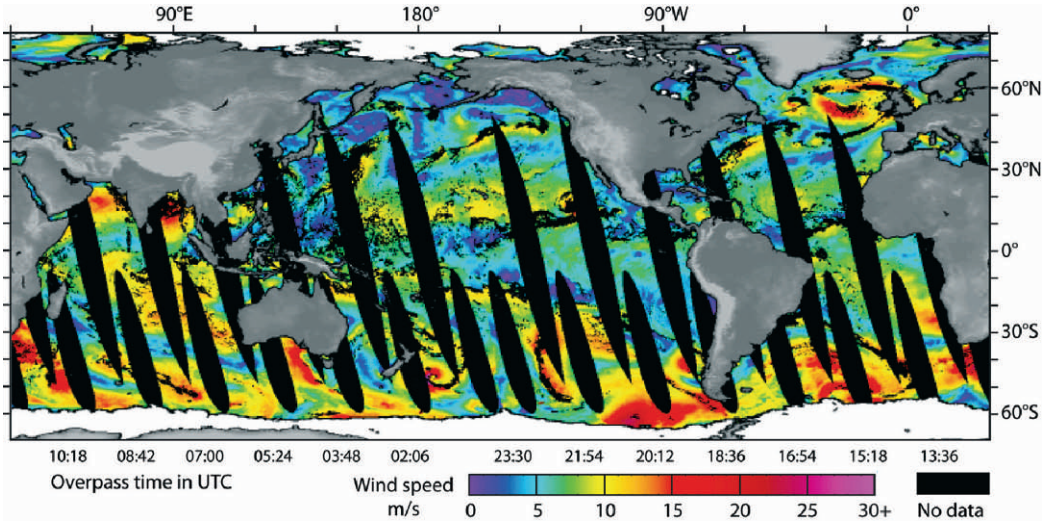


Figure 9.4. Example of wind speeds retrieved from ascending overpasses of the AMSR-E microwave radiometer on August 9, 2008, the same day as the data presented in Figure 9.1. AMSR-E data are produced by Remote Sensing Systems and sponsored by the NASA Earth Science REASoN DISCOVER Project and the AMSR-E Science Team (data are available at *www.remss.com*).

algorithms (Wentz, 1997). Measurable sensitivity to wind is found at frequencies between 6 GHz and 37 GHz, and so wind has been retrieved from the Special Sensor Microwave Imager (SSM/I) series of instruments flown since 1987 on the U.S. Defense Meteorological Satellite Program (DMSP) and made publicly available since 1992. The microwave imager TMI on the Tropical Rainfall Measuring Mission (TRMM) satellite has provided daily maps of wind speed since 1997, while the Advanced Microwave Scanning Radiometer (AMSR-E), launched in 2002, also continues to deliver wind speed as one of it several products. Figure 9.4 shows an example of daily global wind speed data from AMSR-E. None of these sensors is able to measure wind direction as well as speed.

In January 2003 the WindSat instrument on the Coriolis spacecraft was launched. WindSat is the first fully polarimetric, spaceborne, microwave radiometer. Its purpose was to measure partially polarized emission from the ocean surface with sufficient accuracy to be able to demonstrate the retrieval of vector winds (i.e., both speed and direction of the wind). Chapter 8 of *MTOFS* introduced the principles of how wind direction can be estimated using polarimetric radiometry. WindSat is largely an experimental mission to serve as a forerunner of the Conically-scanned Microwave Imager and Sounder (CMIS) proposed for the U.S. National Polar-orbiting Operational Environmental Satellite System (NPOESS). The attraction of such a sensor is that it measures several different ocean and atmospheric variables. If it can measure wind speed as reliably and accurately as a scatterometer then it promises to be a cheaper alternative. Problems were encountered at first with the

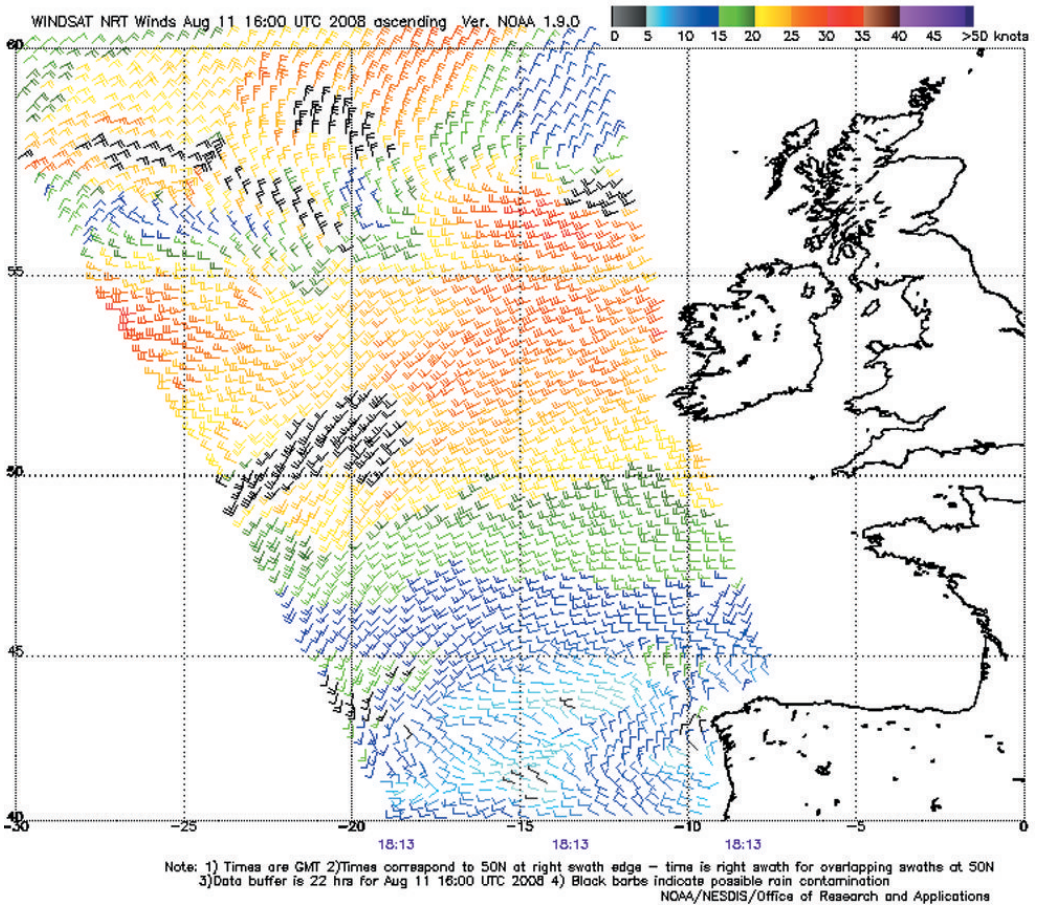


Figure 9.5. Example of wind vector output from Windsat produced in near-real time by NOAA.

onboard calibration load for WindSat but these seem to have been overcome. Evaluation of the polarimetric approach has been published to show that wind direction can be retrieved within 15° at wind speeds above 10 m/s (Brown *et al.*, 2006; Freilich and Vanhoff, 2006) and experimental data products of wind speed and direction are now being produced in near-real time. An example of the NOAA product is shown in Figure 9.5.

It appears that the U.S. NPOESS program has abandoned scatterometry in favor of polarimetric radiometers to meet its requirements to measure vector winds, although there is some uncertainty about how soon their intended operational sensor, CMIS, will be launched. It is no longer part of the payload for the NPOESS Preparatory Project (NPP) due for launch in 2011, while the first NPOESS platform will not launch till at least 2013, well beyond the expected life of QuikScat and WindSat. Meanwhile Eumetsat remain committed to maintaining the scatterometer

route to vector winds. It would seem sensible to run both systems in parallel since the approaches complement each other and some redundancy is necessary to preserve operational reliability, as well as to establish residual uncertainty of independent methods.

9.1.5 The alternatives to satellite measurements

Faced with a requirement to know the wind over the sea, ocean scientists generally turn to global wind forecasts of meteorological agencies, based on numerical weather prediction (NWP) models, if data are required in near-real time. Where oceanographic experiments are being analyzed some time after the event, or where winds are needed to drive ocean circulation models retrospectively, it is generally better to use reanalysis products. These are produced by the same agencies using similar models but with the benefit of including more, and better quality, observational data which has become available since real-time forecasting model runs were made. Two widely used sources for such data are NCEP and ECMWF although many other meteorological agencies now have a capacity for global NWP.

U.S. National Center for Environmental Prediction (NCEP) data, produced with the National Center for Atmospheric Research (NCAR), are available from the Climate Diagnostics Branch of the NOAA Earth System Research Laboratory in a variety of forms of Reanalysis Data Composite within about 2 days after real time. Figure 9.6 shows an example of the surface wind vector plot for August 9, 2008, the same day as satellite-retrieved data in Figures 9.1 and 9.4, for comparison. As well as producing global forecasts, the European Center for Medium-range Weather Forecasts (ECMWF) is producing a consistent reanalysis for the period from mid-1957 to 2001 called ERA-40. Its main objective is to promote the use of a consistent global analysis of the state of atmospheric, land, and surface conditions over that period, making it a very valuable source of wind and other atmospheric data for use in many types of ocean model.

Apart from data supplied by meteorological agencies, oceanographers may obtain their own measurements of atmospheric data from instrumented research vessels and buoys equipped with meteorological sensors. These have the merit that they provide a matching set of ocean and atmospheric measurements that are coincident in space and time. This is ideal for those ocean applications concerned with local air–sea interaction processes, such as wind wave generation or the formation of a local diurnal thermocline. They are less useful if the effect of wind fields needs to be known for a much wider area (e.g., in the generation of storm surges or the occurrence of upwelling). The other problem of ship-based meteorological measurements is that the presence of the vessel distorts the wind field and some of the other atmospheric variables (Yelland *et al.*, 2002). Shipboard anemometer measurements need to be adjusted in both speed and direction in order to represent the conditions at the standard height of 10 m if the ship had not been present. However, on all but the smallest of boats, corrections are difficult to calibrate in relation to the vessel's speed and orientation relative to the wind.

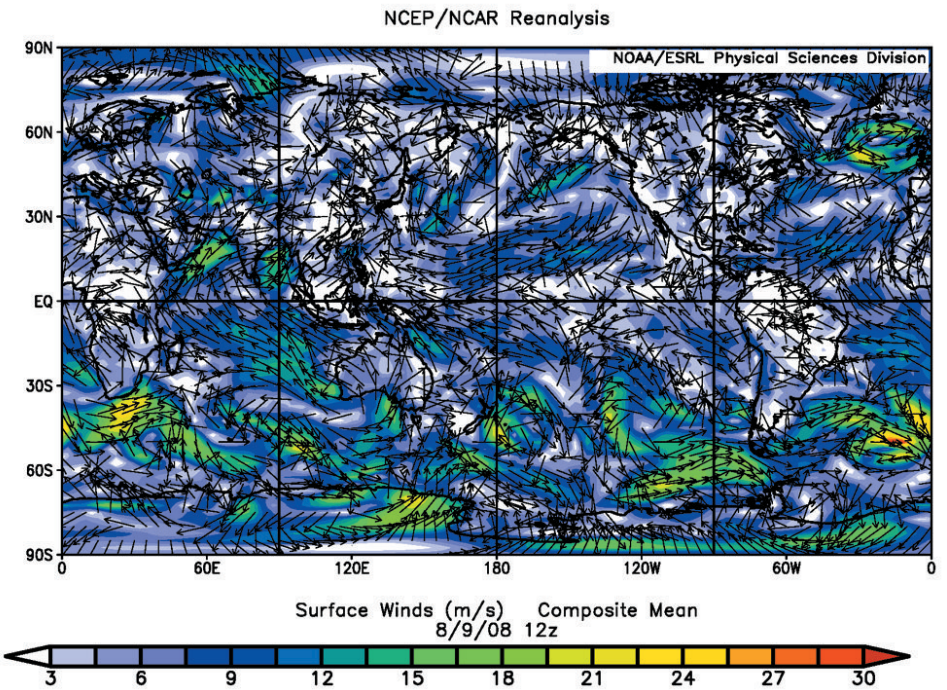


Figure 9.6. NCEP reanalysis of global winds at 12:00 UT on August 9, 2008, the same day as the data presented in Figures 9.1 and 9.4.

9.2 OCEANOGRAPHY AND WIND DATA

Since this chapter is focusing on the use of wind data for ocean applications, it is appropriate to ask whether wind data available from satellites are superior to the main alternative of using analyzed winds from NWP models. The latter can certainly be relied upon to provide a dynamically consistent set of winds, constrained by observations of atmospheric surface pressure and by data from satellite atmospheric-sounding sensors, with operational regularity, typically every 6 hours. Using such data offers a safe way of ensuring knowledge of the wider context of the wind field driving the ocean on scales of several hundred kilometers. However, on lengthscales of tens of kilometers and time scales of a few hours, present NWP forecasts are less reliable, although finer resolution models are being developed.

Thus if spatially detailed knowledge of wind field variability is needed (e.g., to identify wind shadow regions behind large islands, or to pinpoint the position of an atmospheric front) the 25 km resolution of scatterometry or microwave radiometers is attractive, although the limit of two overpasses per day from each of these sensors may be a drawback.. The SAR's unique capability for even finer scale definition of wind distribution in coastal waters opens up the possibility of understanding air–sea

interaction processes at the kilometer scale that could otherwise not be contemplated, even though such data snapshots are available only occasionally. Section 9.4 discusses some of the current applications of high-resolution wind maps.

9.2.1 Differences between analysis winds and satellite winds

It is not only the space-time sampling tradeoff that distinguishes between satellite observations and NWP-analyzed wind data. A very important issue for understanding precisely how the sea experiences wind at its surface concerns the atmospheric boundary layer (ABL). NWP "surface" winds, representing an estimate of the wind vector at the standard 10 m height above the sea surface, are normally parameterized using model wind in the bottom layer of an atmospheric circulation model, on the assumption that there is a neutrally stable ABL. If the boundary layer is unstable then the winds at the sea surface (i.e., at 10 m) will actually be stronger and in a different direction than those estimated on the basis of a stable ABL. NWP models do not presently resolve shear flow at the base of the ABL and no matter how accurate their analysis of wind vectors above this, they are not constructed to accurately project the velocity profile through the lower tens of meters down to the sea surface. Therefore the use of NWP or reanalysis wind fields may not tell oceanographers precisely what they need to know about the magnitude and direction of wind stresses acting directly on the sea surface.

If instead they make use of satellite-derived winds, they can be sure that these correspond precisely to the winds at the very surface, since all satellite radar wind retrieval methods actually use measurements of ocean surface roughness as a proxy for the wind. This is sometimes considered to be a drawback for meteorological applications of satellite-derived winds, because they do not relate precisely to the geostrophic wind vector at the top of the ABL, but for the oceanographer this is a clear benefit. This issue is developed further in Chapter 10 where we consider whether it is strictly wind speed or rather wind stress that is being measured by radar detection of surface roughness.

Overall it seems that there are many advantages to marine science making use of satellite-derived wind fields, although every application ought to be individually considered in the light of its particular needs.

9.2.2 Which type of wind data should be used to study ocean phenomena?

There are many oceanographic situations, in both research and operational applications, where knowledge of the wind speed or vector over the ocean is an essential part of the process. That is another way of saying that air–sea interaction processes are important at all horizontal lengthscales in the ocean.

For the study and simulation of large-scale air–sea interaction that produces wind-driven ocean circulation, analysis winds should be entirely satisfactory. Their smooth distribution in horizontal space and in time is suitable for forcing ocean

general circulation models (GCMs) whereas the less frequent temporal sampling of satellites and the edges of satellite sampling swaths may lead to discontinuities in the satellite-measured field which would require extra preprocessing so that they do not create spurious artifacts in the response of ocean GCMs. Even if there are sharp discontinuities in actual wind forcing (e.g., across atmospheric fronts), these are best smoothed out before forcing large-scale ocean models.

It is when oceanographers want to study smaller, mesoscale features in the ocean, at lengthscales from a few hundred kilometers down to tens of kilometers, that it becomes important to know the variability of the actual surface wind at similar lengthscales. This includes many of the oceanographic features that are discussed in other chapters of this book because they have surface signatures which make them amenable to observation using satellite ocean data. Thus, for example, understanding unexpected phytoplankton blooms, which may be related to vigorous localized wind mixing, ideally requires detailed knowledge of wind distribution at the fine scale achieved by satellite swath sensors. Wind-driven upwelling is another phenomenon where satellite winds can help to show the detailed spatial and temporal variability of the forcing that creates localized centers of upwelling that move from day to day. Analysis winds tend to smooth such variability. While their use in models may produce acceptable monthly averages of upwelling, analysis winds may fail to account for the patchy, day-to-day character of upwelling (as illustrated in Figures 5.3, 5.4, and 5.5).

For those features discussed in Chapter 5 whose geographical bounds are dependent on local wind forcing, such as the wakes behind islands (Section 5.4) and the ocean response to offshore wind jets (Section 5.2) it is essential to have high-resolution snapshots of the wind field that can delineate the extent of wind jets or sheltered regions. The same is true for certain wind-driven features in coastal and shelf seas (described in Chapter 13) and the storm surges mentioned in Chapter 11. Tropical instability waves (discussed in Section 6.6.2) exhibit complex atmosphere–ocean feedback at scales which NWP models may have difficulty matching, and so satellite winds need to be used. In all these cases the ideal scenario would be to have sufficient numbers of wind-monitoring satellites to be able to deliver detailed maps of wind vectors at the surface with a resolution of at least 25 km every 3 hours, so that the time evolution of wind features could be tracked. Such high-frequency sampling in time is not possible at present but even the daily or twice-daily sampling currently achieved is adequate to ensure that the phenomena are not missed altogether. Of course it also makes sense to consult analysis winds as well as satellite winds, to help fill in background trends in the wind field even though fine-scale features are not sharply defined in the analysis field.

There are other larger scale air–sea interaction phenomena where analysis winds are probably adequate to describe the main forcing of wind on the sea, such as the El Niño/La Niña phenomenon in the tropical Pacific, and monsoon-related changes in the circulation patterns of the Indian Ocean (both described in Chapter 11). Even so, in both these cases finer resolved satellite winds are worth consulting to explore the smaller scale variability that may be important for understanding how phenomena evolve as they do.

9.3 TROPICAL CYCLONES OVER THE OCEAN

Taking the perspective of the satellite oceanographer, this section considers how remote sensing of the ocean contributes to the detection and prediction of tropical cyclones, and then explores what can be learned about the impact on the ocean that can be observed following their passage.

9.3.1 Detecting and predicting tropical cyclones

It is beyond the scope of this book to discuss the meteorological science of hurricanes about which much has been written[1] (Elsberry, 1995; Emanuel, 2003; Landsea, 2007), but it is useful to note what is distinctive about these phenomena. *Tropical cyclone* (TC) is a generic term referring to a synoptic-scale, low-pressure weather system that develops without fronts over tropical or subtropical waters. Its primary driving mechanism is a core of rising moist warm air which gains latent heat as it loses moisture through heavy rainfall, promoting strong convection which draws in more moist air from the base, thus driving strong cyclonic circulation characterized by thunderstorms and with very strong surface winds. The fuel for this process is the moisture evaporated by strong surface winds blowing over the sea spirally towards the low-pressure core. The warmer the sea surface the greater the amount of evaporation leading to higher rainfall, stronger convection, and lower pressure at the core. Hence their generic name of tropical cyclones, since they can develop only over warm seas in the tropics and subtropics.

TCs are referred to by different names depending on their severity, as represented by the maximum sustained speed of surface winds. If this is below 17 m/s it is called a *tropical depression*, above 17 m/s it is referred to as a *tropical storm*, and above 33 m/s it is called a *hurricane* if it occurs in the Atlantic or East Pacific oceans, a *typhoon* in the northwest Pacific, a *severe tropical cyclone* in the southwest Pacific, *a severe cyclonic storm* in the North Indian Ocean, or a *tropical cyclone* in the southwest Indian Ocean.

Because of high wind speeds and associated large-amplitude ocean waves, hurricanes are obviously a severe hazard for shipping, and can cause tremendous damage when they make landfall. Once they move inland they are eventually starved of the moisture that fuels them, but not before the very strong winds and heavy rainfall have caused much damage. In addition the combination of wind blowing over the sea and low atmospheric pressure may generate a storm surge that raises the sea level, causing flooding if this overtops or breaches sea defenses. In 2005, Hurricane Katrina which struck New Orleans was accompanied by an 8 m storm surge that flooded the city. Moreover high surface waves on top of the raised mean level cause further damage.

[1] An accessible but authoritative introduction to meteorological knowledge about tropical cyclones, fully supported by technical references, can be found at the "Frequently asked questions" webpages of the Hurricane Research Division of the NOAA Atlantic Oceanographic and Meteorological Laboratory (Landsea, 2007).

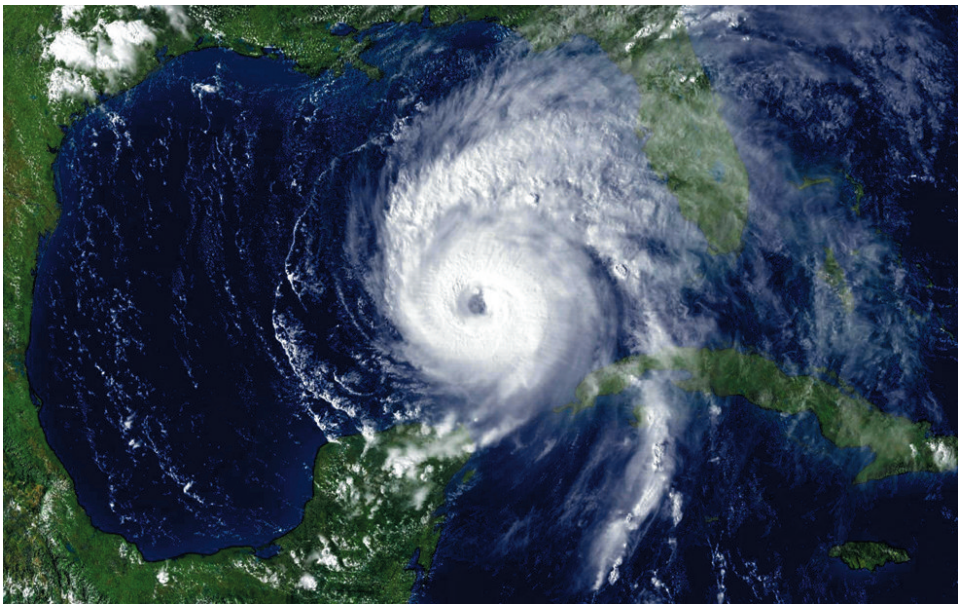


Figure 9.7. Hurricane Ivan over the Gulf of Mexico, as revealed by the NOAA AVHRR visible waveband radiometer on September 14, 2004, showing the characteristic patterns of spiral clouds and a cloud-free eye of a tropical cyclone (image obtained from NOAA website).

The most obvious evidence of hurricanes comes from standard satellite meteorology techniques using visible and infrared geostationary and polar meteorological satellites. Visible radiometry reveals the extent and patterns of clouds while the thermal infrared can detect the high clouds of convective cyclones from cold cloud top temperatures. Figure 9.7 shows a typical example. It is Ivan, a Category 5 hurricane that lasted for 22 days in September 2004. Using such images and microwave sounding techniques to measure the water vapor and rain in the atmospheric column allows tropical storms to be monitored from their first emergence until they die out or grow into a major hurricane. There are several popular hurricane-monitoring websites providing such data, especially during the North Atlantic hurricane season, typically between July and November. However, there is a limit to the amount of useful information available from such data for forecasting the fate of a tropical storm. This is where satellite oceanography methods have a role to play.

Satellite observations can detect and monitor the track of hurricanes when far out at sea by measuring associated sea level winds, detecting the sea state and wave height associated with them, or by monitoring the raised sea level caused by the inverse barometer effect of low atmospheric pressure at the core. The severity of hurricanes is such that sensors on buoys or ships may be damaged or malfunction and vessels would normally try to avoid being in the track of the storm. Satellite sensors are not hindered in this way and microwave sensors that penetrate the intense cloud cover can continue recording properties of the sea's surface. The

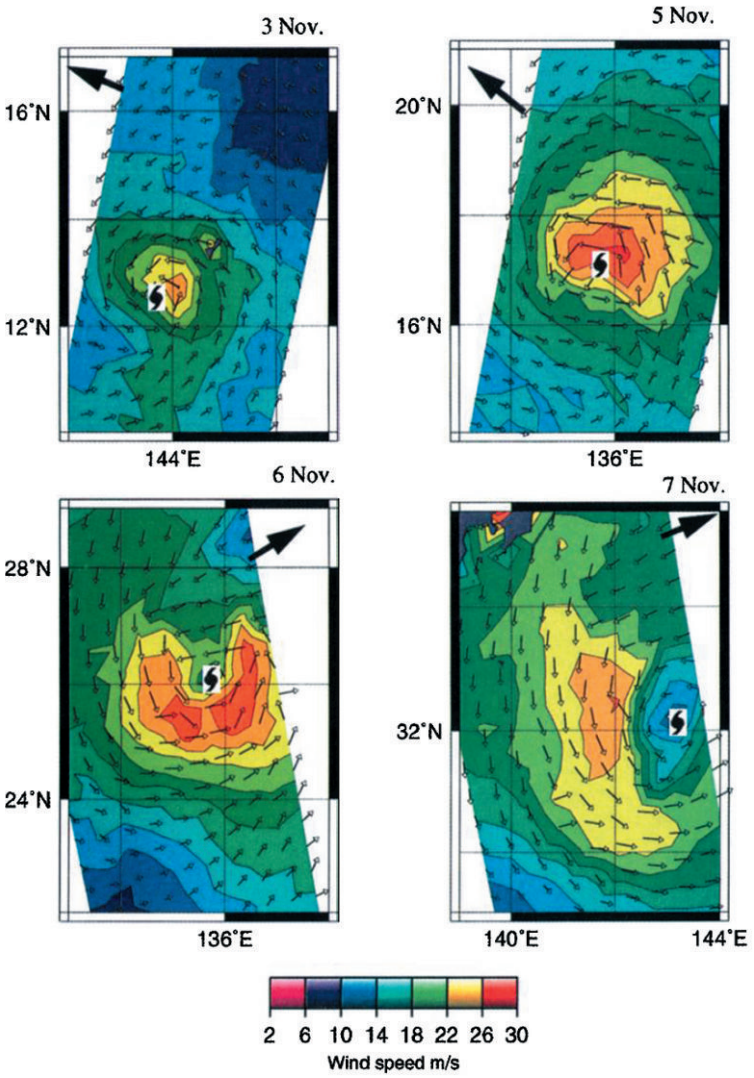


Figure 9.8. Surface wind field from the ERS-l scatterometer at a resolution of 25 km in Tropical Cyclone Elsie: (a) November 3, (b) November 5, (c) November 6, and (d) November 7, 1992. Location of the tropical cyclone center determined from scatterometer measurements and its direction are indicated by the cyclone symbol and the large arrow, respectively (figure copied from plate 2 of Quilfen *et al.*, 1998).

only uncertainty is that sea state and wind speeds in some sectors of the hurricane are typically higher than the range of values over which wind or wave retrieval algorithms have been calibrated and validated. Their extrapolation beyond the validated range therefore attaches greater uncertainly to extreme retrieved values.

Scatterometry was first used for monitoring hurricanes in the 1990s using the ERS-1 scatterometer data. Figure 9.8 shows an example of how Tropical Cyclone Elsie was tracked in the tropical West Pacific during November 1992 (Quilfen *et al.*, 1998). This work showed that processing the scatterometer to retrieve the vector wind field at 25 km resolution (sampled every 12.5 km) allowed the main characteristics of the TC to be defined, such as the location of the center area of low winds and

the outer core of maximum winds. They could also define the wind divergence field. The use of scatterometry at this resolution led to better understanding of the TC mechanism, with improved estimates of wind radius and better storm surge forecasting. It was shown that use of the 50 km resolution standard product from the ERS scatterometer was unable to define the TC well enough to be useful. This early use of the ERS scatterometer also identified the need for better, high-speed wind algorithms which have subsequently been produced (Donnelly *et al.*, 1999; Snoeij *et al.*, 2005; Esteban *et al.*, 2006; Hersbach *et al.*, 2007).

Although the benefits of using scatterometry for monitoring hurricanes was demonstrated by the ERS AMI, they were not used operationally by hurricane-forecasting agencies because of the relatively sparse global coverage of the single swath, and the need to develop backscatter models reliable at high winds. Since QuikScat has been in place, with its wider coverage which normally can observe a given TC at least once per day, increasing reliance has been placed on the operational use of scatterometer wind vectors in hurricane-forecasting centers (Katsaros *et al.*, 2001; Von Ahn *et al.*, 2006). This is reported to have made a very positive impact on operations at the NOAA Ocean Prediction Center, where forecasters appreciate the reliability and timeliness of wind data that are provided to their workstations in a comprehensive form, as well as the wide-swath coverage of QuikScat and the wide range of wind speeds it is capable of detecting. "The ability to distinguish between storm and HF winds has revolutionized the short-term oceanic wind-warning process at the Ocean Prediction Center" (Von Ahn *et al.*, 2006). The provision of 12-hourly updates of detailed wind maps of TCs also allows forecasters to assess how well numerical models are performing and in practice this has given them confidence to place more reliance on model forecasts.

9.3.2 Use of ocean remote sensing to study hurricane–ocean interaction

The capacity to follow a TC's track and observe the detailed evolution of its wind field using a scatterometer has stimulated further research to understand the processes and environmental parameters that control them. For example, there has been growing interest in the impact of sea surface temperature (SST) on both the course and severity of TCs. It is generally accepted that SST needs to be at least 26°C to maintain the TC, although this is by no means the only factor in their initiation, which also requires low vertical wind shear in the troposphere and a large-scale process to produce the vorticity needed to spin up the TC when convection develops.

While TCs respond to SST, the reverse is also true. The cyclonic wind stress field produces a strongly divergent surface ocean current field promoting upwelling of deeper cooler water that leaves a cool wake at the surface along the track of the TC. One of the first applications of ocean remote sensing to hurricane research was to monitor these wakes (Bates and Smith, 1985; Stramma *et al.*, 1986; Cornillon *et al.*, 1987) although it proved difficult using infrared radiometry because of the large amount of cloud surrounding the TC. A considerable time may elapse after the core of the TC passes before a cloud-free view is obtained. During that time the

SST may change, making it difficult to estimate upwelling rates from the SST that is eventually measured. Therefore when the first, reliable, microwave radiometry, SST measurements became available from TMI, this was one of the first phenomena to be explored.

In an early review of the potential of SST measured by the TMI (Wentz *et al.*, 2000) an example is shown of the cold wake left behind by Hurricane Bonnie in late August 1998 as it approached the U.S. Atlantic coast. Three days later, Hurricane Danielle followed along a very similar track until it encountered the 25°C cool wake left by Bonnie, decreased in intensity, and deflected towards the north. It appeared that the new opportunity to measure SST beneath the clouds using microwave radiometry might present a very useful tool to hurricane scientists.

Yet it is too simplistic to suggest that the SST by itself has a dominating control over the track of TCs. As Hurricane Opal crossed the Gulf of Mexico from south to north between August 29 and September 5, 1995, it exhibited sudden unexpected intensification about 24 hours before making landfall. There was a warm core ring sitting in the center of the Gulf at the location where the intensification took place, which might appear to reinforce the connection between SST and the strength of a TC. However, a thorough analysis of the heat exchange and the response of the ocean to the passage of the hurricane focused not so much on the SST as on the heat content in the upper layer of the ocean above the 26°C isotherm (Shay *et al.*, 2000). The study made use of TOPEX altimeter data to map the sea surface height anomaly (SSHA) which domes up over a warm core ring, indicating a greater thickness of warm water near the surface compared with conditions outside the ring. By comparing the SSHA before and after the passage of Opal, they detected a lowering of SSHA by 20 cm, which equates approximately to the mixed layer depth (measured from the 20°C isotherm) reducing by about 50 m. This is a consequence of strong Ekman pumping along the hurricane track which raises the 20°C isotherm. At the same time heat (latent and sensible) is being exchanged from the ocean to the atmosphere through the surface. Normally as a hurricane crosses the ocean the associated upwelling cools the whole mixed layer, fairly soon reducing SST below 26°C which is too low to fuel the hurricane, and thus regulating any intensification. The study concluded that in the absence of the warm core ring the rapid intensification would not have occurred, an interpretation supported by parallel numerical modeling analysis (Hong *et al.*, 2000).

A similar situation appears to have fueled Hurricane Katrina, and was monitored in a similar way using satellite altimetry (Scharroo *et al.*, 2005). They mapped SST from polar-orbiting sensors and used a combination of altimeters to map ocean dynamic topography during the hurricane's passage (as illustrated in Figure 9.9). This shows that hurricane intensification occurred mainly as it crossed regions of high dynamic topography, corresponding to where the mixed layer is deeper. It appears that, although superficially SST was everywhere well above 26°C, the hurricane rapidly exhausted the limited amount of available heat except where there was a thick enough warm surface layer to sustain upwelling without cooling the SST below the temperature at which it could feed Katrina's growth.

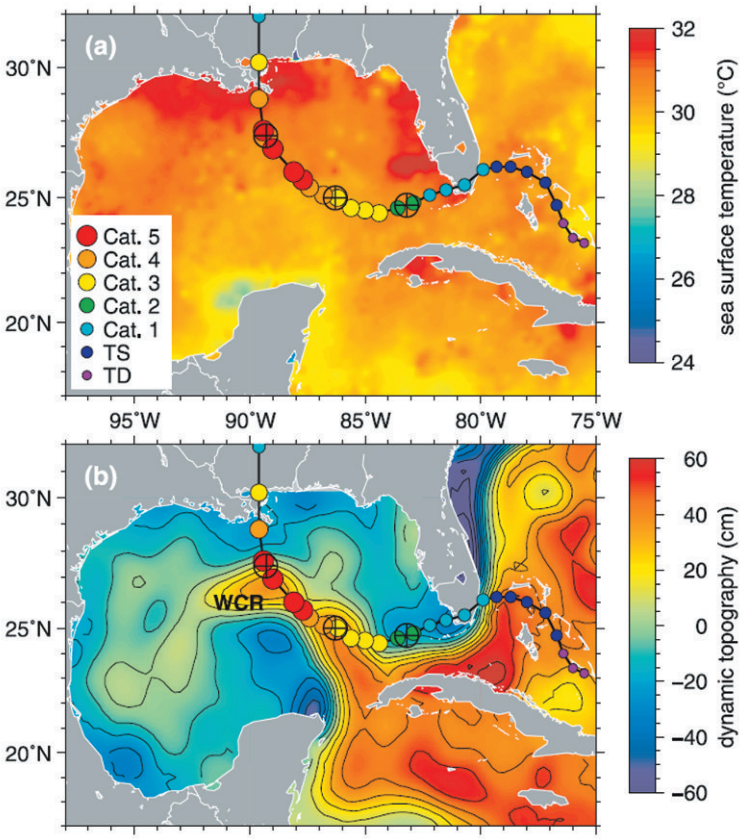


Figure 9.9. The location and intensity of Hurricane Katrina at intervals of 6 hours (circles indicate data from National Hurricane Center advisories) show two intensification events from TS (wind speed <33 m/s) to Cat 1 (33–42 m/s) and Cat 2 (43–49 m/s) to Cat 5 (>70 m/s). (a) Map of sea surface temperature (from POES high-resolution infrared data) shows little correlation with intensification. (b) Map of ocean dynamic topography (from Jason 1, TOPEX, Envisat, and GFO sea surface height data), where highs correlate well with intensifications. The Loop Current can be seen entering the Gulf south of Cuba and exiting south of Florida; the warm-core ring (WCR) is the prominent high shedding from the Loop Current in the center of the Gulf (from figure 2 of Scharroo *et al.*, 2005).

This suggests that, while SST >26°C is a necessary condition for sustaining TCs, SST is not the most useful indicator of which geographical regions present oceanographically favorable conditions for hurricane growth. That depends more on the thermal structure and thickness of the surface layer of water above the 26°C isotherm. Such information can be deduced from a combination of *in situ* profilers and satellite altimetry, as demonstrated by Pun *et al.* (2007) in the western North Pacific. Their study showed that intensity changes of Typhoon Dianmu in 2004 depended on the relative magnitude of tropical cyclone heat potential (TCHP) as determined by

the upper-ocean thermal structure, and the typhoon's self-induced cooling of the upper ocean layer by upwelling.

This appears to be a fertile area of ongoing research where satellite ocean data have an important part to play in providing information about background conditions in the ocean before and after a TC passes over it. There remain some apparently conflicting views about the importance of SST by itself as a driver for TC intensification (e.g., Kafatos *et al.*, 2006). On closer inspection, discrepancies between the conclusions of different studies seem to be a consequence of their different perspectives (e.g., the testing of different hypotheses), rather than contradictions between observational evidence. Ultimately it is important that we do understand how the incidence and severity of TCs is related to the ocean state. As the ocean changes in response to global warming we need to be able to predict with confidence based on scientific evidence whether or not this will inevitably result in more or stronger TCs (Trenberth, 2005). It appears that the role of ocean remote sensing will be more complex than simply monitoring SST with greater accuracy. The challenge is to draw on a combination of satellite data, *in situ* observations, and numerical models to characterize the thermal structure of the upper ocean.

Finally another important oceanographic consequence of the passage of a hurricane must not be overlooked. This is the impact on biology. Davis and Yan (2004) have explored this using SeaWiFS image data, comparing chlorophyll concentrations before and after the passage of a hurricane, since during the hurricane no cloud-free views are available. They found clear evidence of enhancement of chlorophyll concentration by the passage of the hurricane and concluded that this was caused by strong upwelling and vertical mixing raising nutrients into the photic zone. The study was performed off the northeast coast of the U.S.A. between Cape Hatteras and Cape Cod where hurricanes occur mainly in late August, September, and October. This is toward the end of summer when phytoplankton have exploited all the nutrients in the upper ocean layer, so the passage of the hurricane injects a fresh supply while there is still enough light for this to be exploited by phytoplankton. Effectively the passage of the hurricane triggers an early start to the fall bloom. In this region, hurricanes also "generated" long filaments of enhanced primary production, in which nutrients appear to have been injected into the north wall of the Gulf Stream and the blooming population is transported northeastwards into the Atlantic. A similar study in the oligotrophic center of the Atlantic Subtropical Gyre found very similar results (Babin *et al.*, 2004). Here the chlorophyll concentration returned approximately to pre-hurricane conditions within about 2 weeks of the event.

9.4 SATELLITE WINDS FOR OFFSHORE WIND FARMS

The final section of this chapter looks briefly at the way satellite-measured winds have a role to play in mapping the fine-scale wind field in relation to operational activities in coastal waters. For very fine–resolution winds, SAR is the only sensor available. As discussed in section 10.8 of *MTOFS*, a number of interesting, small-

scale, coastal, atmospheric phenomena have been revealed in SAR images, and examples of wind shadows behind islands and headlands, katabatic wind outbreaks over the sea, and land breeze fronts are all illustrated in that chapter. How these and other atmospheric features come to have SAR signatures is discussed in more detail by Alpers (1995) and Alpers *et al.* (1999).

Because there are so few SARs in operation their high-resolution views of the ocean are infrequent, with revisit intervals of 10 to 15 days. Therefore they cannot be put to operational use for local forecasting, although the acquisition of SAR images of an area over several years has created an archive of information about typical patterns of wind distribution in complex coastal regions. This can be exploited for locating coastal and offshore engineering and new port installations. Zones that regularly experience higher winds than the surrounding region can be avoided. On the other hand those are the very places that may provide the optimum locations for offshore wind power installations. This application of satellite wind measurement has received quite a lot of interest in recent years as nations attempt to increase the proportion of power generated from renewable energy sources.

Mapping global wind power potential

In broad terms, the identification of regions around the world where wind power is concentrated, and thus potentially suitable for efficient exploitation by offshore wind energy farms, is a task for coarser resolution, satellite wind sensors, such as the scatterometer. Wind speed climatologies are available from satellite measurements (Risien and Chelton, 2006) and show where the strongest winds are found in different seasons. It may, however, be misleading to consider only mean wind speed because at a given location the strongest wind events contribute disproportionately to total wind power available. The instantaneous energy flux density of wind passing through a plane normal to the wind direction is $\frac{1}{2}\rho_a U^3$ per unit area of the plane—where ρ_a is atmospheric density; and U is wind speed. A simple way of understanding this is that wind transports its own kinetic energy, $\frac{1}{2}\rho_a U^2$, through a distance U per unit time. This power per unit area, or power density, is apparently available for exploitation, but in reality typical outputs of practical wind turbines are between 25% and 35%.[2]

Given the nonlinear relation between available power and wind speed, it is important to know the climatology of high winds (Sampe and Xie, 2007). This requires knowledge of the probability distribution function (PDF) of wind speed at each location as well as its mean (Liu *et al.*, 2008). In principle, assessment of the potential power output at an offshore wind farm site involves an integration over time of the instantaneous wind power density, but Liu *et al.* (2008) use knowledge of

[2] There is a theoretical maximum energy extraction factor of 16/27 (59.3%) (the Betz limit), best understood by noting that if all the wind's kinetic energy were extracted there would be no flow left to transport that energy into the device. The most efficient turbine can convert no more than 70% of what is theoretically extractable, so achieving 40% conversion of wind energy flux is an ideal design target. Practical operating issues, such as ensuring that storms do not damage devices, typically reduce this to less than 30% depending on the operating regime.

the PDF to cope with the nonlinear dependence of instantaneous power on wind speed. Their approach uses a parametric expression for E, mean available wind power density. This is

$$E = 0.5\rho_a c^3 \Gamma(1 + 3/k), \tag{9.1}$$

where $c = \bar{U}/\Gamma(1 + 1/k)$ is a scale parameter; $\Gamma(\ )$ is the gamma function; $\bar{U}$ is the mean and σ is the standard deviation of the population of U samples; and $k = (\bar{U}/\sigma)^{1.086}$ is the dimensionless shape parameter for the Weibull distribution that is assumed to represent the PDF $p(U)$. This has the form

$$p(U) = (k/c)(U/c)^{k-1} \exp[-(U/c)^k]. \tag{9.2}$$

Thus, instead of having to integrate all individual values of instantaneous power density from the time series of U, in order to evaluate mean power density over a given period of time (say, a month), it can be calculated from (9.1) using just the mean and standard deviation of U over that period. This is readily done at all locations using a satellite-derived, gridded, wind climatology. The result is maps like Figure 9.10 which show mean power density distribution over the ocean for (a) the three boreal winter months and (b) the three summer months, based on 8 years of QuikScat data. The broadly expected pattern emerges in which regions of large available power change hemisphere with the season, but there are a number of places where wind power is moderate to high in both local summer and winter. These could make suitable locations for efficient wind farming using offshore, floating, wind turbines. The study by Liu *et al.* (2008) highlights particular locations such as the Oregon coast, the Caribbean, and the Japanese coast where there is high annual mean wind power.

Up till now offshore wind farms have been placed in shallow water close to the coast. Maps like Figure 9.10 are still useful pointers to suitable coastal locations, but for near-shore installations it is important to know about the wind distribution at a much finer spatial scale in order to identify sites where islands or headlands create localized wind shadows. These are generally to be avoided if wind power potential is to be most effectively exploited, although an ideal site might be one which is open to dominant, prevailing winds but offers some shelter from extreme storm winds. This is where SAR-derived wind maps are useful (Johannessen and Bjorgo, 2000; Hasager *et al.*, 2002). Because there are very limited numbers of SAR images and even fewer have been processed to retrieve wind vectors, it is not easy to derive reliable mean winds or vector wind PDFs from SAR (Barthelmie and Pryor, 2003; Pryor *et al.*, 2004) although this approach has been used in some cases.

As an alternative a procedure has been developed (Furevik and Espedal, 2002; Furevik *et al.*, 2003) in which the wind directional climatology for the region (derived from local wind sensors located clear of any wind shadows, from offshore scatterometer data or from model forecasts) is consulted to identify the dominant patterns of wind most frequently encountered in the region. SAR scenes that are representative of these typical wind directions are then analyzed to reveal the detailed wind distribution and to identify where wind shadows occur in relation to particular, prevailing wind directions. This is then factored into estimation of avail-

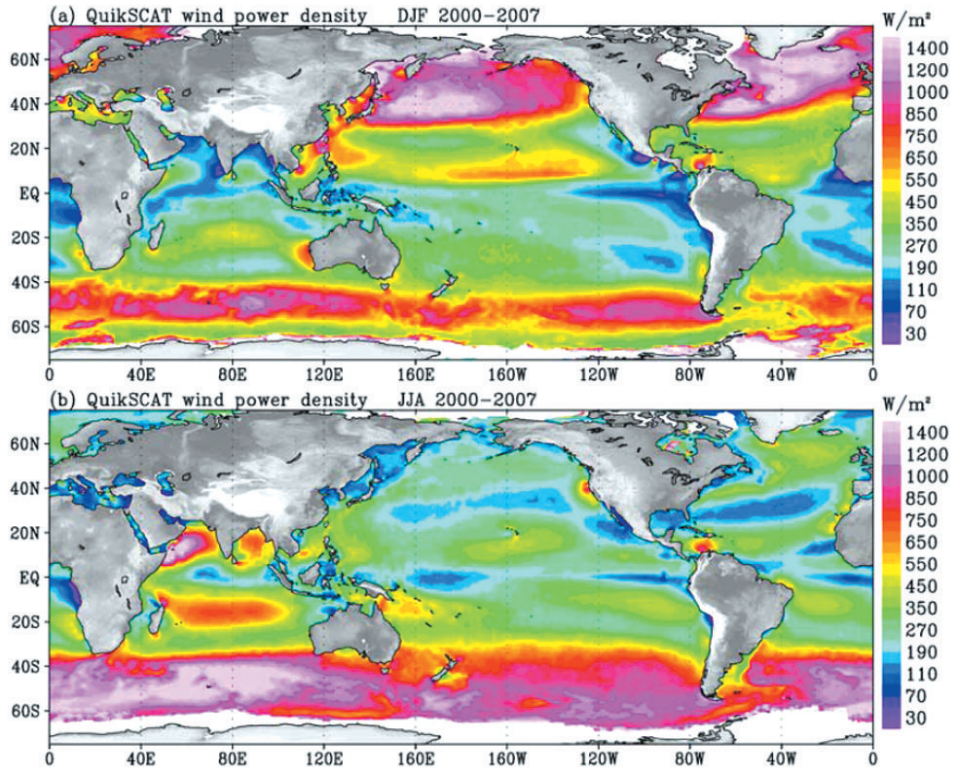


Figure 9.10. Distributions of wind power density derived from QuikSCAT for (a) boreal winter (December, January, and February) and (b) boreal summer (June, July, and August), for an 8-year period between 2000 and 2007. The gray scale is used to show land topography (figure from Liu *et al.*, 2008).

able power for a variety of possible sites for wind farms within the region. The method was demonstrated at locations off the Norwegian and Danish North Sea coasts, and as part of the exercise the SAR-retrieved wind fields were validated against *in situ* wind measurements.

Another use for SAR wind fields in support of offshore wind farming is to detect the extent and the length of wind shadows caused by existing wind farms (Christiansen and Hasager, 2005, 2006). Information about the wind shadows of turbines first demonstrates how effective they are at removing wind energy, and also shows how far downwind the shadow reaches before it would be effective to harness further wind energy. Figure 9.11 shows an example of the shadow or wake of a wind farm installation at Horns Rev off the Danish coast (within the white trapezoid marked on the images). The wake stretches more than 20 km downstream but does not spread. It is also interesting to note much stronger wakes farther north that are the wind shadows of 30 m high sand dunes along the coast. This illustrates the type of naturally occurring wind shadows that must be avoided

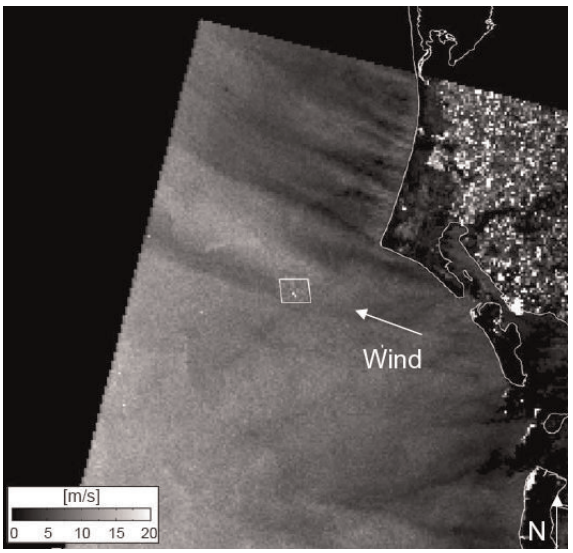


Figure 9.11. Wind speed map derived from an ERS-2 SAR scene acquired on February 25, 2003. The wind farm at Horns Rev is indicated by the white trapezoid. Wind wake is seen as dark pixels downstream of the turbines. Wind direction = 110°, wind speed = 6.0 m/s (as measured at the mast beside the installation).

when deciding the location of a wind farm, and which can clearly be revealed in a SAR image.

It is to be expected that further engineering applications like this will emerge as confidence grows in the validity of wind fields that are retrieved from SAR, and as engineers come to depend on the information to be extracted and interpreted from an image such as Figure 9.11.

9.5 REFERENCES

Alpers, W. (1995), Measurement of oceanic and atmospheric phenomena by ERS-1 SAR. *Radio Sci. Bull.*, **275**, 14–22.

Alpers, W., L. Mitnik, L. Hock, and K. S. Chen (1999), *The Tropical and Subtropical Ocean Viewed by ERS SAR*, available at *http://www.ifm.uni-hamburg.de/ers-sar/* (last accessed April 25, 2008).

Babin, S. M., J. A. Carton, T. D. Dickey, and J. D. Wiggert (2004), Satellite evidence of hurricane-induced phytoplankton blooms in an oceanic desert. *J. Geophys. Res.*, **109**(C03043).

Barthelmie, R. J., and S. C. Pryor (2003), Can satellite sampling of offshore wind speeds realistically represent wind speed distributions? *J. Applied Meteorology*, **42**, 83–94.

Bates, J. J., and W. L. Smith (1985), Sea surface temperature: Observations from geostationary satellites. *J. Geophys. Res.*, **90**, 11609–11618.

Brown, G. S., H. R. Stanley, and N. A. Roy (1981), The wind speed measurement capability of spaceborne radar altimetry. *IEEE J. Oceanic Eng.*, **6**, 59–63.

Brown, S. T., C. S. Ruf, and D. R. Lyzenga (2006), An emissivity-based wind vector retrieval algorithm for the WindSat polarimetric radiometer. *IEEE Trans. Geosc. Remote Sensing.*, **44**(3), 611–621.

Chelton, D. B., and F. J. Wentz (1986), Further development of an improved altimeter wind speed algorithm. *J. Geophys. Res.*, **91**, 14250–14260.

Christiansen, M. B., and C. B. Hasager (2005), Wake effects of large offshore wind farms identified from satellite SAR. *Remote Sens. Environ.*, **98**, 251–268.

Christiansen, M. B., and C. B. Hasager (2006), Using airborne and satellite SAR for wake mapping offshore. *Wind Energy*, **9**, 437–455.

Cornillon, P., L. Stramma, and J. F. Price (1987), Satellite measurements of sea surface cooling during hurricane Gloria. *Nature*, **326**, 373–375.

Davis, A., and X.-H. Yan (2004), Hurricane forcing on chlorophyll-a concentration off the northeast coast of the U.S. *Geophys. Res. Letters*, **31**(L17304), doi: 10.1029/2004GL020668.

Donnelly, W. J., J. R. Carswell, R. E. McIntosh, P. S. Chang, J. Wilkerson, F. Marks, and P. G. Black (1999), Revised ocean backscatter models at C and Ku-band under high-wind conditions. *J. Geophys. Res.*, **104**(C5), 11485–11497.

Ebuchi, N., H. C. Graber, and M. J. Caruso (2002), Evaluation of wind vectors observed by QuikSCAT/SeaWinds using ocean buoy data. *J. Atm. Ocean. Tech.*, **19**, 2049–2062.

Elsberry, R. L. (Ed.) (1995), *Global Perspectives on Tropical Cyclones* (Tech. Doc. WMO/TD No. 693). World Meteorological Organization, Geneva, Switzerland.

Emanuel, K. (2003), Tropical cyclones. *Ann. Rev. Earth Planet. Sci.*, **31**, 75–104.

Esteban, F. D., J. R. Carswell, S. Frasier, P. S. Chang, P. G. Black, and F. D. Marks (2006), Dual-polarized C- and Ku-band ocean backscatter response to hurricane-force winds. *J. Geophys. Res.*, **111**(C08013), doi: 0.1029/2005JC003048.

Fichaux, N., and T. Ranchin (2002), Combined extraction of high spatial resolution wind speed and wind direction from SAR images: A new approach using wavelet transform. *Can. J. Remote Sensing*, **28**(3), 510–516.

Figa-Saldaña, J., J. J. W. Wilson, E. Attema, R. Gelsthorpe, M. R. Drinkwater, and A. Stoffelen (2002). The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers. *Can. J. Remote Sensing*, **28**(3), 404–412.

Freilich, M. H., and P. G. Challenor (1994), A new approach for determining fully empirical altimeter wind speed model functions. *J. Geophys. Res.*, **99**, 25051–25062.

Freilich, M., and B. A. Vanhoff (2006), The accuracy of preliminary WindSat vector wind measurements: Comparisons with NDBC buoys and QuikSCAT. *IEEE Trans. Geosc. Remote Sensing.*, **44**(3), 622–637.

Furevik, B. R., and H. Espedal (2002), Wind energy mapping using synthetic aperture radar. *Can. J. Remote Sensing*, **28**(2), 196–204.

Furevik, B. R., H. A. Espedal, T. Hamre, C. B. Hasager, O. M. Johannessen, B. H. Jørgensen, and O. Rathmann (2003), Satellite-based wind maps as guidance for siting offshore wind farms. *Wind Engineering*, **27**(5), 327–338.

Gelsthorpe, R. V., E. Schied, and J. J. W. Wilson (2000), ASCAT: Metop's advanced scatterometer. *ESA Bulletin*, **102**, 19–27.

Gommenginger, C. P., M. A. Srokosz, P. G. Challenor, and P. D. Cotton (2002), Development and validation of altimeter wind speed algorithms using an extended collocated buoy/Topex dataset. *IEEE Trans. Geosc. Remote Sensing.*, **40**(2), 251–260.

Gourrion, J., D. Vandemark, S. Bailey, B. Chapron, C. P. Gommenginger, P. G. Challenor, and M. A. Srokosz (2002), A two-parameter wind speed algorithm for Ku-band altimeters. *J. Atm. Ocean. Tech.*, **19**, 2030–2048.

Hasager, C. B., H. P. Frank, and B. R. Furevik (2002), On offshore wind energy mapping using satellite SAR. *Can. J. Remote Sensing*, **28**, 80–89.

Hersbach, H., A. Stoffelen, and S. de Haan (2007), An improved C-band scatterometer ocean geophysical model function: CMOD5. *J. Geophys. Res.*, **112**(C03006), doi: 10.1029/2006JC003743.

Hong, X., S. W. Chang, S. Raman, L. K. Shay, and R. Hodur (2000), The interaction between Hurricane Opal (1995) and a warm core eddy in the Gulf of Mexico. *Mon. Weather Rev.*, **128**, 1347–1365.

Johannessen, O. M., and E. Bjorgo (2000), Wind energy mapping of coastal zones by synthetic aperture radar (SAR) for siting potential windmill locations. *Int. J. Remote Sensing*, **21**(9), 1781–1786.

Kafatos, M., D. Sun, R. Gautam, Z. Boybeyi, R. Yang, and G. Cervone (2006), Role of anomalous warm gulf waters in the intensification of Hurricane Katrina. *Geophys. Res. Letters*, **33**(L17802), doi: 10.1029/2006GL026623.

Katsaros, K. B., E. B. Forde, P. Chang, and W. T. Liu (2001), Quik-SCAT"s SeaWinds facilitates early identification of tropical depressions in 1999 hurricane season. *Geophys. Res. Letters*, **28**, 1043–1046.

Kerbaol, V., B. Chapron, and P. W. Vachon (1998), Analysis of ERS-1/2 synthetic aperture radar wave mode imagettes. *J. Geophys. Res.*, **103**, 7833–7846.

Kidder, S. Q., and T. H. Vonder Haar (1995), *Satellite Meteorology: An Introduction* (466 pp.). Academic Press, San Diego, CA.

Korsbakken, E., J. A. Johannessen, and O. M. Johannessen (1998), Coastal wind field retrievals from ERS synthetic aperture radar images. *J. Geophys. Res.*, **103**, 7857–7874.

Landsea, C. (2007), *Frequently Asked Questions, Version 4.2: June 1, 2007*. Hurricane Research Division, National Oceanic and Atmospheric Administration, Silver Springs, MD, available at *http://www.aoml.noaa.gov/hrd/tcfaq/tcfaqHED.html* (last accessed August 18, 2008).

Lehner, S., J. Horstmann, W. Koch, and W. Rosenthal (1998), Mesoscale wind measurements using recalibrated ERS SAR images. *J. Geophys. Res.*, **103**, 7847–7856.

Liu, W. T. (2002), Progress in scatterometer application. *J. Oceanogr.*, **58**, 121–136.

Liu, W. T., and X. Xie (2006), Measuring ocean surface wind from space. In: J. F. R. Gower (Ed.), *Remote Sensing of the Marine Environment: Manual Remote Sensing* (Vol. 6, Third Edition, pp. 149–178). American Society for Photogrammetry and Remote Sensing, Bethesda, MD.

Liu, W. T., W. Tang, and X. Xie (2008), Wind power distribution over the ocean. *Geophys. Res. Letters*, **35**(L13808), doi: 10.1029/2008GL034172.

Pryor, S. C., M. Nielsen, R. J. Barthelmie, and J. Mann (2004), Can satellite sampling of offshore wind speeds realistically represent wind speed distributions? Part II: Quantifying uncertainties associated with distribution fitting methods. *J. Applied Meteorology*, **43**, 739–750.

Pun, I.-F., I.-I. Lin, C.-R. Wu, D.-S. Ko, and W. T. Liu (2007), Validation and application of altimetry-derived upper ocean thermal structure in the western North Pacific Ocean for typhoon-intensity forecast. *IEEE Trans. Geosc. Remote Sensing*, **45**(6), 1616–1630.

Quilfen, Y., B. Chapron, T. Elfouhaily, K. Katsaros, and J. Tournadre (1998), Observation of tropical cyclones by high-resolution scatterometry. *J. Geophys. Res.*, **103**(C4), 7767–7786.

Risien, C. M., and D. B. Chelton (2006), A satellite-derived climatology of global ocean winds. *Remote Sens. Environ.*, **105**, 221–236.

Robinson, I. S. (2004) *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Sampe, T., and S. P. Xie (2007), Mapping high sea winds from space: A global climatology. *Bull. Am. Meteorol. Soc.*, **88**, 1965–1978.

Scharroo, R., W. H. F. Smith, and J. L. Lillibridge (2005), Satellite altimetry and the intensification of hurricane Katrina. *EOS, Trans. Amer. Geophys. Union*, **86**(40), 366–367.

Shay, L. K. G., J. Goni, and P. G. Black (2000), Effects of a warm oceanic feature on Hurricane Opal. *Mon. Weather Rev.*, **128**, 1366–1383.

Snoeij, P., E. Attema, H. Hersbach, A. Stoffelen, R. Crapolicchio, and P. Lecomte (2005), Uniqueness of the ERS scatterometer for nowcasting and typhoon forecasting. Paper presented at *Proc. Geoscience and Remote Sensing Symposium: IGARSS '05* (pp. 4792–4795). Institute of Electrical and Electronic Engineers, Piscataway, NJ.

Stoffelen, A. C. M., and D. L. T. Anderson (1997), Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4. *J. Geophys. Res.*, **102**(C3), 5767–5780.

Stramma, L., P. Cornillon, and J. F. Price (1986), Satellite observations of sea surface cooling by hurricanes. *J. Geophys. Res.*, **91**(C4), 5031–5035.

Trenberth, K. (2005), Uncertainty in hurricanes and global warming. *Science*, **308**(5729), 1753–1754.

Vachon, P. W., and F. W. Dobson (1996), Validation of wind vector retrieval from ERS-1 SAR images over the ocean. *Glob. Atmos. Ocean Syst.*, **5**, 177–187.

Von Ahn, J. M., J. M. Sienkiewicz, and P. S. Chang (2006), Operational impact of QuikSCAT winds at the NOAA Ocean Prediction Center. *Weather and Forecasting*, **21**, 523–539.

Wentz, F. J. (1997), A well calibrated ocean algorithm for SSM/I. *J. Geophys. Res.*, **102**, 8703–8718.

Wentz, F. J., C. Gentemann, D. Smith, and D. Chelton (2000), Satellite measurements of sea surface temperature through clouds. *Science*, **288**(5467), 847–850.

Yelland, M. J., B. I. Moat, R. W. Pascal, and D. I. Berry (2002), CFD model estimates of the airflow distortion over research ships and the impact on momentum flux measurements. *J. Atm. Ocean. Tech.*, **19**(10), 1477–1499.

10

Fluxes through the air–sea interface

Co-authored with Susanne Fangohr[1]

10.1 INTRODUCTION

Considering the World Ocean as a continuous body of water, oceanographers are mainly concerned with the internal processes that control the distribution of its physical and chemical properties. However, there are two regions in which the conditions determining the characteristics and behavior of this water mass are distinctly different from the rest. These are its boundaries with the solid earth below and with the atmosphere above. It is here that liquid ocean waters meet either solid or gas, giving rise to a range of processes not encountered in other parts of the ocean. This chapter is concerned with the interface between the ocean and the atmosphere, the ways in which fluxes from one medium to the other can be measured, and how satellite-derived ocean data can be used with the goal of improving the geographical reach and accuracy of air–sea flux estimates.

From a remote-sensing perspective, the air–sea interface is the part of the ocean most accessible to sensors in space. For electromagnetic waves in those parts of the spectrum that can pass through the atmosphere with little attenuation, the sea surface is the principal encounter point which determines what a satellite remote-sensing instrument observes. While for many oceanographic applications it might be preferable to look through this barrier to see into the deep ocean, for those wishing to study processes centered around the air–sea interface, remote sensing appears to be an ideal tool. Using satellites we can measure or deduce a number of the ocean parameters that influence air–sea fluxes. Unfortunately it is not so straightforward for atmospheric remote-sensing methods to measure air properties close to the sea surface at the bottom of the atmospheric boundary layer (ABL). This creates serious challenges for estimating fluxes on the basis of satellite data alone.

[1] Dr. Susanne Fangohr is a research fellow in the School of Ocean and Earth Sciences at the National Oceanography Centre, Southampton, U.K.

Now that global warming and climate change are recognized as issues of considerable public interest, there has been a substantial increase in the perceived importance of research to understand the interactions and exchanges between the atmosphere and the ocean, mediated by fluxes across the air–sea interface. The study of only one side of the interface—ocean or atmosphere—cannot offer a complete answer to questions relating to development of the Earth's climate. Coupled ocean–atmosphere models which aim to provide an accurate description and prediction of these developments depend on parameterizations of processes at the air–sea boundary and of fluxes between the two media. Gases such as oxygen (O_2) and carbon dioxide (CO_2) as well as heat, momentum, and humidity (water vapor) are quantities that are exchanged between the ocean and the atmosphere at all times, driven by gradients across the interface. Characteristic properties of the interface itself play an important role in determining the magnitude of this exchange and thus they affect the speed of equilibration following a change of gas concentration or physical quantity in the ocean, the atmosphere, or both.

Characterizing a moving interface between two media at small spatial scales representative of the governing molecular processes is a difficult problem on its own. Add to this the vast horizontal extent of the sea surface, its inaccessibility, and the difficulties of measuring from a ship or buoy, which itself is a moving platform, and it becomes obvious why studying the air–sea interaction poses a complex challenge to classical oceanography. In comparison, the advantages of remote sensing in terms of spatial and temporal coverage and ease of access have a lot to offer. Indeed it is difficult to envisage a system for monitoring air–sea fluxes with global coverage, spatial resolution at the oceanic mesoscale, and time-sampling every few days which does not make extensive use of satellite remote sensing. However, since the air–sea flux of the properties of interest cannot be measured directly from space, the ongoing scientific task for 20 years has been to develop suitable parameterizations that link those quantities that are measurable by satellite sensors to the air–sea fluxes we wish to estimate.

Outlining the progress made in this task forms the content of this chapter, which discusses the air–sea fluxes of heat and gases and how they can be measured from space. Although the flux of momentum is mentioned—because it provides a basis for understanding other fluxes—it is not developed here. In fact the impact of momentum exchange crops up in many other parts of this book, wherever wind stress is implicated in driving an oceanic process such as upwelling (Chapter 5), surface waves (Chapter 8), and wherever wind mixing of the upper ocean is mentioned. The next section explains some of the basic principles underlying parameterization of air–sea exchange processes, and is followed by a review of the remote-sensing methods used to determine the important parameters required for flux estimation. Section 10.4 outlines the current state of the art concerning the global mapping of gas and heat fluxes using satellite data. The final section reflects on what more needs to be done before the benefits of satellite data are fully exploited in global systems for routine monitoring of air–sea fluxes that can supply valuable knowledge about short-term climate change.

10.2 DETERMINING FLUXES

10.2.1 General principles

It is a characteristic of the Earth's hydrosphere that flows of air and water are naturally turbulent except at very short lengthscales. Consequently turbulent mixing transfers heat, mass and momentum, eroding gradients of temperature, concentration of dissolved constituents, and velocity throughout the oceans and atmosphere on timescales which are several orders of magnitude faster than those of molecular diffusion. However, at the air–sea interface, turbulent mixing occurs independently on either side of the interface but cannot penetrate it. The size of eddies performing the mixing action decreases with increasing proximity to the interface. Directly at the interface, two thin viscous sublayers exist through which transport can occur only by molecular processes. This can be conceptually modeled as illustrated in Figure 10.1, which shows the two "bulk" domains (ocean and atmosphere) governed by turbulent mixing and within which properties are almost uniform over lengthscales of centimeters or more. The bulk regions are separated by two thin molecular sublayers immediately on either side of the interface, less than 1 mm thick but across which the property can change considerably. While usually the difference in concentration between sea and air of any parameter will be measured in the well-mixed bulk of the atmosphere and the ocean, fluxes are often limited by the rate at which they cross molecular sublayers from one medium into the other. This means that processes on lengthscales of molecular diffusion which are difficult to measure directly can determine how much mass, heat, or momentum is exchanged between the ocean and the atmosphere.

From the point of view of a remote-sensing scientist, there are several surface measurements such as temperature, wind, sea surface roughness, and wave height that can be made directly from space and which provide information on the air–sea interface that can be used in the estimation of fluxes between the ocean and atmosphere. To be able to use these observations for calculating fluxes requires the physical concepts of flux processes to be expressed in equations whose measurable properties are variable parameters (as developed in the next subsection).

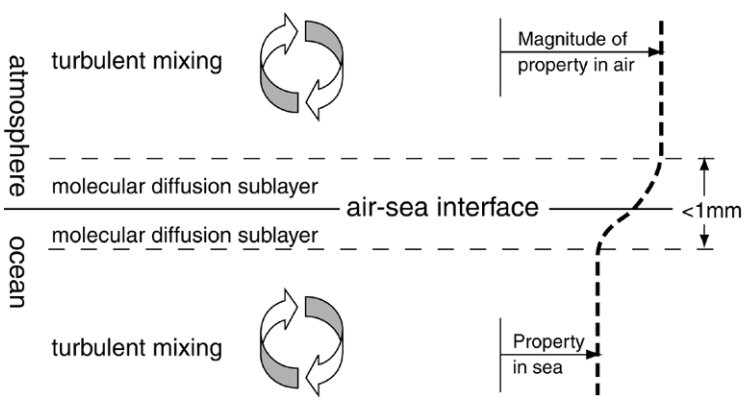


Figure 10.1. Two-layer model of air–sea interaction showing a layer dominated by turbulent mixing and one governed by molecular diffusion on either side of the air–sea interface.

10.2.2 Theoretical basis of flux parameterizations

The most commonly used parameterizations of air–sea fluxes of gases, and of turbulent heat and momentum fluxes, all follow a common scheme. Generically the flux F is calculated from bulk formulas as the product of a transfer coefficient K_x, another parameter R describing the ease of transfer across the interface which depends on characteristics of the air and sea directly adjacent to the interface, and a gradient across the air–sea interface of a quantity related to the flux, ΔX, which drives and determines the direction of the flux:

$$F = K_x \cdot R \cdot \Delta X. \tag{10.1}$$

In the case of gas transfer, substituting the relevant parameters for the air–sea flux of sparingly soluble gases such as oxygen and carbon dioxide, Equation (10.1) becomes

$$F_{\text{gas}} = s \cdot k \cdot (pX_w - pX_a), \tag{10.2}$$

where F_{gas} is gas flux from ocean to atmosphere; s is the solubility of the gas in seawater at temperature T_s and salinity S (e.g., Weiss, 1974; Wanninkhof, 1992); k is the transfer velocity across the interface given in units of centimeters per hour; and pX_w and pX_a are the partial pressures of gas on the sea and air side of the interface, respectively. Nightingale and Liss (2004) give a more complete overview of the derivation of this equation.

For the flux of momentum from atmosphere to ocean, expressed by the wind stress τ, Equation (10.1) becomes

$$\tau = \rho \cdot C_D \cdot (u_z - u_s)^2, \tag{10.3}$$

where ρ is the density of air at temperature T_a and pressure p_z at height z; C_D is the drag coefficient (Large and Pond, 1981); u_z is horizontal wind speed at height z (normally standardized to 10 m); and u_s is horizontal wind speed at the sea surface, often approximated to be zero.

Net heat exchange Q at the sea surface can be divided into four main components (Liu *et al.*, 1979):

$$Q = Q_S + Q_b + Q_H + Q_E, \tag{10.4}$$

where Q_S is net shortwave radiation (incoming from the Sun); Q_b is net longwave radiation; Q_H and Q_E are sensible and latent heat fluxes, respectively. Determination of the radiative part of net flux (i.e., $Q_S + Q_b$) requires a different approach from that applied to turbulent exchange across the air–sea interface. It is described separately in Section 10.4.1, but its relative importance to the overall heat budget is considered alongside Q_H and Q_E when global heat flux observations are discussed in Section 10.4.3. In this section, we will concentrate on turbulent heat fluxes (i.e., $Q_H + Q_E$).

Latent heat flux is one of the dominant components in the exchange of energy between the atmosphere and the ocean. It broadly balances energy from shortwave solar flux. Latent heat flux occurs when thermal energy is drawn from the sea to

cause a phase change of water at the sea surface into water vapor, which is then freed into the atmosphere, transferring both heat and water from the ocean to the atmosphere. Later condensation of the water vapor releases latent heat in the atmosphere and provides energy to feed atmospheric circulation, especially in tropical regions (Jourdan and Gautier, 1995). In the case of latent heat flux from ocean to atmosphere, Q_E, the generic flux equation (10.1) is expressed as

$$Q_E = \rho \cdot L \cdot C_E \cdot (u_z - u_s) \cdot (q_s - q_a), \tag{10.5}$$

where ρ is the density of air at temperature T_a and pressure p_z at height z; L is the latent heat of vaporization of water at air temperature T_a; C_E is the transfer coefficient for latent heat (also referred to as the Dalton number); u_z is wind speed at height z; and q_s and q_a are the specific humidity at the sea surface (often assumed to equal the saturation humidity at T_s and p_s) and at height z.

For the flux of sensible heat from ocean to atmosphere, Q_H, Equation (10.1) becomes

$$Q_H = \rho \cdot c_p \cdot C_T \cdot (u_z - u_s) \cdot (T_s - T_a), \tag{10.6}$$

where ρ is the density of air at temperature T_a and pressure p_z at height z; and c_p is the specific heat at constant pressure; C_T is the exchange coefficient for sensible heat, also referred to as the Stanton number.

These equations provide a practical method for estimating air–sea fluxes from average measurements of temperature, wind, and water vapor density assuming that the value of the exchange coefficients are known. Where an atmospheric property is required, such as gas pressure X_a (in 10.2), wind speed u_z (in 10.3), specific humidity q_a (in 10.5), and air temperature T_a (in 10.6) the measurement is expected to be normalized to a standard height, normally 10 m, above the sea surface but still within the lower part of the ABL. This follows the standard practice of *in situ* observational experiments in boundary layer meteorology.

The following section considers the extent to which remote sensing can provide the required input data. Section 10.4 then discusses how these conceptual equations are implemented in practical flux retrieval systems, and notes the uncertainties that remain with regard to the exchange coefficients for heat and gas fluxes.

10.3 SATELLITE DATA AVAILABLE FOR SURFACE FLUXES

Having identified in Equations (10.1) to (10.6) the type of parameters that must be known in order to estimate gas and heat fluxes, this section considers which of them can be measured from satellites. Also mentioned are properties that need to be known but which are not retrieved directly from satellites although remote sensing can contribute partly to defining their global distribution. These are water vapor, air temperature, and gas concentrations, all in the lower part of the ABL, and gas concentration in the upper layer of the ocean.

10.3.1 Sea surface temperature

Chapters 7 and 8 in *MTOFS* explain how infrared and microwave radiometers operate to measure sea surface temperature (SST). It is important when dealing with SST in the context of surface fluxes to note that satellites measure the temperature of the actual surface or skin of the sea. This is not the same as conventional databases of sea "surface" temperature acquired from *in situ* measurements that are normally made at a depth between 1 m and 8 m. Sometimes in coupled modeling studies the SST used to drive radiative heat flux is the temperature of the upper mixed layer that forms the top layer of the ocean model. Both of these differ from the true skin temperature seen by infrared satellites, for the reasons discussed in detail in section 7.3 of *MTOFS*.

Two factors need to be considered. First, because the flow of heat out of the surface must be driven though a microlayer at the surface where the absence of turbulence makes the thermal conductivity of the water orders of magnitude less than normal, a temperature gradient builds up until the skin is about 0.17°C cooler than the subskin temperature at a depth of about 1 mm (Donlon *et al.*, 1999). Second, the occurrence of diurnal variability of the temperature profile in the upper 10 meters of the water column in response to solar heating may result in the skin temperature rising by about 1°C during the day and then cooling again at night. Although difficult to monitor and predict because it is very dependent on wind stress (a sudden strong wind burst can rapidly destroy the diurnal thermocline) its occurrence is widespread (Gentemann *et al.*, 2003; Stuart-Menteth *et al.*, 2003) and in places the warming may be as much as 5°C.

Infrared radiometers measure radiation emitted at the temperature of the top skin of the ocean surface. Sensors such as the Along-Track Scanning Radiometer (ATSR) deliver an SST product that explicitly aims to represent skin SST. Others such as the Advanced Very-High Resolution Radiometer (AVHRR) also observe the skin temperature although many of their derived SST products are calibrated against *in situ* SST measurements which introduces uncertainty because of the two factors noted above. Infrared sensors on polar satellites provide global coverage at length-scales down to 1 km with up to two samples a day and are capable of resolving temperature differences as small as 0.1°C, with an absolute accuracy of 0.2°C. On geostationary platforms, infrared SST coverage is not global, but within the field of view the sampling interval is 1 hour or less, and spatial resolution 2 km to 5 km depending on how oblique the view is. When intercalibrated using polar sensors, the accuracy can be around 0.3°C.

Passive microwave sensors provide sea surface temperature measurements that are independent of cloud cover but at reduced spatial resolution (typically 50 km although sampled every 25 km). Because microwaves can penetrate the nonturbulent sublayer (the thermal skin layer), the temperature measurement retrieved from microwave radiometers in principle approximates subskin SST (i.e., the layer between 1 mm and 1 cm depth, immediately below the cool skin mentioned above).

These measurements are of obvious benefit for determination of heat fluxes across the air–sea interface. Moreover, since temperature has a strong influence on

the solubility of gases in seawater it also plays an important role in determining gas fluxes (Ward *et al.*, 2004). Lastly, momentum flux depends strongly on the stability of stratification of the marine atmospheric boundary layer (ABL), so SST measurements are of interest in this context.

Subtle differences between the SST measured by different satellite sensors and conventionally can be important when used in the context of air–sea fluxes. Data need to be treated with caution and might require bias adjustment before they are used together for calculations of fluxes. In recent years, measurements of SST from diverse sensors (discussed by Robinson and Donlon, 2003) have been harmonized by the Group for High-Resolution SST (GHRSST),[2] as discussed further in Chapter 14. Level 2 SST data in the GHRSST "L2P" format (Donlon *et al.*, 2007) contain the original SST retrievals from each particular producer, but they are also accompanied by ancillary data which should allow users to remove known biases between different products before blending them for use in flux estimations. Building on the GHRSST initiative, applying optimal interpolation to combinations of bias-adjusted SST products from several sensors, a number of agencies are developing new, daily, global SST analysis products for near-real time operational use (e.g., Donlon *et al.*, 2010), with the potential for use in flux estimation. Also expected in the near future are reanalysis products for climate applications.

10.3.2 Wind

Wind speed is a parameter derived operationally from a variety of remote-sensing instruments (as discussed in Chapter 9). Scatterometers, synthetic aperture radars, altimeters, and passive microwave sensors have all been used to obtain wind speed, although scatterometers are the only instrument providing proven wind direction along with speed and are probably used most widely today. At the same time, wind speed determines the dominant horizontal movement at the air–sea interface (assuming that ocean current speeds are usually significantly less than wind speeds). It is not surprising that the most commonly used flux parameterizations for momentum, heat, moisture, and gases depend on wind speed at the sea surface, typically represented by measurements or estimates of u_{10}, which is wind speed normalized to a height of 10 m assuming a neutrally stable ABL. However, it is also commonly acknowledged that other physical and biochemical processes, which do not scale with wind speed, can play an important role in determining fluxes but these are neglected in flux parameterizations that depend solely on wind speed.

Another parameter frequently used in the parameterization of fluxes instead of wind speed is the friction velocity, u^*. In contrast to u_{10}, friction velocity is a function of wind stress τ at the sea surface and air density ρ,

$$u^* = \sqrt{\tau/\rho}, \tag{10.7}$$

which can play an important role for air–sea fluxes. The magnitude of the friction velocity usually lies one to two orders of magnitude below that of u_{10}. Friction

[2] See *http://www.ghrsst-pp.org*

velocity can be obtained from satellites using dual-band altimetry such as the C-band and Ku-band data from the Poseidon altimeters on TOPEX and Jason. For further details see Elfouhaily (1998) and other references in chapter 11 of *MTOFS*. The use of u^* must be considered in relation to the question of whether it is more appropriate to use satellite retrievals of wind speed or wind stress in estimations of air–sea flux (as discussed in the next subsection).

10.3.3 Sea surface roughness

Wind speed data gathered with the help of remote sensors are based on measurements of the normalized radar cross section, σ_0. This is related to mean square slope, S_ζ^2, of the surface waves in a certain waveband, from which u_{10} is then calculated based on empirical models (as outlined in Chapter 9). The theoretical relationship between surface roughness and wind, which is assumed by these empirical backscatter models to apply in all conditions, is outlined in section 9.5.2 of *MTOFS* and fully explained by Kraus and Businger (1994). At its core is the idea that the roughness height, z_0, of the surface is proportional to u^{*2}, scaled by acceleration due to gravity, g, through a constant of proportionality, α, known as the Charnock Constant (Charnock, 1955). That is

$$z_0 = \frac{\alpha u^{*2}}{g}. \tag{10.8}$$

This simple dependence of air–sea fluxes on wind can be complicated by the presence of surface films. For a given wind speed these tend to reduce actual surface stress, which is what influences the fluxes, so that sea surface roughness could be a better predictor than wind speed or friction velocity derived from wind. Moreover, even when no surface film is present, the normally assumed relationships between u_{10} and surface stress are based on a neutrally buoyant ABL and may not hold for unstable ABLs. The presence of a strong swell may also change the relationship between wind speed and small-scale roughness that characterizes σ_0. These situations are ones where surface roughness is not what would be predicted by wind speed alone and so the Charnock constant may not be constant in these situations. In that case calculating wind speed from roughness, before then deriving flux from wind speed, introduces unnecessary uncertainty.

Figure 10.2a represents schematically how air–sea fluxes will be relatively low at

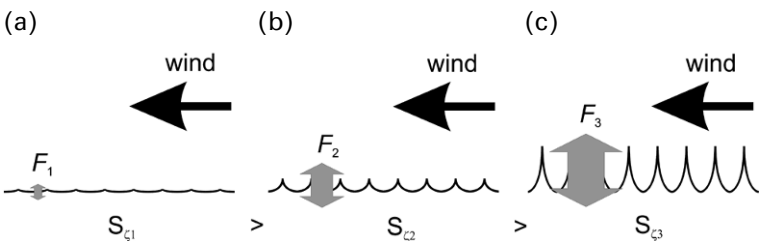


Figure 10.2. Schematic of three levels of mean square sea surface slope S_ζ^2 and air–sea flux F at given wind speeds.

a smooth surface. With increasing surface roughness—shown in (b) and (c)—both the sea surface area in contact with the atmosphere and the interaction with mean wind increase give rise to larger air–sea exchange. However, through the different stages of development of a wave field, the relationship between wind and waves is not necessarily unique since it depends on parameters such as fetch and surfactant damping, so identical wind speeds could produce quite different degrees of roughness. In this context radar remote-sensing instruments provide an opportunity to use direct measures of sea surface roughness instead of wind speed for parameterizations of air–sea fluxes.

Establishing these direct relationships between flux and sea surface roughness is a relatively young technique which has developed specifically as a result of satellite measurements, since roughness at small scales does not constitute a classical oceanographic parameter. The exact wavelength to which individual sensors are sensitive depends on the wavelength(s) of the sensor and the incidence angle at which the sea surface is monitored. For oblique-viewing radar it depends on Bragg radar scattering (explained in more detail in chapter 10 of *MTOFS*). Recent developments using σ_0 directly for flux measurements are mentioned in Section 10.4.2 in relation to gas transfer velocity algorithms (Glover *et al.*, 2002; Woolf, 2005; Fangohr and Woolf, 2007).

10.3.4 Significant wave height and wave age

Significant wave height, H_S, can be obtained from altimetry as described in Chapter 8. Wave age (mentioned in Section 8.2.3) may also be estimated from satellites. Both these parameters contain information about whether surface roughness is the same as would be expected if the sea state were in equilibrium with the wind. Thus they could be used to bring to flux estimates extra information about factors such as swell, surface films, white-capping, and sea spray that are partially or completely independent of the wind.

Therefore H_S and wave age are potential candidates as inputs to new flux parameterizations, as an alternative or to supplement algorithms based on σ_0 (as discussed in the previous subsection). However, despite the seemingly obvious link between air–sea fluxes and wave properties describing the shape of the sea surface, significant wave height data have not yet generally been used operationally in air–sea flux estimations.

The one exception is within an upgrade of an algorithm used for bulk parameterization of fluxes in connection with the Coupled Ocean–Atmosphere Response Experiment, known as the COARE algorithm (Fairall *et al.*, 1996). A revised algorithm (Fairall *et al.*, 2003) contains two optional alternatives to the standard approach that specifies the Charnock constant for determining roughness height of the sea surface as a function of wind speed. In one alternative (Taylor and Yelland, 2001) roughness height is given as:

$$z_0 = 1{,}200 H_s (H_s/L_p)^{4.5}, \tag{10.9}$$

where L_p is the wavelength of the dominant frequency in the wave spectrum. The numerical coefficients are based on an empirical fit to a large number of measurements. This takes the place of Equation (10.8). The second alternative (Oost *et al.*, 2002) represents the Charnock factor (no longer a "constant") as

$$\alpha = 50(C_p/u^*)^{-2.5}, \tag{10.10}$$

where C_p is the phase speed of the dominant wave; and (C_p/u^*) is a measure of wave age.

Although the COARE algorithm is mainly used in numerical-modeling studies, it can be applied to the analysis of observational data (including satellite data) and so these are pointers to how global air–sea flux calculations might be developed to include significant wave height or wave age measurements from satellites.

10.3.5 Water vapor

The amount of water vapor contained in the atmosphere can be derived from passive-microwave radiometers such as the SSM/I, TMI, AMSR-E, and WindSat (see Table 2.6 for details about these sensors). While water vapor is considered a meteorological rather than oceanographic measurement, it quantifies a property that, if known at the air–sea interface, is important for air–sea heat flux since it allows derivation of the mixing ratio (Liu, 1985) and its measurement at the 10 m height essential for estimating latent heat flux. Standard water vapor products from microwave radiometers actually retrieve total column water. It is possible to retrieve an average over the lower 500 m of the atmosphere (Schulz *et al.*, 1993) but this normally differs considerably from water vapor at 10 m which is characterized in the literature as surface specific humidity, Q. Several attempts have been made to estimate this globally from other satellite-derived quantities.

The most recent of these (Zong *et al.*, 2007) gives a review of previous algorithms (including Schulz *et al.*, 1993, 2003) from which theirs is developed. It uses measurements of SST, total column water vapor (W), and wind speed (U), all derived daily from AMSR-E, in an empirical algorithm of the form:

$$Q = a + b\,\mathrm{SST} + c(\mathrm{SST})^2 + dW + eW^2 + fU.$$

The coefficients, a to f, are determined empirically by regressing on coincident values of Q from the NCEP reanalysis for 2003. Separate algorithms are determined for daily mean values and monthly mean values. When tested against 2004 data, the r.m.s. error for the global dataset was found to be 1.05 g/kg for daily retrievals and 0.61 g/kg for monthly estimates. These results show some promise, considering that the global range of NCEP values of Q is between about 1.5 g/kg and 22 g/kg. The poorest results appear to be where air–sea temperature differences are larger than normal (e.g., over western boundary currents), and it may be possible to derive special algorithms for these regions. Nonetheless, it seems unreasonable to expect microwave radiometers such as SSMI or AMSR-E to detect variations in sea surface humidity that are independent of total column water vapor. Therefore undue reli-

ance on these satellite-retrieved estimates of Q risks missing important variability in the true value of Q that may be significant in the evaluation of latent heat fluxes. It would seem prudent to evaluate these products, and the sensitivity of fluxes derived from them, before using satellite-derived Q as a substitute for *in situ* measurements.

10.3.6 Air temperature at sea level

There are no ways of directly measuring air temperature at 10 m height, T_a, using a remote-sensing instrument on a satellite. Although the infrared and microwave radiation reaching a satellite sensor is influenced to a small extent by T_a it is virtually impossible to isolate that information from all the other controlling factors such as sea temperature and the profiles of temperature and water vapor through the whole atmosphere. Instead, attempts to retrieve T_a are based implicitly on the assumption that it is related to those other atmospheric and ocean properties in such a way that it can be determined from knowledge of them, or more particularly from those properties which can be reliably measured such as SST and total atmospheric water content. This results in a very similar approach to that adopted for water vapor at 10 m outlined in the previous paragraph.

Thus Gautier *et al.* (1998) created an artificial neural net (ANN) model that predicts monthly mean T_a from inputs of SST (in their case derived from NCEP rather than directly by remote sensing), and total precipitable water (W) derived from the SSM/I microwave radiometer. Although the ANN performs quite well in validation tests (Jones *et al.*, 1999) with standard deviation of 0.72°C, the usefulness of monthly means is limited. In order to exploit the daily remote-sensing sampling capability of SST and W, a model for deriving instantaneous T_a is needed. This has now been reported, in one case (Singh *et al.*, 2006) based on SST from AVHRR and W and total column water vapor (W_b) both from SSM/I. In another case (Jackson *et al.*, 2006) several regression models for T_a are examined, combining satellite microwave observations from the Advanced Microwave Sounding Unit-A (AMSU-A), SSM/I, and the Special Sensor Microwave Temperature Sounder (SSM/T-2). The most promising T_a retrieval model is based on inputs from SSM/I and AMSU-A. However, as with estimates of Q, it would be unwise to adopt the use of satellite-derived T_a before thorough evaluation of the errors and their impact on sensible heat flux calculations.

10.3.7 Gas concentrations in the surface sea and the ABL

Gas concentration in water is usually defined in terms of its partial pressure, pX (thus for CO_2 we refer to pCO_2). There are no known ways of measuring or estimating the gas concentration in surface waters of the ocean directly using remote-sensing techniques. This obviously presents problems for estimation of global fluxes of gases between ocean and atmosphere. Such studies normally make use of accumulated observational data. For CO_2 the global coverage is reasonable and data from *in situ* measurements have been accumulated monthly on a grid of 4° latitude × 5° longitude (Takahashi *et al.*, 2002).

The use of climatological values does not allow interannual variability to be identified, although in the case of CO_2 the steady increase of atmospheric CO_2 is factored in. They offer a reasonable picture of how gas concentrations in the upper ocean vary with location and season, which allows the dependence of global fluxes on physical flux drivers (wind and temperature) to be explored. But ideally we want to know how surface pCO_2 varies in response to the processes that may drive it (e.g., primary production or upwelling), so that we might understand how air–sea gas fluxes mediate these changes between the ocean and atmosphere or *vice versa*.

It may be that a relationship can be found between surface pCO_2 and other more readily measurable ocean variables which would allow us to estimate pCO_2 from these proxies with tolerable accuracy. Steps have been made in this direction as a result of a program of *in situ* monitoring of pCO_2 from ships of opportunity along regular shipping routes, allowing pCO_2 to be modeled as a function of temperature, latitude, and longitude (Lefèvre and Taylor, 2002; Lefèvre *et al.*, 2005). Should such techniques eventually prove robust then satellite data could be used to provide some of the proxy data. As an alternative, numerical ocean models with integrated biogeochemical components including phytoplankton primary production are being developed, with a view to predicting surface pCO_2 concentration for air–sea flux estimation (Hemmings *et al.*, 2008). These use satellite-derived ocean color and temperature data in their assimilation scheme.

For atmospheric gas concentrations, it is very difficult to obtain regularly updated observations defining the spatial distribution of gas partial pressures in the ABL. Each particular gas must be considered separately, but in general it can be expected that over the ocean and away from specific sources the horizontal variability lengthscale of gas concentration in the atmosphere is at least 1,000 km, so that widely spread *in situ* measurements should be adequate to define global distributions. Many studies use climatologies or time-evolving databases of surface gas concentrations such as Globalview (2007) which contains CO_2, CH_4, and CO. To convert gas concentrations to partial pressure they must be scaled by the difference between sea level barometric pressure and water vapor pressure.

There are also atmospheric-sounding sensors in space, such as SCIAMACHY on Envisat and AIRS on Aqua, from which estimates of CO_2 and other gases can be obtained (Barkley *et al.*, 2006). These results offer more precision over land than over sea, and measure total gas content through the whole atmospheric column, but they show promise for more precise atmospheric gas sampling in the lower part of the atmosphere in future (Barkley *et al.*, 2007).

10.4 MEASURING FLUXES FROM SPACE

10.4.1 Radiative flux

Determination of radiation terms in net heat flux (Q_b and Q_S in Equation 10.4) requires radiative transfer modeling to account for the effects of the atmosphere

under a given set of conditions (e.g., Pinker and Laszlo, 1992). Longwave net flux Q_b is computed from atmospheric back radiation $R_L^{\downarrow}$ retrievable from satellite data and sea surface temperature T_s. Adopting the sign convention that fluxes are positive in the direction from sea to atmosphere, we write:

$$Q_b = \bar{\varepsilon}\sigma T_s^4 - \bar{\varepsilon}R_L^{\downarrow}, \tag{10.11}$$

where $\bar{\varepsilon}$ is spectrally integrated surface emissivity which is close to 0.89 (Gardashov *et al.*, 1988); and σ is the Stefan–Boltzmann constant. Further details of this procedure are described, among others, by Schanz and Schlüssel (1997).

The ready availability of daily-updated global SST gives satellite data a strong part to play in the evaluation of net longwave flux. It is important to note that it is skin temperature (within about 100 μm of the surface) that controls longwave radiation, so ideally skin SST is required. It should be suspected that radiant fluxes based on *in situ* observations of SST fail to allow for the slightly reduced radiation because of the cool skin effect, and do not include the excess radiation associated with local diurnal warming events. While these two effects act in opposition to each other, that does not justify their neglect. On the other hand, care must be taken when introducing skin SST into radiation models since any tunable coefficients may previously have been optimized to ensure that the use of "bulk" SST data achieves thermodynamic closure.

Incident, solar, short-wave flux, Q_S, either passes through the sea surface or is reflected, but has very little effect on conditions in the sea or air close to the boundary. That is because water is quite transparent to visible wavelength light, especially in the blue where the peak of the solar spectrum occurs. It is only when a photon of light interacts with a water molecule and is absorbed that any thermal energy transfers into the water. Solar heating of the upper ocean should therefore be treated like internal heating, distributed through the water column although exponentially decreasing with depth as light attenuates. Thus whereas the effects of the other terms in Equation (10.4) are all applied to water at its surface, solar heating is smoothly applied over depth. That is why, when solar insolation is strong, but there is heat transfer from the ocean to the atmosphere through evaporation driven by wind, the cool skin of the ocean remains cooler than water below the surface microlayer, even though this may at first seem counter-intuitive.

The magnitude of solar shortwave energy is required to understand the vertical temperature and density structure of the upper ocean and can be determined by remote-sensing methods as outlined in Section 7.3.3 in relation to PAR (which can be considered to be a spectrally filtered version of Q_S expressed in quanta rather than standard units of energy). Satellite ocean color estimates of the diffuse attenuation coefficient, K_D, for solar radiation are also important for determining the depth over which solar radiation is absorbed.

10.4.2 Gas flux

Using Equation (10.2) to calculate gas fluxes requires knowledge of the solubility of gas at the temperature of the sea surface, gas transfer velocity k, and the gradient of the partial pressure of the gas across the interface.

Partial pressure gradient

The partial pressure gradient of gases across the interface cannot be measured from space or reliably parameterized in a straightforward way (as discussed in Section 10.3.7). Instead, global climatologies such as the one produced by Takahashi *et al.* (2002) for carbon dioxide are the current, state-of-the-art products providing temporal and spatial coverage of variability at scales relevant to global remote sensing. Such climatologies will be improved through increased coverage of the World Ocean by *in situ* measurements.

Ideally an aim should be to monitor the time-evolving distribution of ocean surface pCO_2 or other gases with a time resolution of, say, one week, in order to detect its response to biological and physical processes in the upper ocean, and hence to understand the impact on air–sea gas exchange of time-varying gas concentration in the ocean. To achieve that is likely to require the use of the biogeochemical models mentioned in Section 10.3.7, supported by remotely sensed visible and infrared data, but confidence in such an approach probably still lies some years ahead.

Solubility

Meanwhile, satellite data already have an important part to play in implementing the other terms in Equation (10.2). Weiss (1974) and Wanninkhof (1992) provide values of solubility, s, in units of concentration/pressure as a function of SST for a variety of gases. These are quite strongly dependent on sea temperature; for example, the solubility of CO_2 reduces by more than half when temperature rises from 0°C to 20°C (as shown in Figure 10.3a), although this is substantially offset by the contrary temperature dependence of the k term in (10.2). Strictly it is surface skin temperature that controls gas solubility in the water at the air–sea interface. Satellite measurements can provide daily updates of SST, although until recently many studies used SST climatologies based on *in situ* (subsurface) measurements. Ideally account should be taken of diurnal warming activity since recent modeling has shown it to have measurable effect (Jeffery *et al.*, 2008; Kettle *et al.*, 2008), although logistically it may prove difficult to do so. In addition there is salinity dependence (also illustrated in Figure 10.3 by lines of different shading), but this results in much smaller changes in s across the ocean than the variability associated with global SST distribution.

Gas transfer velocity

The remaining term to consider from Equation (10.2) is gas transfer velocity, k, and here satellite-derived global wind distributions have made a very considerable impact, allowing localized estimates of gas exchange to be expanded to the global scale. Over the last two decades, a few parameterizations of the velocity of gas

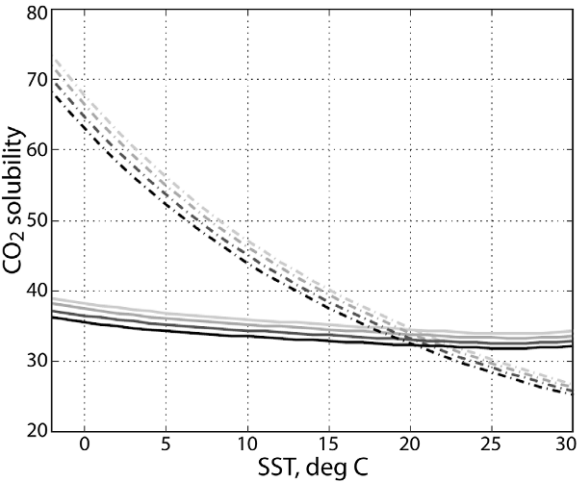


Figure 10.3. Variation with water temperature of the solubility of CO_2 (dashed lines) and the product $s \cdot (Sc)^{-0.5}$ (solid lines). In each case the set of four differently shaded lines corresponds to four salinities; 26 psu (palest line), 30 psu, 35 psu, 40 psu (darkest line) (these plots are based on formulations in Wanninkhof, 2002).

transfer through the air–sea interface have been published that follow a common scheme (Liss and Merlivat, 1986; Wanninkhof, 1992; Wanninkhof and McGillis, 1999; Nightingale *et al.*, 2000). They rely on wind speed as the main parameter determining the magnitude of gas flux, expressing the transfer velocity k as a polynomial of u_{10} and the Schmidt number, Sc.

Sc is the ratio (μ/D) between the kinematic viscosity of the water, μ, and the diffusivity, D, of the gas in seawater. Sc can be calculated for a specific gas as a function of SST (see Wanninkhof, 1992 for values). Note that for CO_2 the $(Sc)^{-0.5}$ term increases with temperature and almost compensates for the solubility v temperature effect, but not completely (as shown in Figure 10.3b). For example, the magnitude of the product $s \cdot (Sc)^{-0.5}$ reduces by about 10% between 0°C and 20°C, but starts to increase as temperatures rise above 27°C. Thus SST dependence of CO_2 flux is not dominant but remains important.

Table 10.1 gives details of the four most commonly used versions of specific parameterizations of gas transfer velocity. Figure 10.4 shows that the four parameterizations vary significantly in their predictions for k, especially at moderate to high wind speeds where significant amounts of gas are transferred across the interface, indicating that the processes influencing gas transfer are not fully explained by wind speed and temperature alone.

Integrating nonlinear terms

The nonlinear dependence of gas flux on u_{10} which applies to all parameterizations in Table 10.1 raises an important issue. Quadratic dependence means that doubling the wind speed will result in four times the gas flux, or eight times in the case of cubic dependence. If total gas flux over a period of time, say a month, is to be accurately evaluated it is important to sample the wind as often as possible so that peak bursts of wind that make a disproportionately large contribution to time-integrated flux are

Table 10.1. Parameterizations of gas transfer velocity k.

Published by	*Parameterization of transfer velocity*	
Liss and Merlivat (1986)	$k = 0.17u_{10} \cdot \left(\frac{Sc}{660}\right)^{-1/2}$	$(u_{10} < 3.6\,\mathrm{m/s})$
	$k = (2.85u_{10} - 9.65) \cdot \left(\frac{Sc}{660}\right)^{-1/2}$	$(3.6\,\mathrm{m/s} > u_{10} > 13\,\mathrm{m/s})$
	$k = (5.9u_{10} - 49.3) \cdot \left(\frac{Sc}{660}\right)^{-1/2}$	$(u_{10} > 13\,\mathrm{m/s})$
Wanninkhof (1992)	$k = 0.31u_{10}^2 \cdot \left(\frac{Sc}{660}\right)^{-1/2}$	(instantaneous winds)
Wanninkhof and McGillis (1999)	$k = 0.0283u_{10}^3 \cdot \left(\frac{Sc}{660}\right)^{-1/2}$	(instantaneous winds)
Nightingale *et al.* (2000)	$k = (0.333u_{10} + 0.222u_{10}^2) \cdot \left(\frac{Sc}{660}\right)^{-1/2}$	

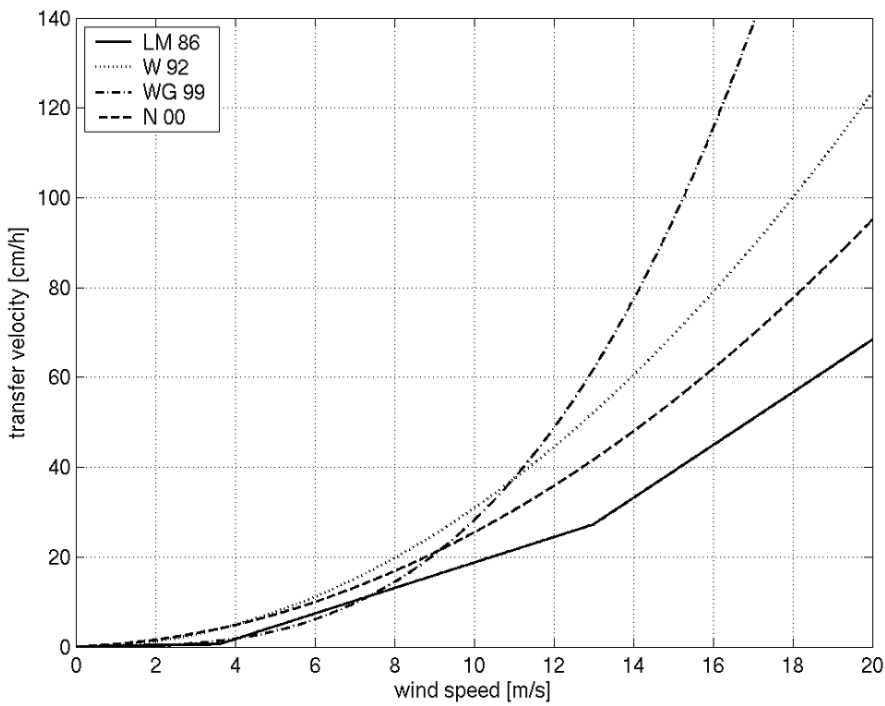


Figure 10.4. Parameterizations of the gas transfer velocity k for different wind speeds at a 20°C sea surface temperature. LM 86: Liss and Merlivat (1986); W 92: Wanninkhof (1992); WG 99: Wanninkhof and McGillis (1999); N 00: Nightingale *et al.* (2000).

not missed. Six-hourly wind data should effectively sample actual wind variability, and the present satellite capability of once-daily or twice-daily sampling over much of the globe is adequate. Since the temperature dependence of fluxes, though less than wind dependence, may change from day to day it is also important to update SST daily, if possible.

However, when calculating fluxes globally over many years to produce climatologies, researchers often prefer to simplify the integration. Monthly averaged values of wind, $\bar{u}_{10}$, are used in a single calculation of the transfer velocity for that month to produce $k(\bar{u}_{10})$ instead of evaluating $k(u_{10})$ for each daily or more frequent sample of u_{10} and then averaging those results to produce $\bar{k}(u_{10})$ which represents the true average. Because k is nonlinear in u_{10} these do not produce the same result. To allow for this a correction factor R must be introduced, where

$$\bar{k}(u_{10}) = R \cdot k(\bar{u}_{10}) \tag{10.12}$$

and in general $R \neq 1$ because of the nonlinearity of the function. To evaluate R requires knowledge of the probability distribution of u_{10} in each of the locations and over the time interval for which the flux is being evaluated. However, it has been common practice to assume a Rayleigh distribution of wind speeds in all places at all times, leading to assigning $R = 1.25$ for a quadratic dependence of k on wind speed and 2.17 for a cubic relationship.

Wanninkhof (2002) pointed out that using these values of R will introduce errors over large oceanic regions since realistic frequency distributions can be quite different from a Rayleigh distribution. Figure 10.5 shows values of R that are needed at each location around the world to apply an acceptably accurate correction when using the Wanninkhof (1992) quadratic expression for gas flux (see Table 10.1) to evaluate monthly fluxes using monthly mean QuikScat winds instead of integrating from 12-hourly samples (Fangohr *et al.*, 2008). It shows that the relatively constant speeds of the trade winds in the tropics produce lower values of R. In contrast, R values exceed 1.25 in high-latitude regions which often experience high wind speed events.

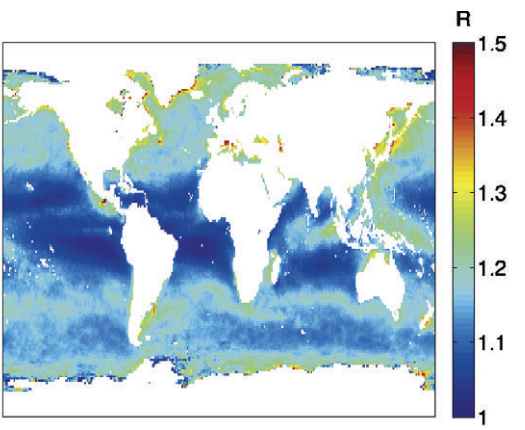


Figure 10.5. Global values of the correction factor R as defined in Equation (10.12) required to account for the nonlinear dependence of the Wanninkhof (1992) flux parameterization, when using monthly mean winds instead of integrating 12-hourly samples. It is based on the frequency distribution of wind speeds derived from 2 years of 12-hourly QuikSCAT data (this is the same as figure 7a of Fangohr *et al.*, 2008).

Kettle and Merchant (2005) show that a similar effect exists for the covariance of wind speed and air pressure, directly influencing the partial pressure of any atmospheric gas. They find that this can produce further inaccuracies in flux estimates of up to 22% when monthly mean values of partial pressure are used.

Non-wind based parameterizations

As discussed in Sections 10.3.2–10.3.4, it can be argued with some justification that gas transfer velocity is more closely dependent on actual surface roughness, characterized by mean square slope, than by wind (Frew *et al.*, 2004). The presence of surfactant films results in surface roughness being less than normal for a given wind, so that wind-based parameterizations overestimate fluxes in such conditions. It should be stressed that there is no evidence that the film itself restricts the flow of gas but its surfactant effect mechanically reduces wave steepness and turbulence compared with conditions in the absence of a film and this is what influences k.

In order to avoid this problem, Glover *et al.* (2002) presented a new parameterization that relates transfer velocity directly to roughness of the sea surface, using mean square slope values from the dual-frequency altimeter TOPEX. Use of the return signal at two frequencies allows them to isolate wave spectra in the region of 6.3 cm to 16.5 cm which is relevant to gas transfer. Further assessment of this approach (Frew *et al.*, 2007) confirms that resulting transfer velocity fields, evaluated every 7 km along the altimeter track and then gridded monthly at 2.5° resolution over the whole ocean, are generally consistent in magnitude and dynamic range with k evaluated by the parameterizations in Table 10.1.

Another new algorithm (Woolf, 2005) follows a similar approach using dual-frequency altimeter backscatter to represent the transfer velocity, k_d, for direct gas exchange, but adds to this an additional transfer velocity, k_b, associated with bubble-mediated gas transfer, so that $k = k_d + k_b$. Implementation and assessment of this method (Fangohr and Woolf, 2007) shows that it also gives reasonable results, although there is a shortage of field observations for confirming what the relative balance should be between k_d and k_b.

Towards global budgets of CO_2 flux?

As a result of continuing uncertainties about which coefficients and wind or roughness dependence functions are most appropriate for global application, researchers are cautious about attempting to present a wholly realistic description of global geographical and seasonal variation of gas flux, although some have pointed out the sensitivity of global CO_2 flux estimates to different algorithms even when global mean k does not change (e.g., Fangohr and Woolf, 2007). It is important to note how sensitive the calculation of net global air–sea CO_2 flux can be to relatively small changes in the geographical distribution of k. The reason for this is the way in which the driving gas gradient (pCO_{2w}–pCO_{2a}) varies geographically not only in magnitude but also in its sign.

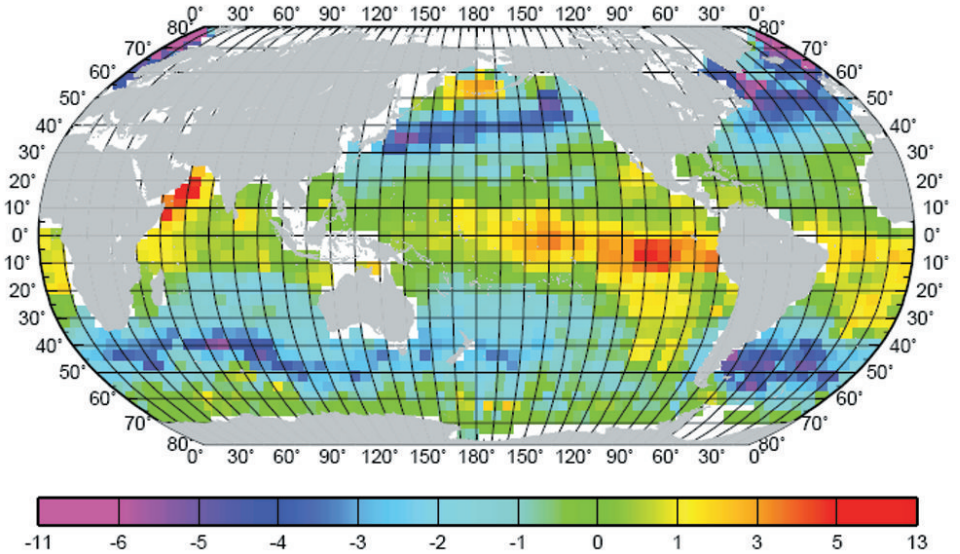


Figure 10.6. Mean annual net sea-to-air flux for CO_2 (mole CO_2 m^{-2} per year) in 1995 as calculated by Takahashi *et al.* (2002). It is based on the following information: (a) climatological distribution of surface water pCO_2 for the reference year 1995, (b) NCEP/NCAR 41-year mean wind speeds, (c) long-term wind speed dependence of the sea–air CO_2 transfer velocity of Wanninkhof (1992), (d) the concentration of atmospheric CO_2 in dry air in 1995 obtained from the GLOBALVIEW CO_2 2000 database, and (e) climatological barometric pressure and sea surface temperature from the 1994 *NODC Atlas of Surface Marine Data.*

Figure 10.6 (based on Takahashi *et al.*, 2002) presents the estimated annual mean distribution of CO_2 flux for 1995, showing clearly the large areas where this dataset implies CO_2 flows from ocean to atmosphere (mainly in the tropics) and other large regions (mainly subpolar ocean gyres) where the reverse is true. This indicates that net global exchange is a relatively small difference between two larger figures representing total CO_2 drawdown and total outgassing. If we are unlucky, and the geographic distribution of errors in transfer velocity are correlated with the direction of gas flux, then the impact on net gas exchange could be rather large. Uncertainties of a few percent in the parameterization of gas flux could amplify to much larger errors in net gas exchange.

Until we can be sure that our flux estimates are regionally and seasonally accurate it would be prudent to be very cautious about drawing global conclusions from present estimates. This highlights the importance of seeking to estimate fluxes with fine spatial and temporal resolution. Even though remote sensing is not able to give us all the data we need, its capacity to measure SST and winds is providing a sound foundation on which this important field of research can build, and new roughness-based algorithms show considerable promise for eventually being able to operate a global flux–monitoring system (Glover *et al.*, 2007) that will be able to identify any changes in the distribution of oceanic sources and sinks for CO_2.

10.4.3 Turbulent heat flux

Input variables and coefficients for heat flux parameterizations

Equations (10.5) and (10.6) for the calculation of latent and sensible heat flux across the air–sea interface require a variety of input variables and parameters. Some of these can be supplied directly by remote sensing: sea surface temperature (T_s) and wind speed (u_z). Other parameters are well known: latent heat of vaporization of water (L) and isobaric specific heat of water (c_p) (Xue *et al.*, 2000). Others can be assumed with little expected loss of accuracy: specific humidity immediately above the sea surface (q_s) is taken to equal saturation humidity at T_s and p_s; u_s is taken to be zero. Air pressure at sea level (p_s) can be obtained with sufficient accuracy from meteorological forecasts or reanalyses. From this and an approximate estimate of air temperature, air density (ρ) can be evaluated.

The two properties which present the greatest challenge for global mapping of latent and sensible heat fluxes are atmospheric variables for gradient terms in flux equations: water vapor (or specific humidity, q_a) and air temperature (T_a) in the ABL, at a height of 10 m. As discussed in Sections 10.3.5 and 10.3.6, some progress has been made in estimating these from microwave radiometry using empirical algorithms tuned to *in situ* measurements. However, the reliability of these data products remains open to question. To use them in Equations (10.5) or (10.6) introduces uncertainty into resulting heat flux estimates. The alternative is to use climatologies of *in situ* measurements and/or meteorological analysis values.

The remaining two variables in Equations (10.5) and (10.6) that must be quantified are the transfer coefficients, C_T (Stanton number) and C_E (Dalton number). These have often been assumed to be equal or similar to the better understood drag coefficient, C_D, used to estimate shear stress and momentum transfer. C_D is constant under neutral conditions for wind speeds up to 11 m/s ($C_D = 1.2 \times 10^{-3}$) and then increases with the wind (Large and Pond, 1981). Published values of C_E adjusted to neutral stratification and 10 m height (C_{EN}) lie between 1.0×10^{-3} and 1.5×10^{-3} for unstable conditions. Values for C_{TN} are usually around (1.0×10^{-3}) for unstable conditions and lower than that ($0.66–0.8 \times 10^{-3}$) for slightly stable conditions (Large and Pond, 1982; DeCosmo *et al.*, 1996; Bentamy *et al.*, 2003). Figure 10.7, from Grassl *et al.* (2000), demonstrates the potential global range of values of C_E computed as a function of stability and wind speed based on climatological data.

There is still scientific debate about these values, because of a possible dependence of C_T and C_E on stability, wind speed, or parameters of the wave field. There is discussion about whether Monin–Obukhov similarity theory (Monin and Obukhov, 1954) is applicable to open-ocean heat transfer (Oost *et al.*, 2000; Edson *et al.*, 2004). Furthermore, the influence of sea spray on the turbulent structure of the marine boundary layer and thus the transfer of heat and moisture requires quantification (Andreas and DeCosmo, 2002). As mentioned in Section 10.3.4 there is ongoing debate (Fairall *et al.*, 2003) about adjustments to the Charnock parameter (see Equations 10.9 and 10.10), with perspectives shifting as improved experimental techniques change understanding of the detailed processes in

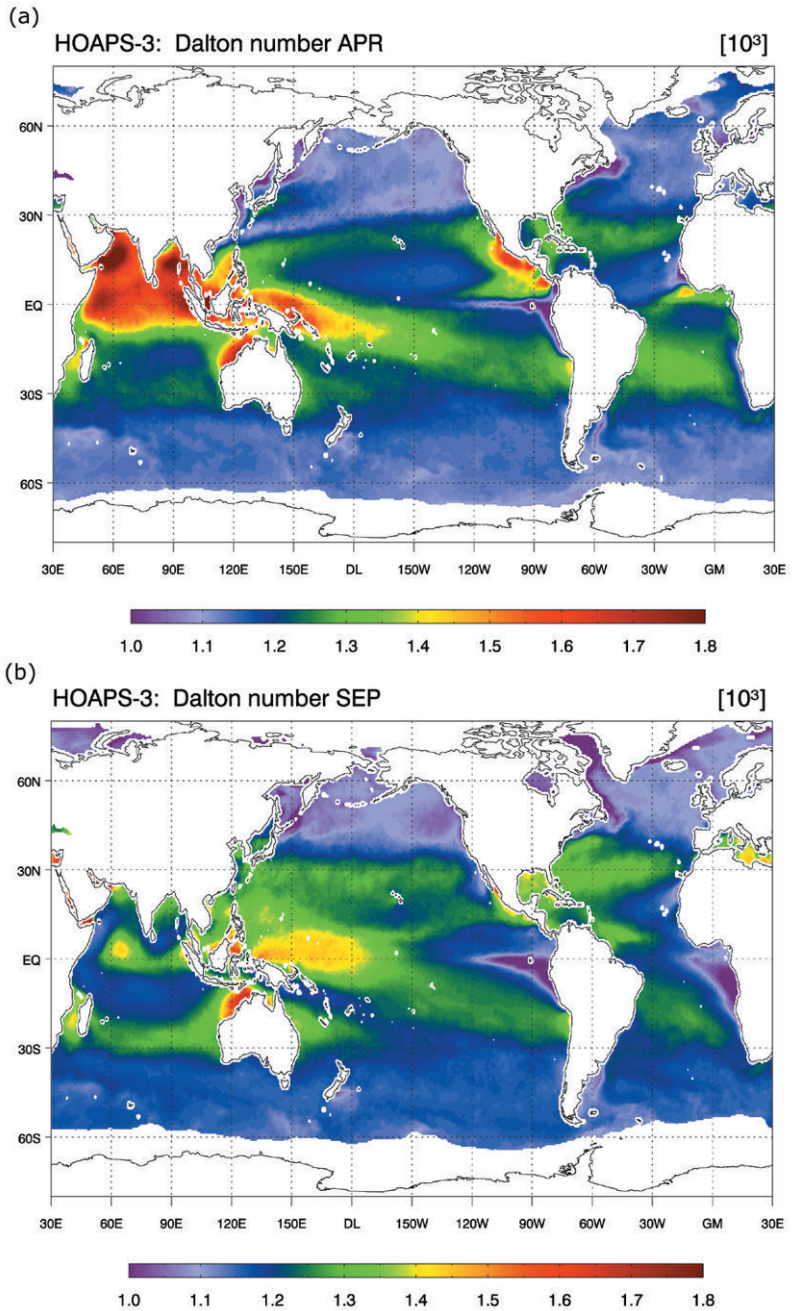


Figure 10.7. Global distribution of monthly mean values of C_E, the Dalton number, calculated as a function of atmospheric stability and wind speed, for (a) April and (b) September. Climatology based on data from 1987–2005 (images obtained from the HOAPS 3 database—Andersson *et al.*, 2007).

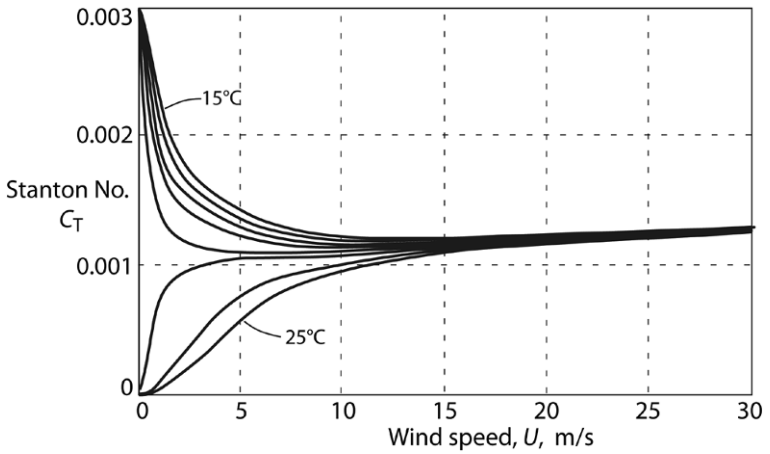


Figure 10.8. Variation of C_T with wind speed for different air temperatures. In this realization of the parameterizations used in the COARE-3.0 model, SST is 20°C and air temperatures vary from a most unstable 15°C (uppermost curve) through 17°, 18.5°,19.6°, 20.4°, 21°, 23°, to 25°C (lowest curve) which is most stable.

the ABL (Yelland and Taylor, 1996; Yelland *et al.*, 1998; Taylor and Yelland, 2000). Based on parameterizations used in the COARE-3.0 model, Figure 10.8 shows estimates of how C_T is predicted to vary with wind speed and for different temperature differences between sea surface and atmosphere at 10 m. The dependence of C_E is very similar. When air is considerably warmer than water, this formulation indicates that C_T and C_E fall well below 1.0×10^{-3} at lower wind speeds because of enhanced stability.

It is important to acknowledge that this is an active research field where complete consensus of how to understand, and then parameterize, complexities of the turbulent ABL has not yet been reached. A pragmatic way forward for remote sensing is to make use of more widely accepted parameterizations and cautiously to apply them to available data. Then the sensitivity of fluxes to changes in parameterisations of C_T or C_E will be revealed by changes in the space-time distributions of fluxes.

Flux estimates based on satellite data

Calculations of latent heat flux, which makes up the majority of air–sea turbulent heat transport, have now been performed using only satellite data inputs (apart from air density estimates) allowing them to provide global coverage on a regular basis. Different application systems show varying degrees of success when compared with *in situ* flux measurements (Jourdan and Gautier, 1995; Xue *et al.*, 2000; Bentamy *et al.*, 2003; Jo *et al.*, 2004). Discrepancies remain in those areas where there are large air–sea temperature differences and high wind speeds.

Sensible heat flux on the other hand poses a bigger problem, given the need for knowledge of air temperature. Since retrievals of air temperature (mentioned in

Section 10.3.6) are somewhat speculative, a number of alternatives have been considered to solve this problem. Atmospheric temperatures can be obtained from models or *in situ* measurements. Alternatively, specific conditions in certain areas (atmospherically convective regions of the tropics) allow modifications of bulk formulas which make direct measurement of sea level air temperature unnecessary (Fairall *et al.*, 1996; Jo *et al.*, 2004; Pan *et al.*, 2004).

Schulz *et al.* (1997) discuss the errors introduced by assuming constant pressure and various assumptions for air temperature within the bulk approach. While errors in surface pressure compensate each other through their contrary effects on q_s and ρ, errors in T_a directly enter the equation for sensible heat flux and alter the values of the transfer coefficients in all flux equations (Grassl *et al.*, 2000).

Gautier *et al.* (1998) and Jones *et al.* (1999) adopted the approach to satellite-derived air temperature and specific humidity mentioned in Sections 10.3.5 and 10.3.6. They proposed derivation of both air temperature and ocean surface specific humidity from a neural network which relies on total precipitable water and sea surface temperature, derived by passive-microwave remote sensing, for input. This allows derivation of both latent and sensible heat flux based predominantly on satellite measurements. Results of their studies are shown in Figure 10.9, illustrating the average annual mean of two turbulent flux components for a 15-year period over the tropical oceans (Jones *et al.*, 2003).

One of the first major projects to derive global air–sea heat fluxes from satellite data was the Hamburg Ocean–Atmosphere Parameters and Fluxes from Satellite data (HOAPS) project. Longwave net radiation, latent and sensible heat flux, as well as evaporation, precipitation, and freshwater flux were all computed for the period 1987–1998 providing mean monthly, seasonal, and annual data fields (Grassl *et al.*, 2000). These have since been extended to 2005 (Andersson *et al.*, 2007). Examples of latent heat flux are shown in Figure 10.10 and of sensible heat flux in Figure 10.11. However, since *in situ* measurements of heat flux are sparse and in themselves prone to errors, validation of these methods remains difficult and their accuracy is uncertain and likely to be regionally variable.

With that caution in mind, it is still instructive to consider what can be learned from Figures 10.10 and 10.11. First of all it should be noted that the range of latent heat flux is greater than sensible heat flux by a factor of almost 3. Latent heat flux is always positive, whereas sensible heat flux can be slightly negative in the summer hemisphere when the air is warmer than the sea, and in tropical regions where upwelling maintains zones of cooler surface waters in plumes off the western coasts of America and Africa. It is evident that there are large differences in the geographical distribution of fluxes between winter and summer. The high values of sea–air heat fluxes, both sensible and latent, in the winter North Atlantic confirm the importance of the ocean in preserving the moderate climate of northwest Europe.

There is little doubt that, in principle, the superior sampling capacity of remotely sensed SST, surface wind speed, and water vapor greatly improve the potential for obtaining global measurements of turbulent heat fluxes that are of geographically uniform quality and capable of being updated daily. This contrasts with conventional flux climatologies which are typically monthly accumulations of data, and

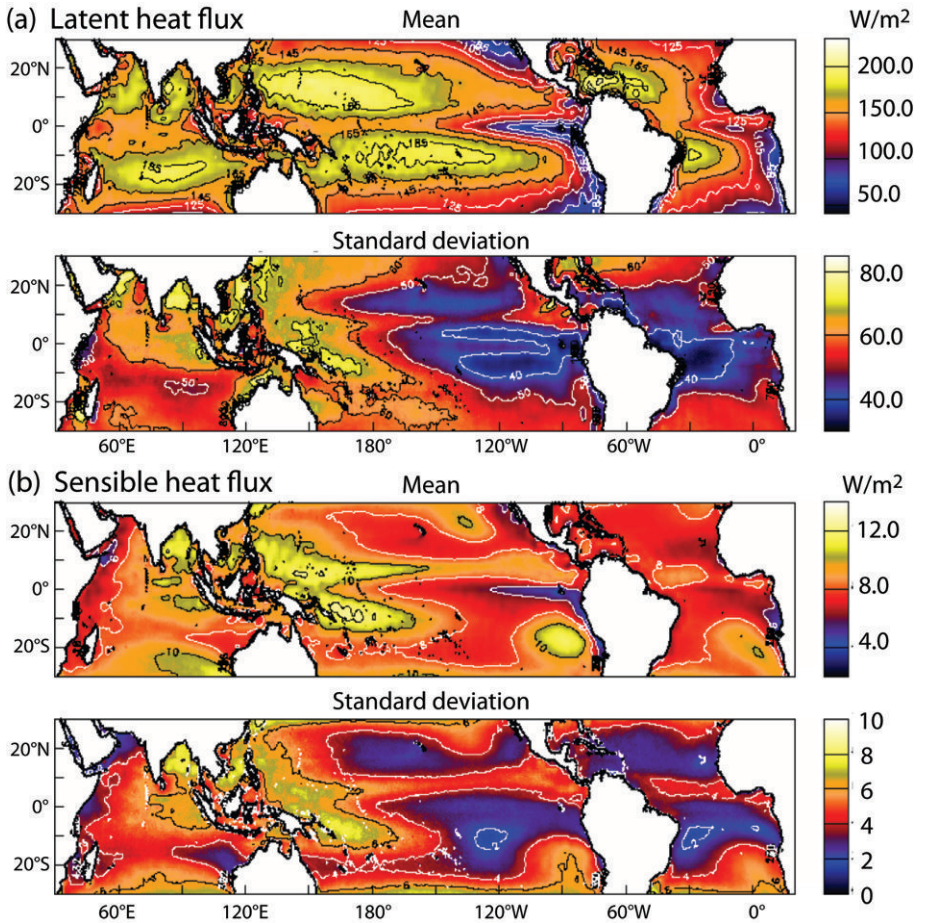


Figure 10.9. Fifteen-year mean and standard deviation of (a) latent and (b) sensible heat flux derived from satellite data for October 1987 to September 2002 (figures adapted with permission from Jones *et al.*, 2003).

whose quality tends to be highest in regions where there are many samples and poorest in remote parts of the ocean. Nonetheless, the question of whether flux climatologies derived from remote sensing can match the best quality of those derived from conventional data (such as Yu and Weller, 2007), remains to be definitively answered.

10.5 SATELLITE FLUX MEASUREMENTS IN FUTURE?

Estimating air–sea fluxes of momentum, heat, moisture, and gases from remotely sensed data is a fairly young area of research. Only recently have we reached a stage

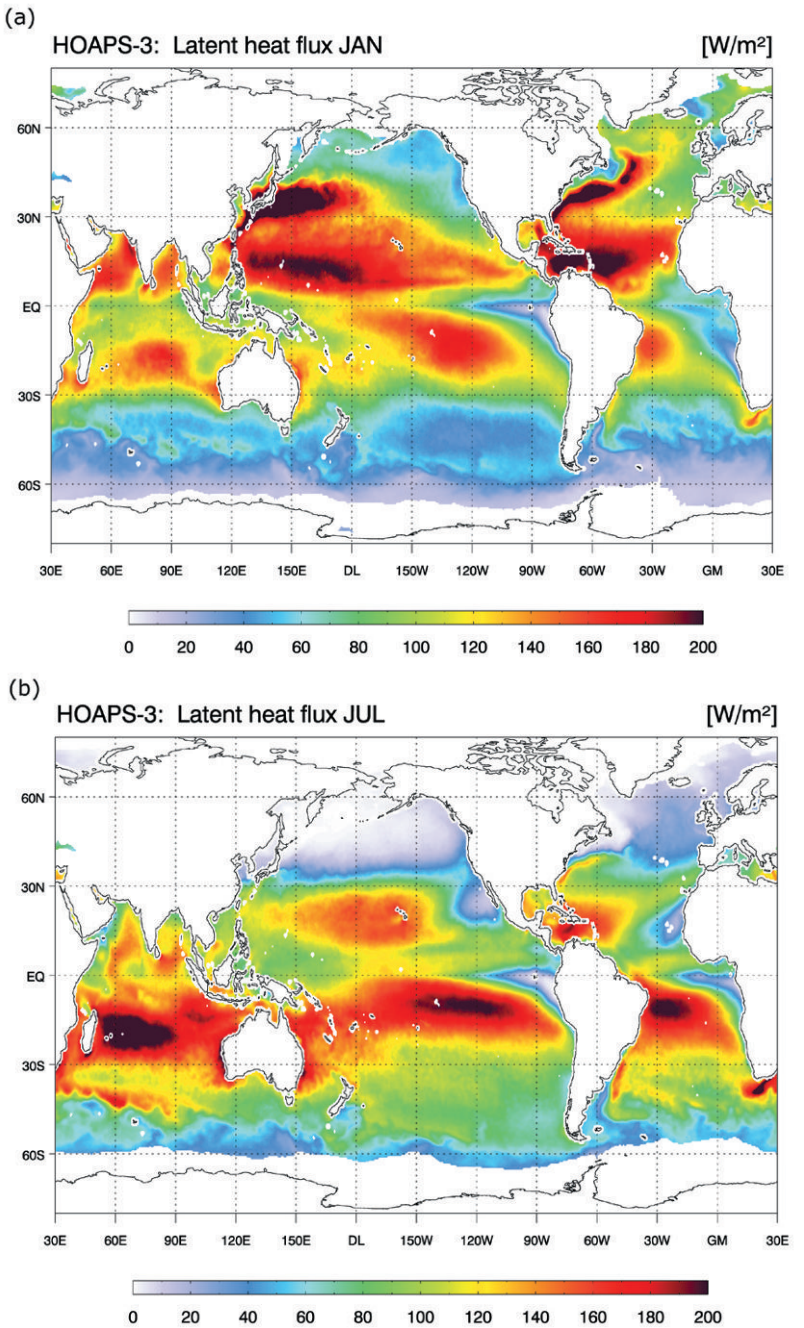


Figure 10.10. Monthly mean climatology of global latent heat fluxes calculated on the basis of satellite data for the period 1987–2005. (a) January, (b) July (from HOAPS 2 database—Andersson *et al.*, 2007).

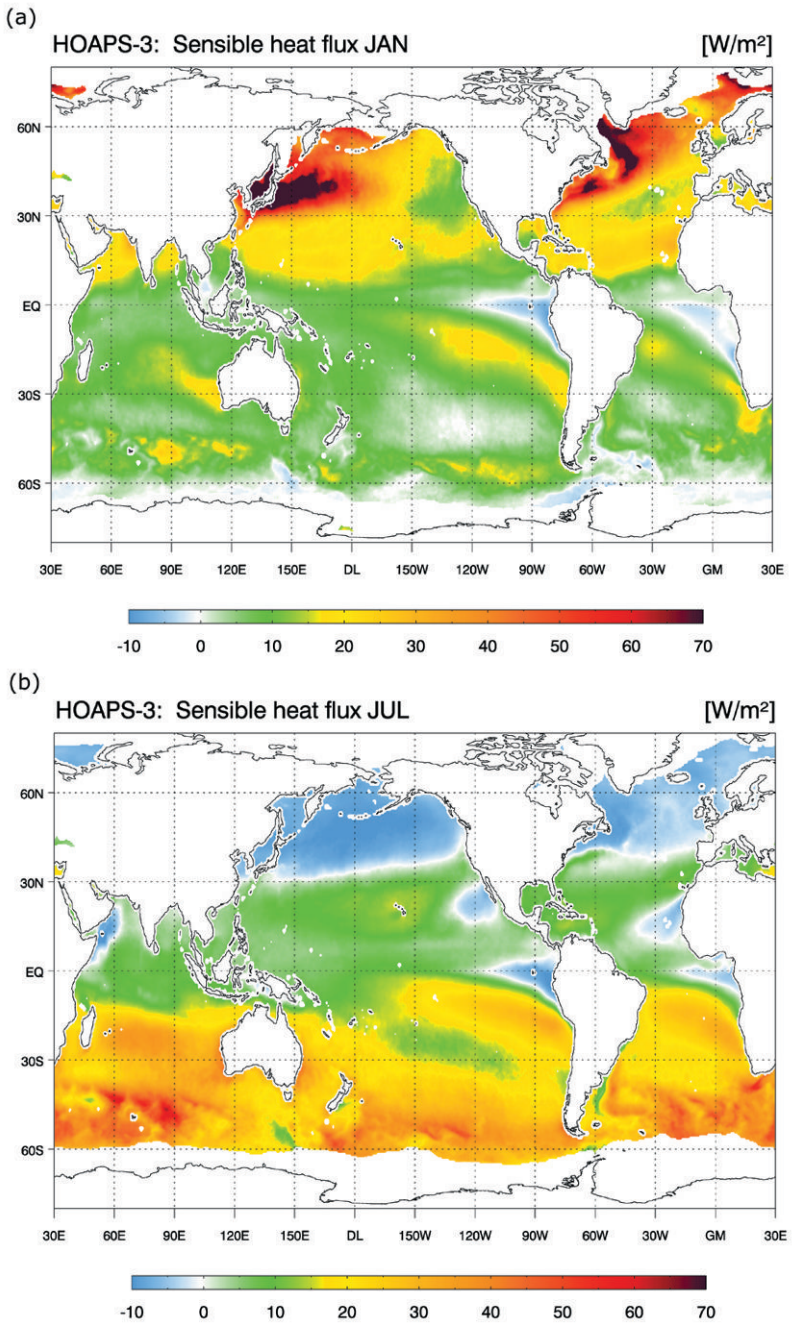


Figure 10.11. Monthly mean climatology of global sensible heat fluxes (sea to air positive) calculated on the basis of satellite data for the period 1987–2005. (a) January. (b) July (from HOAPS 3 database—Andersson *et al.*, 2007).

where operational applications are becoming feasible, because of widening spatial and temporal coverage by satellite sensors, their improving accuracy, and the nearly simultaneous availability from one or more platforms of variables that influence air–sea fluxes. Satellite data offer a promising route to improve our understanding of air–sea fluxes on a global scale, including those regions that are inaccessible to regular *in situ* measurements.

The more detailed knowledge of fluxes that is now accruing from satellite observations could have a fundamental impact on the modeling and prediction of Earth's climate, one of the major political and environmental concerns of today. As the Earth system adjusts in response to global warming, we may expect there to be shifts in the geographical distribution of ocean–atmosphere exchanges which have the potential to exert positive or negative feedbacks on climate change. We need globally detailed climatologies of air–sea fluxes so that we are able to recognize changes, however subtle, while they are occurring. There is therefore an urgency to continue the task of constructing a reliable and comprehensive satellite-based flux measurement system, in conjunction with effective *in situ* validation observations. As space agencies plan future commitments to ocean-monitoring programs, the satellite oceanography community needs to be able to present clear, evidence-based requirements for the type and density (in space and time) of ongoing measurements needed to accurately measure ocean–atmosphere fluxes.

What are the ways forward for research in this field? Although the physical and biogeochemical processes that control fluxes are still not fully understood, nor parameterized with complete confidence, the use of satellite data has opened up possibilities for new approaches that are worth further investigation. These include the incorporation of parameters describing the wave field, such as wave breaking, significant wave height, wave period, or wave age which can all, in principle, be derived from remotely sensed data. There is scope for further research using direct radar measurements of surface roughness instead of, or to supplement, wind fields, especially in cases where surface films or uncertain atmospheric stability conditions compromise the use of standard, wind-based, bulk parameterizations.

Such research will require *in situ* experiments and further measurements at sea, coincident with satellite data retrievals. Ways must be developed to combine the different benefits of field-based and remotely sensed observations. For example, where the inputs needed by parametric flux models are too patchy to deliver global flux climatologies, they may be improved by applying objective analysis (OA) or optimal interpolation (OI) techniques. These blend data from different sources (satellites, buoys, and ships) to fill at least some of the space-time sampling gaps, as demonstrated by Yu and Weller (2007) who have produced what some experts in the field consider to be the most satisfactory flux climatology to date.

Modeling studies also have an important part to play. For example, one-dimensional ocean turbulence models incorporating the COARE air–sea flux model (Fairall *et al.*, 1996, 2003) have been shown (Jeffery *et al.*, 2007; Kettle and Merchant, 2008) to explicitly resolve some processes that are parameterized in bulk flux equations. Such models not only provide a testbed for experimenting with

parameterizations, but may also generate artificial datasets from which the coefficients of new parametric flux models could be optimized.

Collaboration between remote-sensing scientists and those developing three-dimensional, numerical, coupled, ocean–atmosphere and biogeochemical models offers yet another way forward, through data assimilation. Ocean color data, scatterometer winds, sea surface height, and SST are assimilated into a variety of models for parameter and state estimation (e.g., Gregoire *et al.*, 2003; Hemmings *et al.*, 2008). Stammer *et al.* (2004) assimilated sea surface height from ERS-1 and 2 and TOPEX/Poseidon, Reynolds SST (Reynolds and Smith, 1994), and wind stress fields from ERS, NSCAT, QuikSCAT, and a range of *in situ* data. They aimed to calculate those air–sea fluxes of momentum, heat, and moisture that best produce what is observed of the time-evolving ocean state. Air–sea fluxes are used as part of the control vector that is adjusted to bring ocean models into agreement with observational data in such a way that the model describes the temporal and spatial evolution of the oceanic state in a dynamically consistent manner. This use of satellite data has only recently become feasible because of the quality and quantity of available data and the maturity of existing models. Air–sea fluxes form only a small part of the results of such studies and there is scope for further research to exploit existing long-term datasets in similar ways.

Nonetheless, we are still a long way from reliable, globally detailed, ocean–atmosphere simulations. Even when numerical modeling and assimilation are more advanced there will always remain a need for flux observations that are independent of models. There is also the question of extreme events, such as hurricanes, which are relatively infrequent in time and space but which may make a disproportionately large contribution to globally or zonally integrated fluxes. Flux parameterizations that are satisfactory for normal conditions may not apply to these extreme events and need special treatment. As discussed in Section 9.3, satellite data give access to details of tropical cyclones that are otherwise inaccessible, and so translating those data into reliable estimates of gas and heat fluxes in hurricanes presents another challenge that needs to be met.

10.6 REFERENCES

Andersson, A., S. Bakan, K. Fennig, H. Grassl, C.-P. Klepp, and J. Schulz (2007), *Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data* (HOAPS-3, monthly mean, online database), available at *http://www.hoaps.org/*, doi: 10.1594/WDCC/HOAPS3_MONTHLY6

Andreas, E. L., and J. DeCosmo (2002), The signature of sea spray in the HEXOS turbulent heat flux data. *Boundary-Layer Meteorology*, **103**, 303–333.

Barkley, M. P., P. S. Monks, and R. J. Engelen (2006), Comparison of SCIAMACHY and AIRS CO_2 measurements over North America during the summer and autumn of 2003. *Geophys. Res. Letters*, **33**(L20805), doi: 10.1029/2006GL026807.

Barkley, M. P., P. S. Monks, A. J. Hewitt, T. Machida, A. Desai, N. Vinnichenko, T. Nakazawa, M. Y. Arshinov, N. Fedoseev, and T. Watai (2007), Assessing the near

surface sensitivity of SCIAMACHY atmospheric CO_2 retrieved using (FSI) WFM-DOAS. *Atmos. Chem. Phys.*, **7**, 3597–3619.

Bentamy, A., K. B. Katsaros, A. M. Mestas-Nunez, W. M. Drennan, E. B. Forde, and H. Roquet (2003), Satellite estimates of wind speed and latent heat flux over the global oceans. *J. Climate*, **16**, 637–656.

Charnock, H. (1955), Wind stress on a water surface. *Quart. J. Roy. Meteorol. Soc.*, **81**, 639.

DeCosmo, J., K. B. Katsaros, S. D. Smith, R. J. Anderson, W. A. Oost, K. Bumke, and H. Chadwick (1996), Air–sea exchange of water vapour and sensible heat: The humidity exchange over the sea (HEXOS) results. *J. Geophys. Res.*, **101**, 12001–12016.

Donlon, C. J., T. J. Nightingale, T. Sheasby, J. Turner, I. S. Robinson, and W. J. Emery (1999), Implications of the oceanic thermal skin temperature deviation at high wind speed. *Geophys. Res. Letters*, **26**(16), 2505–2508.

Donlon, C. J., I. S. Robinson, K. S. Casey, J. Vazquez, E. Armstrong, O. Arino, C. L. Gentemann, D. May, P. Le Borgne, J.-F. Piolle *et al.* (2007), The Global Ocean Data Assimilation Experiment (GODAE) High Resolution Sea Surface Temperature Pilot Project (GHRSST-PP). *Bull. Am. Meteorol. Soc.*, **88**(8), 1197–1213, doi: 10.1175/BAMS-88-8-1197.

Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, and E. Fiedler (2010), The operational sea surface temperature and sea ice analysis (OSTIA). *Remote Sens. Environ.*, submitted to AATSR Special Issue.

Edson, J. B., C. J. Zappa, J. A. Ware, and W. R. McGillis (2004), Scalar flux relationships over the open ocean. *J. Geophys. Res.*, **109**(C08S09), doi: 10.1029/2003JC001960.

Elfouhaily, T., D. Vandemark, J. Gourrion, and B. Chapron (1998), Estimation of wind stress using dual-frequency TOPEX data. *J. Geophys. Res.*, **103**, 25101–25108.

Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young (1996), Bulk parameterization of air–sea fluxes in TOGA COARE. *J. Geophys. Res.*, **101**, 3747–3767.

Fairall, C. W., E. F. Bradley, J. E. Hare, A. A. Grachev, and J. B. Edson (2003), Bulk parameterization of air–sea fluxes: Updates and verification for the COARE algorithm. *J. Climate*, **16**(2), 571–591.

Fangohr, S., and D. K. Woolf (2007), Application of new parameterizations of gas transfer velocity and their impact on regional and global marine CO_2 budgets. *J. Marine Systems*, **66**(1/4), 195–203.

Fangohr, S., D. K. Woolf, C. D. Jeffery, and I. S. Robinson (2008), Calculating long-term global air–sea flux of carbon dioxide using scatterometer, passive microwave, and model re-analysis wind data. *J. Geophys. Res.*, **113**(C09032), doi: 10.1029/2005JC003376.

Frew, N. M., E. J. Bock, U. Schimpf, T. Hara, H. Haußecker, J. B. Edson, W. R. McGillis, S. P. McKenna, R. K. Nelson, B. M. Uz, and B. Jähne (2004), Air–sea gas transfer: Its dependence on wind stress, small-scale roughness, and surface films. *J. Geophys. Res.*, **109**(C08S17), doi: 10.1029/2003JC002131.

Frew, N. M., D. M. Glover, E. J. Bock, and S. J. McCue (2007), A new approach to estimation of global air–sea gas transfer velocity fields using dual-frequency altimeter backscatter. *J. Geophys. Res.*, **112**(C11003), doi: 10.1029/2006JC003819.

Gardashov, R. G., K. S. Shifrin, and J. K. Zolotova (1988), Emissivity, thermal albedo, and effective emissivity of the sea at different wind speeds. *Oceanologica Acta*, **11**, 121–124.

Gautier, C., P. Peterson, and C. Jones (1998), Ocean surface air temperature derived from multiple data sets and artificial neural networks. *Geophys. Res. Letters*, **25**, 4217–4220.

Gentemann, C. L., C. J. Donlon, A. R. Stuart-Menteth, and F. J. Wentz (2003), Diurnal signals in satellite sea surface temperature measurements. *Geophys. Res. Letters*, **30**(3), 1140, doi: 10.1029/2002GL016291.

Globalview (2007), *Cooperative Atmospheric Data Integration Project*. National Oceanic and Atmospheric Administration, ESRL, Boulder, CO, available at http://www.esrl.noaa.gov/gmd/ccgg/globalview/ (last accessed August 24, 2008).

Glover, D. M., N. M. Frew, S. J. McCue, and E. J. Bock (2002), A multiyear time series of global gas transfer velocity from the TOPEX dual frequency, normalized radar backscatter algorithm. In M. A. Donelan, W. M. Drennan, E. S. Saltzman, and R. Wanninkhof (Eds.), *Gas Transfer at Water Surfaces* (Geophys. Monograph 127, pp. 325–331). American Geophysical Union, Washington, D.C.

Glover, D. M., N. M. Frew, and S. J. McCue (2007), Air–sea gas transfer velocity estimates from the Jason-1 and TOPEX altimeters: Prospects for a long-term global time series. *J. Mar. Syst.*, **66**, 173–181.

Grassl, H., V. Jost, R. Kumar, J. Schulz, P. Bauer, and P. Schlüssel (2000), *The Hamburg Ocean–Atmosphere Parameters and Fluxes from Satellite Data (HOAPS): A Climatological Atlas of Satellite-derived Air-Sea-Interaction Parameters over the Oceans* (Report No. 312, ISSN 0937-1060, 132 pp.) Max Planck Institute for Meteorology, Hamburg, Germany.

Gregoire, M., P. Brasseur, and P. F. J. Lermusiaux (2003), Special issue: The use of data assimilation in coupled hydrodynamic, ecological and bio-geo-chemical models of the ocean. *J. Marine Systems*, **40/41**, 1–406.

Hemmings, J. C. P., R. M. Barciela, and M. J. Bell (2008), Ocean color data assimilation with material conservation for improving model estimates of air–sea CO_2 flux. *J. Marine Res.*, **66**, 87–126.

Jackson, D. L., G. A. Wick, and J. J. Bates (2006), Near-surface retrieval of air temperature and specific humidity using multisensor microwave satellite observations. *J. Geophys. Res.*, **111**(D10306), doi: 10.1029/2005JD006431.

Jeffery, C. D., D. K. Woolf, I. S. Robinson, and C. J. Donlon (2007), 1-d modelling of convective CO_2 exchange in the Tropical Atlantic. *Ocean Modelling*, **19**, 161–182.

Jeffery, C. D., I. S. Robinson, D. K. Woolf, and C. J. Donlon (2008), The response to phase-dependent wind stress and cloud fraction of the diurnal cycle of SST and airsea CO_2 exchange. *Ocean Modelling*, **23**, 33–48.

Jo, Y.-H., X.-H. Yan, J. Pan, W. T. Liu, and M.-X. He (2004), Sensible and latent heat flux in the tropical Pacific from satellite multi-sensor data. *Remote Sens. Environ.*, **90**, 166–177.

Jones, C., P. Peterson, and C. Gautier (1999), A new method for deriving ocean surface specific humidity and air temperature: An artificial neural network approach. *J. Applied Meteorology*, **38**, 1229–1245.

Jones, C., P. Peterson, and C. Gautier (2003), *Satellite Estimates of Air Temperature, Specific Humidity, Latent and Sensible Heat Fluxes over the Global Tropics* (Technical Report, 40 pp.). Institute for Computational Earth System Science, University of California, Santa Barbara.

Jourdan, D., and C. Gautier (1995), Comparison between global latent heat flux computed from multisensor (SSM/I and AVHRR) and from *in situ* data. *J. Atm. Ocean. Tech.*, **12**, 46–72.

Kettle, H., and C. Merchant (2005), Systematic errors in global air–sea CO_2 flux caused by temporal averaging of sea-level pressure. *Atmospheric Chemistry and Physics*, **5**, 1459–1466.

Kettle, H., and C. Merchant (2008), Modeling ocean primary production: Sensitivity to spectral resolution of attenuation and absorption of light. *Prog. Oceanogr.*, **78**, 135–146.

Kettle, H., C. J. Merchant, C. D. Jeffery, M. J. Filipiak, and C. L. Gentemann (2008), The impact of diurnal variability in sea surface temperature on the atlantic air–sea CO_2 flux. *Atmos. Chem. Phys. Discuss.*, **8**, 15825–15853.

Kraus, E. B., and J. A. Businger (1994), *Atmosphere–Ocean Interaction* (Second Edition, 362 pp.). Clarendon Press, Oxford, U.K.

Large, W. G., and S. Pond (1981), Open ocean momentum flux measurements in moderate to strong winds. *J. Phys. Oceanogr.*, **11**, 324–336.

Large, W. G., and S. Pond (1982), Sensible and latent heat flux measurements over the ocean. *J. Phys. Oceanogr.*, **12**, 464–482.

Lefèvre, N., and A. Taylor (2002), Estimating pCO_2 from sea surface temperatures in the Atlantic gyres. *Deep-Sea Res. I*, **49**, 539–554.

Lefèvre, N., A. J. Watson, and A. R. Watson (2005), A comparison of multiple regression and neural network techniques for mapping in situ pCO_2 data. *Tellus*, **57B**, 375–384.

Liss, P., and L. Merlivat (1986), Air–sea gas exchange rates: Introduction and synthesis. In: P. Buat-Menard (Ed.), *The Role of Air–Sea Gas Exchange in Geochemical Cycling* (pp. 113–129). Kluwer Academic Publishers, Dordrecht, The Netherlands.

Liu, W. T. (1985), Statistical relationship between monthly mean precipitable water and surface-level humidity over global oceans. *Mon. Weather Rev.*, **114**, 1592–1602.

Liu, W. T., K. B. Katsaros, and J. A. Businger (1979), Bulk parameterization of air–sea exchanges of heat and water vapor including the molecular constraints at the interface. *J. Atmos. Sci.*, **36**, 1722–1735.

Monin, A. S. and A. M. Obukhov (1954), Basic laws of turbulent mixing in the ground layer of the atmosphere. *Akad. Nauk. SSSR Geofiz. Inst.*, **151**, 163–187.

Nightingale, P., G. Malin, C. Law, A. Watson, P. Liss, M. Liddicoat, J. Boutin, and R. Upstill-Goddard (2000), In situ evaluation of air–sea gas exchange parameterizations using novel conservative and volatile tracers. *Global Biogeochem. Cycles*, **14**, 373–387.

Nightingale, P. D., and P. S. Liss (2004), Gases in seawater. In: H. Elderfield (Ed.), *Treatise on Geochemistry, Vol. 6: The Oceans and Marine Geochemistry* (pp. 49–81). Elsevier Science.

Oost, W. A., C. M. J. Jacobs, and C. van Oort (2000), Stability effects on heat and moisture fluxes at sea. *Boundary-Layer Meteorology*, **95**, 271–302.

Oost, W. A., G. J. Komen, C. M. J. Jacobs, and C. van Oort (2002), New evidence for a relation between wind stress and wave age from measurements during ASGAMAGE. *Boundary-Layer Meteorology*, **103**, 409–438.

Pan, J., X.-H. Yan, Y.-H. Jo, Q. Zheng, and W. T. Liu (2004), A new method for estimation of the sensible heat flux under unstable conditions using satellite wind vectors. *J. Phys. Oceanogr.*, **34**, 968–977.

Pinker, R. T., and I. Laszlo (1992), Modeling surface solar irradiance for satellite applications on a global scale. *J. Applied Meteorology*, **31**, 194–211.

Reynolds, R. W., and T. S. Smith (1994), Improved global sea surface temperature analyses. *J. Climate*, **7**, 928–948.

Robinson, I. S., and C. J. Donlon (2003), Global measurement of sea surface temperature from space: Some new perspectives. *J. Atm. Ocean Sci.* (previously *The Global Atmosphere and Ocean System*), **9**(1), 19–37.

Schanz, L., and P. Schlüssel (1997), Atmospheric back radiation in the tropical Pacific: Intercomparison of in-situ measurements, simulations, and satellite retrievals. *Meteor. Atmos. Phys.*, **63**, 217–226.

Schulz, J., P. Schlüssel, and H. Grassl (1993), Water vapor in the atmospheric boundary layer over oceans from SSM/I measurements. *Int. J. Remote Sensing*, **14**, 2773–2789.

Schulz, J., J. Meywerk, S. Ewald, and P. Schlüssel (1997), Evaluation of satellite derived latent heat fluxes. *J. Climate*, **10**, 2782–2795.

Singh, R., B. Simon, and P. C. Joshi (2003), A technique for direct retrieval of surface specific humidity over oceans from IRS/MSMR satellite data. *Boundary-Layer Meteorology*, **106**, 547–559.

Singh, R., P. C. Joshi, and C. M. Kishtawal (2006), A new method to determine near surface air temperature from satellite observations. *Int. J. Remote Sensing*, **27**(14), 2831–2845.

Stammer, D., K. Ueyoshi, A. Köhl, W. G. Large, S. A. Josey, and C. Wunsch (2004), Estimating air–sea fluxes of heat, freshwater, and momentum through global ocean data assimilation. *J. Geophys. Res.*, **109**(C05023), doi: 10.1029/2003JC002082.

Stuart-Menteth, A. R., I. S. Robinson, and P. C. Challenor (2003), A global study of diurnal warming using satellite derived sea surface temperature. *J. Geophys. Res.*, **108**(C5), 3155, doi: 3110.1029/20025C001534.

Takahashi, T., S. C. Sutherland, C. Sweeney, A. Poisson, N. Metzl, B. Tilbrook, N. Bates, R. Wanninkhof, R. Feely, C. Sabine *et al.* (2002), Global sea–air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects. *Deep-Sea Res. II*, **49**, 1601–1622.

Taylor, P. K., and M. A. Yelland (2000), On the apparent "imbalance" term in the turbulent kinetic energy budget. *J. Atm. Ocean. Tech.*, **17**, 82–89.

Taylor, P. K., and M. A. Yelland (2001), The dependence of sea surface roughness on the height and steepness of the waves. *J. Phys. Oceanogr.*, **31**, 572–590.

Wanninkhof, R. (1992), Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res.*, **97**, 7373–7382.

Wanninkhof, R. (2002), The effect of using averaged winds on global air–sea CO_2 fluxes. *Gas Transfer at Water Surfaces* (Geophysical Monograph 127, pp. 351–356). American Geophysical Union, Washington, D.C.

Wanninkhof, R., and W. McGillis (1999), A cubic relationship between air–sea CO_2 exchange and wind speed. *Geophys. Res. Letters*, **26**(13), 1889–1892.

Ward, B., R. Wanninkhof, W. R. McGillis, A. T. Jessup, M. D. DeGrandpre, J. E. Hare, and J. B. Edson (2004), Biases in the air–sea flux of CO_2 resulting from ocean surface temperature gradients. *J. Geophys. Res.*, **109**(C08S08), doi: 10.1029/2003JC001800.

Weiss, R. (1974), Carbon dioxide in water and seawater: The solubility of a non-ideal gas. *Marine Chemistry*, **2**, 203–215.

Woolf, D. K. (2005), Parameterization of gas transfer velocities and sea-state dependent wave breaking. *Tellus*, **57B**, 87–94.

Xue, Y., D. T. Llewellyn-Jones, S. P. Lawrence, and C. T. Mutlow (2000), On the Earth's surface energy exchange determination from ERS satellite ATSR data, Part 3: Turbulent heat flux on open sea. *Int. J. Remote Sensing*, **21**(18), 3427–3444.

Yelland, M. A., and P. K. Taylor (1996), Wind stress measurements from the open ocean. *J. Phys. Oceanogr.*, **26**, 541–558.

Yelland, M. A., B. I. Moat, P. K. Taylor, R. W. Pascal, J. Hutchings, and V. C. Cornell (1998), Measurements of the open ocean drag coefficient corrected for airflow disturbance by the ship. *J. Phys. Oceanogr.*, **28**, 1511–1526.

Yu, L., and R. A. Weller (2007), Objectively analyzed air–sea heat fluxes for the global ice-free oceans (1981–2005). *Bull. Am. Meteorol. Soc.*, **4**, 527–539, doi: 10.1175/BAMS-88-4-527.

Zong, H., Y. Liu, Z. Rong, and Y. Cheng (2007), Retrieval of sea surface specific humidity based on AMSR-E satellite data. *Deep-Sea Res. I*, **54**(7), 1189–1195.

11

Large ocean phenomena with human impact

11.1 INTRODUCTION

The goal throughout this book is not simply to present the variety of applications of remote sensing in marine science, but particularly to identify unique contributions to revealing and understanding ocean phenomena made possible by the observational capabilities of sensors viewing the sea from orbiting satellites. In some cases the discovery of new scientific knowledge is mainly of academic interest, but in others it reaches more widely into human society and its application can make an impact on people's daily lives. This chapter presents examples of such applications. It examines a few separate phenomena that have two things in common: first, they affect the safety and economic wellbeing of significant numbers of people as well as the health of the global environment; second, they rely on, or have the potential to benefit from, satellite oceanography methods to obtain data used for warnings, forecasts, and improved understanding of the phenomenon.

Because of their relevance to the everyday lives of millions of people, the topics discussed here represent part of the cutting edge where science brings benefits to human society. The obligation for governments to grapple with the consequences of these environmental phenomena is part of what has motivated world leaders to agree that a comprehensive Earth observation program, including satellite-based observations, is an essential international "common good".[1] The purpose of this chapter is

[1] The 2002 World Summit on Sustainable Development in Johannesburg highlighted the urgent need for co-ordinated observations relating to the state of the Earth. The Group of Eight Summit in June 2003 in Evian, France, affirmed the importance of Earth observation as a priority activity. The First Earth Observation Summit convened in Washington, D.C., in July 2003 adopted a Declaration establishing the *ad hoc* intergovernmental Group on Earth Observations (*ad hoc* GEO) to draft a 10-Year Implementation Plan. This led to the establishment of the Global Earth Observation System of Systems. To see how this is now being implemented worldwide go to *http://www.earthobservations.org/index.html*

to inform readers about the crucial role that ocean-observing sensors on satellites have to play in monitoring, possibly forecasting, and managing the human impact of some large-scale natural phenomena.

The first topic to be presented is the El Niño–Southern Oscillation (ENSO) phenomenon. This natural but irregular perturbation of weather patterns and the ocean's coupled response in the equatorial Pacific impacts global climate and has caused problems to mankind for centuries, but only in the last two decades has it attracted popular public attention. When an El Niño event occurs, it serves as a cause on which to pin the blame for many ills, irrespective of scientific justification, and for a time becomes a favorite demon of the news media in many nations! In reality it is a complex process involving the coupled and lagged response between ocean and atmosphere. Arguably the capacity of satellite-derived data to visualize the phenomenon has helped to bring it to the public attention, but more constructively satellite oceanography has brought valuable new tools which can monitor the phenomenon and may lead to better prediction.

The subject next addressed is another atmosphere-driven phenomenon which occurs in a more seasonally regular way than El Niño events. This is the monsoon and in particular the ocean's response to monsoon winds. The daily lives of many hundreds of millions of people are affected by seasonal changes of winds in tropical zones. The ocean also is affected, and satellites enable regular monitoring which allows changes to be detected in the fundamental patterns of ocean circulation, hydrography, and primary production from one season to the next. The capacity to understand, or even to predict, interannual variability in the ocean's monsoon response around the world is of considerable importance to those who live and work in the affected areas.

Section 11.4 addresses a distinctly cooler phenomenon, moving from equatorial to polar regions to consider the topic of how sea ice is distributed. This also is of importance to those whose livelihood takes them into ice-covered waters. In the past there were pockets of rich but isolated knowledge based on localized experience of annual variation of sea ice extent, but only with the advent of satellite-based detection of sea ice has it become possible to tell the global story for each polar region. As a subject of wide geographical extent, which is almost impossible to study without the perspective provided by satellite data, it makes an appropriate topic for this chapter.

The final substantive section examines how satellite measurements contribute to our detailed knowledge of sea surface height and its variability. This touches on mapping astronomical tides throughout the World Ocean, but focuses on sea level and how changes and trends are detected by altimetry. It shows how remote sensing can detect storm surges when sea level rapidly rises or falls with possibly calamitous consequences. Tsunamis are also considered, following suggestions that Earth observation technology could be harnessed to mitigate their impact.

Other topics which could also belong in this chapter because of their human impact, but which fitted naturally into other chapters, include hurricanes (Chapter 9) and harmful algal blooms (Chapter 13). Chapter 14 will discuss the role of satellite data in underpinning new, operational, ocean-monitoring-and-forecasting systems.

11.2 EL NIÑO

11.2.1 The ENSO phenomenon

Overview

The El Niño–Southern Oscillation (ENSO) phenomenon is the name given to particular patterns of coupled ocean–atmosphere behaviou-r in and above the tropical Pacific, and linked also to wider atmospheric circulation over the South Pacific. The significant factor which gives this topic worldwide human importance is the way that different phases of the phenomenon are accompanied by very different weather patterns. From time to time, atmospheric circulation and coupled ocean currents depart from their ''normal'' climatic state, flipping into an atypical, but quasistable, pattern of behavior that can persist for several months. One particular form of this, referred to as an El Niño event, suppresses the East Pacific equatorial upwelling, with disastrous consequences for regional fisheries. Even more significant are the associated changes to weather patterns over North and Central America and equatorial East Asia.

Such events occur apparently randomly, with varying strength, at intervals of between 3 and 7 years. The sometimes very strong deviation from typical values of meteorological conditions such as winds and rainfall can have devastating human consequences. Storms and heavy rainfall cause flooding in some regions while drought is experienced elsewhere. Agriculture suffers when crops are lost to flood and drought. Insurance claims confirm widespread damage to buildings and infrastructure. The difficulty of predicting when the phenomenon will occur compounds the problems since the occurrence of an El Niño event is made more memorable when atypical weather events take people by surprise. Moreover, because of coupling to the Southern Oscillation, the consequences of the El Niño phenomenon can be detected in climatological records of many parts of the world.

There is therefore a wide interest in monitoring, understanding, and, if possible, forecasting the state of the complex atmosphere–ocean feedback loop in the tropical Pacific. If the onset of events could be confidently predicted it would be possible for the countries directly affected to plan action in advance to mitigate the worst human impacts of changed weather patterns. Throughout the rest of the world, allowance could also be made for the more diffuse climatological changes that seem to attend a strong El Niño event. Table 11.1 summarizes some of the distinctive changes of weather patterns associated with an El Niño event that have been observed around the world. A more complete view of the global impact of the ENSO phenomenon and the scientific processes which couple the oceanic and atmospheric behavior can be found in books dedicated to the topic (e.g., Philander, 1990; Allen *et al.*, 1996, 2000).

Our purpose here is not to explain in detail the phenomenon itself or its consequences, but to pay attention to how the methods of satellite oceanography can reveal more clearly the evolution of an ENSO event and particularly the changes that take place in the surface ocean. However, in order that readers unfamiliar with the topic can appreciate the contribution made by remote sensing to observing an El

Table 11.1. Changes in meteorological conditions around the world apparently affected by the occurrence of the "Warm Episode", or El Niño phase, of the ENSO phenomenon during boreal winter (December–February) and summer (June–August).

Region	*December–February*	*June–August*
Rainfall		
Northeastern Australasia, Indonesia, Philippines	Abnormally dry	Abnormally dry
Equatorial zone 170°E–160°W	Wet	Wet
Ecuador, Peru coast	Wetter than normal	
Gulf coast of U.S.A.	Wetter than normal	
Mexico, Central America		Drier than normal
Northern Brazil	Drier than normal	
Subtropical South America	Wet (east coast)	Wet
Northeastern India		Reduced monsoon rainfall
Equatorial East Africa	Wet	
Southeast Africa	Drier than normal	
Air temperature		
Indonesia, South East Asia	Warm	
Equatorial zone 80°–180°W	Warm	
Japan	Warm	
Southeast Australia	Warm	
Alaska, West Canada	Warm	
Maritime East Canada	Warm	
South Brazil	Warm	Warm
Southeast Africa	Warm	
Mexico, Central America		Warm
Oceania		Cool

Niño event, the following simple review may be useful. It first describes the physical changes that occur during the cycle, and then outlines how the phenomenon has been detected and monitored using conventional meteorological and *in situ* ocean-observing instruments.

Description of the ENSO cycle

Under normal conditions, when the weather pattern over the equatorial Pacific corresponds to conditions that are encountered most often, the dominant wind along the Equator blows towards the west (referred to as an easterly wind since it comes from the east). The effect of this wind on the ocean is to force a westward current along the Equator, at which latitude the Coriolis effect is zero. Just north of the Equator the Coriolis parameter is positive, and the wind induces a northward Ekman transport. South of the Equator, Ekman transport is southwards (as explained in Section 5.1.1). Consequently strong upwelling is induced along the Equator, drawing up deeper, cooler water into the ocean mixed layer and raising the thermocline closer to the surface towards the east of the ocean (as illustrated in Figure 11.1a). Thus whereas SST might be expected to be highest along the Equator because solar heating is strong, a cool tongue of upwelled water occurs at the east and spreads halfway across the ocean, as shown in near-surface temperature maps of the tropical Pacific (Figure 11.2a) and longitude–depth temperature sections along the Equator (Figure 11.3a).

The upwelling does not reach the western side of the Pacific Ocean. At between 150°E and 160°E the upper layer of the sea, heated by the Sun and not cooled by upwelling, is some 6°C to 7°C warmer than in the east. This is the so-called "Warm Pool" where the high ocean temperature produces strong atmospheric convection and low atmospheric sea level pressure that drives a convective loop of air circulation. Air flows to the east at high altitude, sinks over the American continent and returns as easterly winds at sea level which drive upwelling, thus maintaining a stable pattern of air–sea interaction (see Figure 11.1a). Under these normal conditions, upwelling provides a rich supply of nutrients to the eastern equatorial Pacific which maintains an abundant fishery. The high pressure results in low rainfall over central America and Ecuador, while on the western side of the ocean strong convection produces considerable rainfall over East Asia.

Normal conditions prevail as long as equatorial easterlies are able to maintain upwelling. However, the equatorial convective loop is embedded in the wider atmospheric circulation over the Pacific. If the wider pressure distribution is tending to produce stronger-than-normal westerly winds, these may reduce both equatorial easterlies and the related upwelling, allowing the Warm Pool to move towards the east. If this effect is strong enough, it can create a different pattern (shown in Figures 11.1b, 11.2b, and 11.3b), which persists in a quasistable state for several months. This is the El Niño condition.

Here upwelling has effectively switched off and the thermocline remains deep towards the American coast. The Warm Pool, and with it the main equatorial convection, has moved much farther east to 180°W–160°W, bringing higher rainfall.

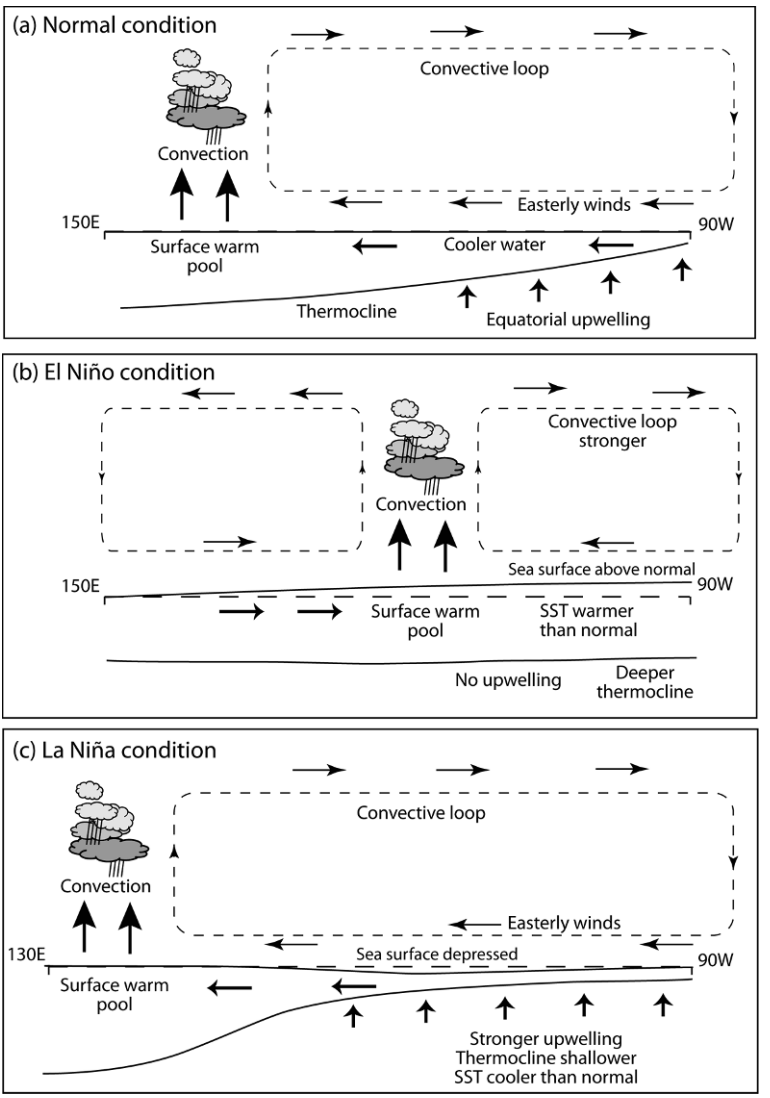


Figure 11.1. Equatorial section from west to east across the Pacific Ocean, schematically outlining air–sea interaction patterns, and the corresponding response of the upper ocean during (a) normal conditions, (b) an El Niño event, (c) a La Niña event.

Rainfall decreases over East Asia and increases over the American continent, causing drought with consequent wildfires in Australia and East Asia, and flooding in Ecuador and California. Removal of the equatorial upwelling severely reduces primary production, and consequently local fisheries fail. It was in the 1800s that Peruvian fishermen named the unusually warm current "El Niño" (the Christ child) because it tended to arrive in December, close to Christmas. They learned from experience that it heralded a collapse of their fishery for the next season.

There is a third state that can occur, which presents conditions almost at the other extreme from an El Niño. Shown in Figures 11.1c, 11.2c, and 11.3c, it is called

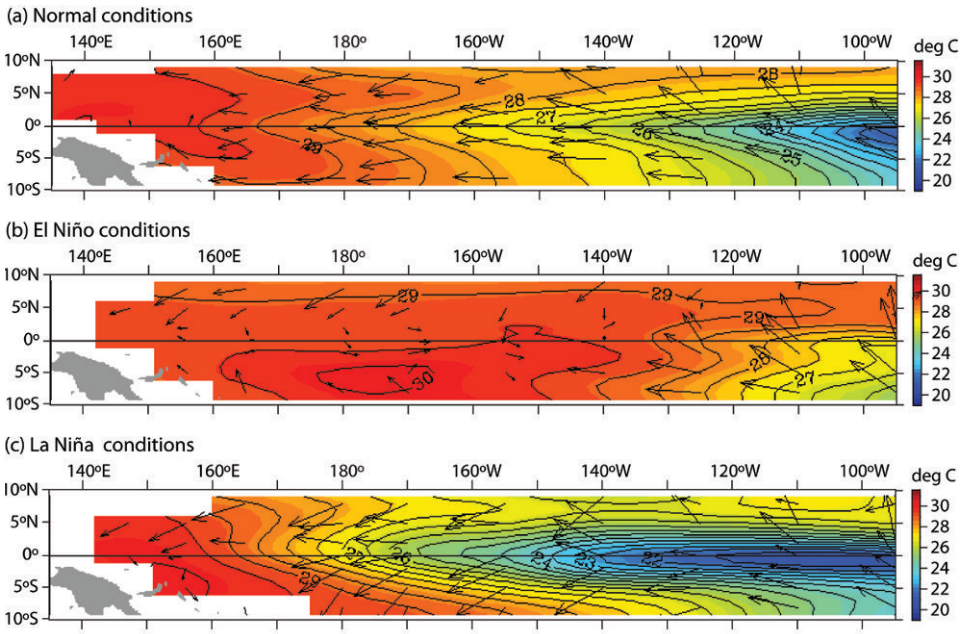


Figure 11.2. Typical maps of monthly mean, near-surface temperature in the tropical Pacific Ocean for different phases of the El Niño cycle. The arrows indicate the direction of monthly mean surface winds. (a) July 1996; normal conditions. (b) November 1997; El Niño conditions. (c) December 1998; La Niña conditions. These data are based on measurements from the TAO/TRITON array of moored buoys (figure based on maps obtained from the data display pages of the website of the Tropical Atmosphere Ocean (TAO) Project, available at *http://www.pmel.noaa.gov/tao/index.shtml)*.

"La Niña". Here westward-blowing winds are stronger than normal, pushing the Warm Pool farther west to 130°E–140°E. This causes increased precipitation over Indonesia. The equatorial upwelling is more vigorous than normal, particularly between 110°W and 170°W enriching primary production and enhancing East Pacific fisheries. The cool equatorial SST also creates conditions under which tropical instability waves can develop (see Section 6.6.2).

Although the temperature distribution maps shown in Figure 11.2 are quite complex, the distribution of the SST anomaly, relative to seasonal climatology defined over a decade or longer, characterizes very clearly the difference between El Niño, La Niña, and normal conditions (as shown in Figure 11.4). El Niño and La Niña SST anomaly signatures are, respectively, tongues of warmer or cooler water spreading from the east to about the dateline, meridionally symmetrical about the Equator. This gives a strong indication that enhanced wind-driven equatorial upwelling, or the lack of it, is the driving factor for ocean temperature behavior during La Niña and El Niño events, respectively.

The SST anomaly can therefore be used to provide a simplified record of the status of El Niño/La Niña perturbation of the ocean at any time. Figure 11.5a shows

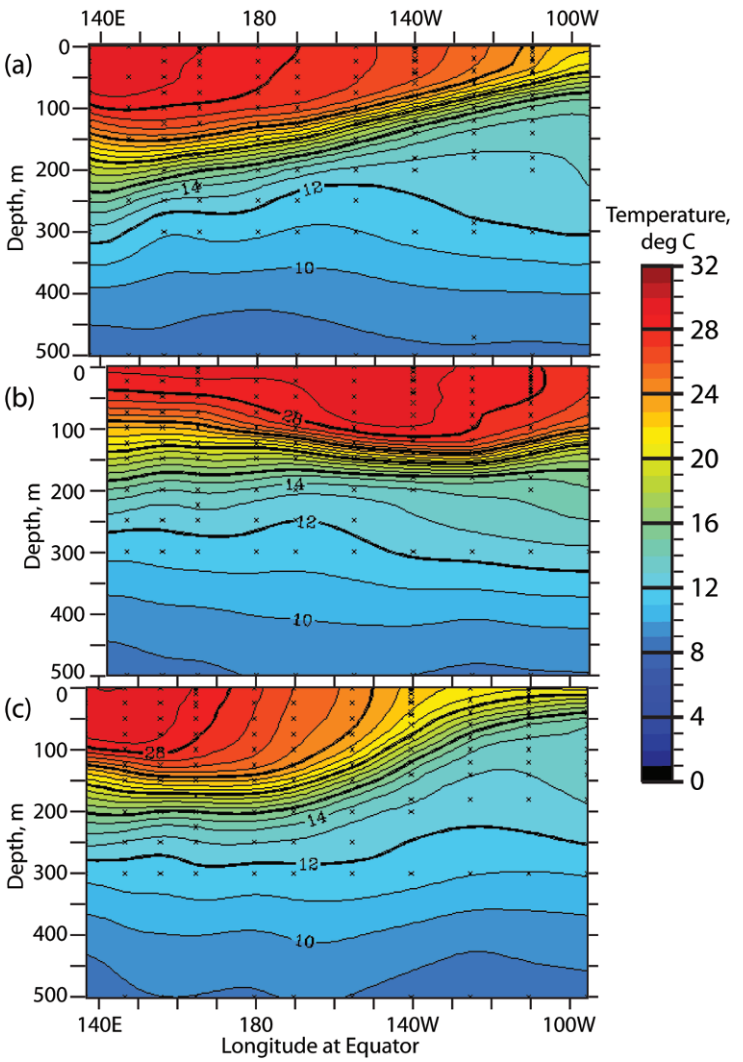


Figure 11.3. Typical longitudinal sections of monthly mean temperature distribution with depth along the Equator in the Pacific Ocean for different phases of the El Niño cycle. (a) July 1996; normal conditions. (b) November 1997; El Niño conditions. (c) December 1998; La Niña conditions. These data are based on measurements from the TAO/TRITON array of moored buoys (figure based on maps obtained from the data display pages of the website of the Tropical Atmosphere Ocean (TAO) Project, available at *http://www.pmel.noaa.gov/tao/index.shtml*).

the time series of the Ocean Niño Index (ONI) from 1950 to the present. The ONI consists simply of monthly samples of SST anomalies averaged spatially over the region between 5°N and 5°N and 120°W and 170°W. This is the so-called "Niño-3/4" region, one of several regions used for defining different indices. The

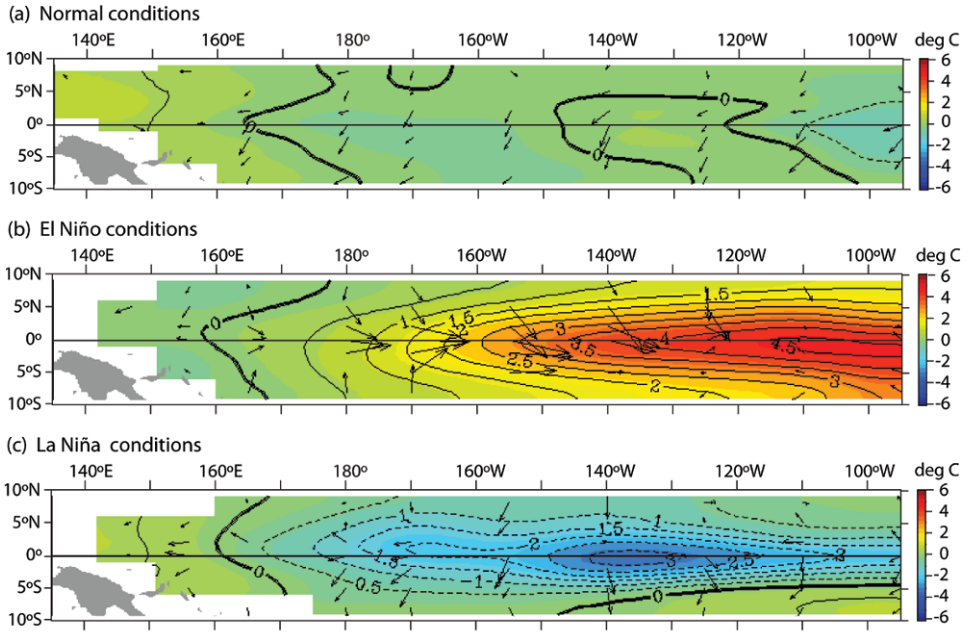


Figure 11.4. Maps of the near-surface temperature anomaly corresponding to monthly mean, near-surface temperature for different phases of the El Niño cycle. The arrows indicate the direction of the monthly mean, wind anomaly field. (a) July 1996; normal conditions. (b) November 1997; El Niño conditions. (c) December 1998; La Niña conditions. These data are based on measurements from the TAO/TRITON array of moored buoys (figure is based on maps obtained from the data display pages of the website of the Tropical Atmosphere Ocean (TAO) Project, available at *http://www.pmel.noaa.gov/tao/index.shtml*).

ONI is a 3-month running mean of the Extended Reconstructed Sea Surface Temperature (ERSST) version 3 dataset produced by the U.S. National Climate Data Center. This global dataset uses *in situ* temperature measurements aggregated onto a 2° latitude $\times$ 2° longitude grid. It dates back to 1854 and is adjusted for long-period variability (Xue *et al.*, 2003) although since 1985 ERSST has also included bias-adjusted satellite SST retrieved from AVHRR. The anomaly used for ONI is based on climatology derived over the base period 1971–2000.

In Figure 11.5a the peaks are colored red when the anomaly is greater than 1.0°C, corresponding to strong El Niño events, and blue for strong La Niña events when the anomaly is less than -1.0°C. The figure shows considerable fluctuation of the index, with occasional events (two or three times per decade), when the anomaly amplitude is greater than 1°C, either positive or negative. Once the ocean temperature switches into an El Niño or La Niña state, it appears to persist for several months, taking at least a year to reach its peak before it returns to more normal levels. There is no immediately obvious repetition of a sequence of events in the ONI, on which a simple predictive capacity might be built. However, it is very

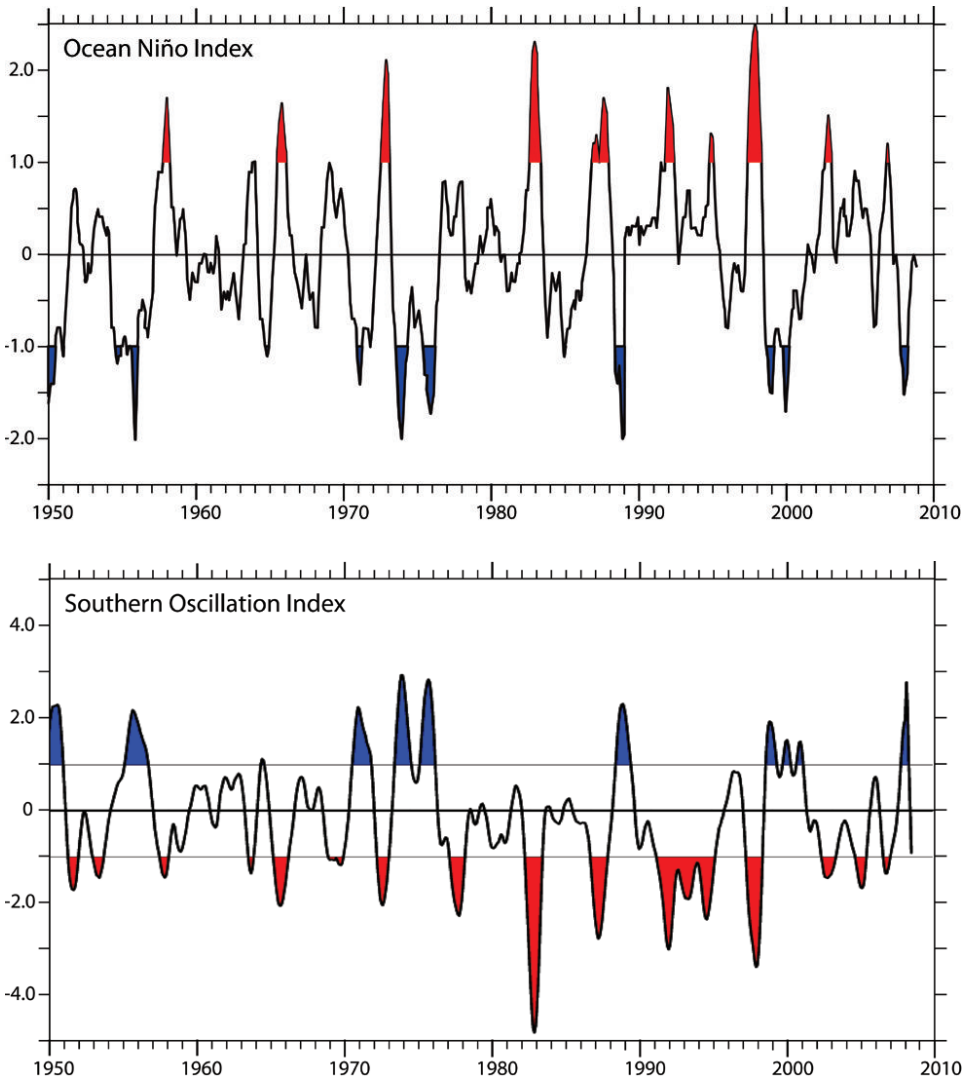


Figure 11.5. Time series of ENSO indicators, 1950 to present. (a) Ocean Niño Index (ONI). (b) Southern Oscillation Index (SOI). Monthly indices were obtained for (a) from NOAA's Climate Prediction Center (*http://www.cpc.noaa.gov*) and data for (b) came from the Climate and Global Dynamics Group of the U.S. National Center for Atmospheric Research (*http://www.cgd. ucar.edu/cas/catalog/climind/soi.html*).

useful to have a readily measurable index for El Niño and La Niña events since once particular events have been identified it allows each of them to be studied individually to determine whether the ocean processes and case history of each are similar. It also encourages the search of available records to discover whether each event is attended by similar meteorological anomalies such as those listed in Table 11.1.

Meteorologists studying El Niño events realized several decades ago that their occurrence was usually correlated with the Southern Oscillation Index (SOI) which is a broad measure of southeast trade winds over the south equatorial Pacific and is used to characterize the large-scale behavior of the atmosphere in the southern hemisphere. It is defined as the normalized difference in surface pressure anomaly between Tahiti (at 150°W, 18°S) and Darwin, Australia which is 80° to the west (at 130°E, 12°S). When it is positive this corresponds to higher-than-average pressure at Tahiti and/or lower-than-average pressure at Darwin, and *vice versa*. Figure 11.5b shows the SOI from 1950 to the present. When the smoothed time series of SOI has a persistently large negative value for several months, implying weaker-than-normal southeast trade winds, in most cases it can be associated with matching El Niño conditions in Figure 11.5a. The same is true for positive peaks and La Niña occurrences. To highlight this in Figure 11.5b the peaks of the index with an amplitude greater than 1 have been colored accordingly, red when negative and blue when positive.

Because of the inverse relationship between ONI and SOI, an El Niño event, which was originally thought of as a local phenomenon relevant to the eastern equatorial Pacific, was linked to the SOI and became known as the El Niño–Southern Oscillation phenomenon, or ENSO. This terminology better indicates that we are dealing with a coupled ocean–atmosphere phenomenon, whose local oceanographic characteristics remain crucially important for East Pacific fisheries, but with a much wider, possibly global impact on weather patterns and climate.

Taken together these time series demonstrate the randomness in the way the episodes occur. There is not even a regular sequence: sometimes El Niño episodes follow each other without an intervening La Niña event. On average there is an El Niño every 3–4 years, but between 1972 and 1982 there was a gap of almost 10 years with only very minor El Niño events, although two strong La Niña episodes occurred during that interval. At the time of writing it has been 10 years since the last major El Niño in 1998. In 2003 and 2007 the signs pointed strongly towards a developing El Niño but in each case the temperature anomaly suddenly flipped back towards normal before serious warming occurred. However, in general the two independent indices do agree remarkably well in identifying events, although their relative magnitude between indices, and the precise timing of peaks (which index leads the other) is also variable. Short-scale variability in these indices, even though they are smoothed, makes it very difficult to be able to predict the future path of the curve from the present or recent trend. Thus by itself these measures are good for recording the events, but have less value in forecasting them, or being able to relate them to other oceanographic processes.

Monitoring ENSO

When an El Niño event in 1982–1983 led to insurance claims of billions of dollars in the U.S.A., stemming largely from the impact of unexpected weather patterns, its economic impact became evident to government agencies. This motivated an ambitious program to set up observational systems for monitoring the phenomenon with

the objective of providing early warning of its onset, and the longer term goal of forecasting it through the use of coupled ocean–atmosphere models. As a consequence the Tropical Ocean–Global Atmosphere (TOGA) international research program (see, e.g., McPhaden *et al.*, 1998) was established and ran from 1984 to 1995. This installed an *in situ* monitoring system for measuring subsurface ocean temperatures and salinity as well as meteorological data, using a set of 70 moorings spread across the tropical Pacific between 8°N and 8°S.

This was called the Tropical Atmosphere–Ocean (TAO) array (Hayes *et al.*, 1991). Completed in 1994 and managed by the U.S. Pacific Marine Environmental Laboratory (PMEL), the TAO array continues to report data in real time, from which interpretations are drawn concerning the status of the tropical Pacific in relation to the ENSO cycle. In January 2000, the monitoring system was renamed TAO/TRITON as the Japan Agency for Marine–Earth Science and Technology (JAMSTEC) contributed its TRITON moorings in the western portion of the array. The array was designed to monitor changes in thermocline depth, equatorial upwelling, and movement of the Warm Pool which allow the different phases of El Niño behavior to be recognized. Thus most of the temperature data on which Figures 11.2, 11.3, and 11.4 are based come from the TAO array, while the density of SST sampling contributing to El Niño temperature indices such as Figure 11.5a has been much greater since 1994.

The El Niño episodes of 1982–1983 and 1997–1998 are the largest documented since any measurements began. Yet it is worth noting that the former caught the world by surprise, and only when its impact was already being felt strongly were oceanographers sure that it was an El Niño episode. In contrast, by 1997–1998, the *in situ* monitoring array was in place and had already detected two small El Niño events in 1992 and 1995. Thus the major El Niño event of 1997 followed by twin La Niña events in 1998 and 1999 were observed in much richer detail than any previous events, leading to increased knowledge and better scientific understanding of the phenomenon (McPhaden, 1999). Moreover forecasting models were able to use the observations (Ji and Leetma, 1997) and made some fairly good forecasts of the 1997 El Niño onset (Anderson and Davey, 1998; Barnston *et al.*, 1999) although not of its intensity. In addition, and of special interest for this chapter, the 1997–1998 ENSO events were also monitored by a number of Earth-orbiting sensors, which demonstrated their capacity to capture the whole sequence of El Niño phases, from preconditioning through onset, evolution, decay, and, in this case, transition into a La Niña (Picaut *et al.*, 2002). We now focus our attention on satellite observations.

11.2.2 Observing an El Niño from satellites

Several remote-sensing methods are very well suited for observing the ENSO phenomenon, and can strongly complement the array of *in situ* sensors that are now well established for operational service. These are sea surface temperature (SST) measurements from infrared and microwave sensors, retrieval of sea surface height anomalies (SSHA) using satellite altimetry, detection of primary production through estimates of chlorophyll concentration derived from satellite ocean color

sensors, the observation of surface winds using scatterometry and microwave radiometry, and measurement of rainfall over the ocean. An outline of these individual methods is given in Chapter 2, while a more thorough account can be found in *MTOFS* (Robinson, 2004). Before showing what each method brings in particular, it is instructive to point out generic reasons why we should expect satellite oceanography to make a strong contribution to the study of ENSO.

First, space-time sampling from remote sensors is appropriately matched to the scales of the El Niño phenomenon. Weekly measurements are sufficiently frequent to capture time evolution, while sampling two-dimensional surface fields with a resolution of 0.25° latitude $\times$ 0.25° longitude is adequate for mapping dynamical structures. Note that with a spatial resolution of just 1°, a remote-sensing sensor would achieve 2,400 samples within the region 150°E–90°W, 10°S–10°N, whereas the *in situ* measurement array contains about 70 moorings. Clearly satellite data can add a more complete and detailed spatial view to buoy measurements, and this may be particularly important for the supply of data for assimilation into forecasting models. Nonetheless it is important to emphasize that the moorings are vital for the many essential subsurface and atmospheric measurements made at each location. Intercomparison of satellite and buoy measurements of variables such as SST allow for mutual quality control of the data. There should therefore be no question of satellites replacing buoys, but maximum advantage needs to be taken of the complementarity between the two measurement methods.

Second, remote-sensing measurements, when repeated regularly for many years using a sustained monitoring system, allow for the construction of mean climatologies, from which anomaly maps can be plotted in near-real time (as discussed in Section 6.2.1). The capacity to do this is very useful for a phenomenon such as El Niño which has a timescale of the same order as the seasonal (annual) cycle, from which it needs to be separated if its evolving structure is not to be hidden by, or confused with, normal seasonal variability. This is particularly the case for SST as shown in Section 11.2.3.

Third, satellite data provide a view not only of the equatorial Pacific region where the main El Niño drama takes place, but also of the whole globe through which the impact may be felt. ENSO has been described as "the major disturbing factor of the Earth's climate on seasonal to interannual timescales" (Picaut *et al.*, 2002). Satellite data allow us very easily to see how the anomalies of SST and SSHA, which depict so clearly the main event in the equatorial Pacific, are behaving elsewhere in the world, both before, during, and after it. They provide the potential for analytical searches for correlations between conditions in other oceans and what is happening in the El Niño–La Niña arena.

The following subsections individually explore the ways in which different ocean remote-sensing techniques are applied to observe ENSO events, concluding by considering the synergy that arises when they are combined. Many of the examples mentioned below come from data acquired during the 1997–1998 ENSO episode. However, it should be remembered that not all today's technological capabilities were available at that time. For example, the first microwave radiometer capable of reliable temperature measurements, TMI, was not launched until late 1997, and

global SST data from AMSR-E were not produced until 2001. The major workhorse for ocean color, SeaWiFS, was operational only from September 1997, and so did not witness the complete history of that El Niño.

11.2.3 Observing an El Niño in sea surface temperature from satellites

The El Niño–La Niña phenomenon is most clearly revealed by the temperature of the upper ocean in the equatorial band between 5°N and 5°S, and in particular its variation with longitude, as first demonstrated by Legeckis (1986) for the 1982–1983 El Niño. Figure 11.6 shows monthly composites of SST measured by AVHRR.

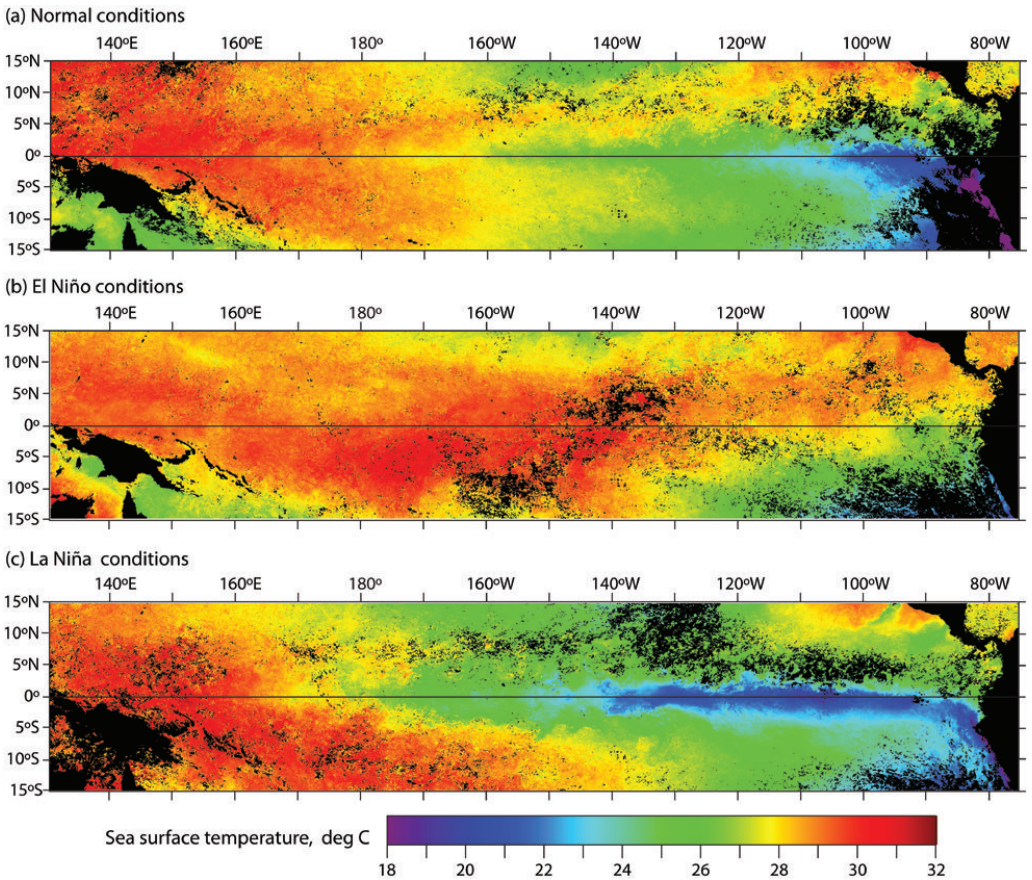


Figure 11.6. Monthly composite SST distributions over the equatorial Pacific Ocean, based on the 4 km resolution Pathfinder version 5 dataset from night-time retrievals from AVHRR. (a) July 1996; normal conditions. (b) November 1997; El Niño conditions. (c) December 1998; La Niña conditions. These images were created using data provided by the U.S. National Ocean Data Center (see *http://www.nodc.noaa.gov/sog/pathfinder4km/userguide.html*), accessed through the NASA JPL PO.DAAC using the data extraction tool POET (*http://poet.jpl.nasa.gov/*).

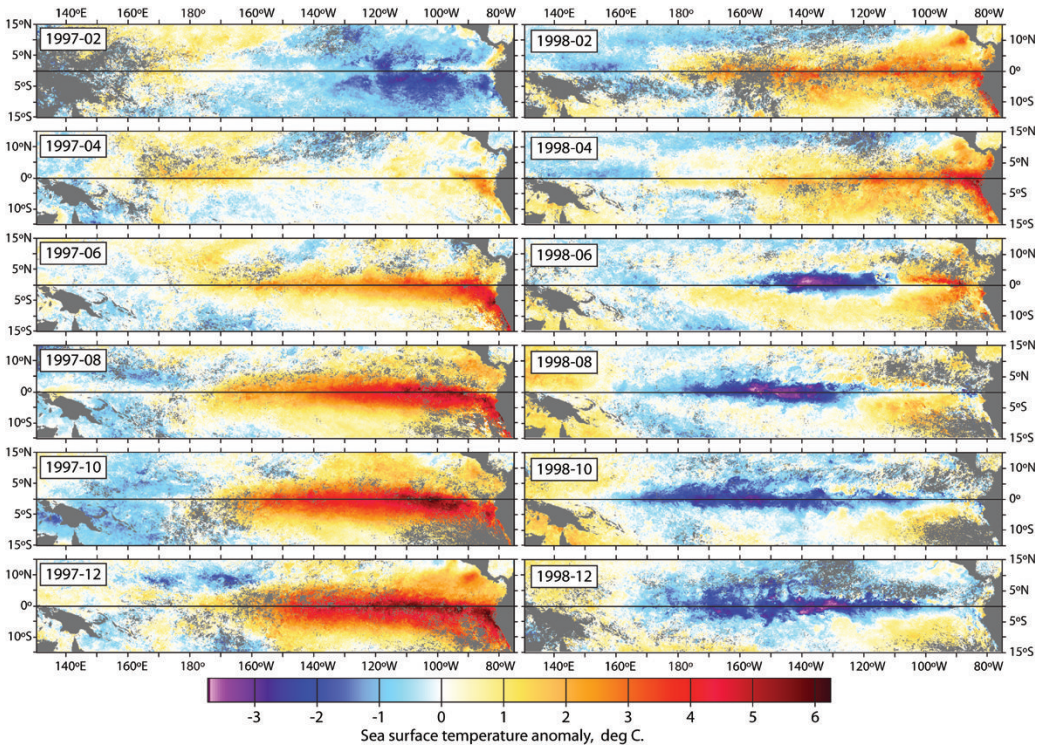


Figure 11.7. Sequence of monthly SST anomaly maps of the equatorial Pacific, every 2 months during 1997–1998. SST anomaly data for these images were specially produced from 4 km monthly Pathfinder version 5 night-time SST (downloaded from http://poet.jpl.nasa.gov/) and the 4 km monthly Pathfinder climatologies (downloaded from *ftp://data.nodc.noaa.gov/pub/data.nodc/pathfinder/Version 5.0_Climatologies/Monthly/Night*). A 5×5 median filter was applied to the anomaly maps. Note that an asymmetric color scale was used because maximum positive anomalies of El Niño are almost twice the maximum negative anomalies of La Niña. Gray shading indicates either land or data dropout caused by persistent cloud.

These are from the Pathfinder version 5 dataset using only night-time retrievals to avoid any effects of diurnal warming. Comparison with the *in situ* observations displayed in Figure 11.2 confirms that both reveal the same characteristic differences in SST distribution between normal, El Niño, and La Niña events. Satellite data contain more spatial detail, although they also contain blank pixels caused by persistent cloud cover which prevents infrared sensors from retrieving SST.

The time sequence of satellite-derived SST anomaly maps shown in Figure 11.7 highlights the evolving pattern of temperature during the 1987–1988 El Niño and La Niña events. Note how the use of anomalies removes seasonal variability and the mean spatial structure of SST to focus attention on ENSO effects alone. For the first few months of 1997 there was no hint in SST of a possible El Niño until the slight warming near the coast of Ecuador in April, which might be no more than an

isolated and local perturbation until May and June confirmed a persistent positive anomaly all along the Equator as the normal upwelling reduced. The positive anomaly grew during the next 2 months into the pattern characteristic of an El Niño, appearing as a warm anomaly plume with a magnitude of more than 3°C attached to the Ecuador coast reaching westwards to 160°W and spreading over 10° of latitude. In this case the anomaly reached its maximum around December 1997 in what was described as the strongest El Niño of the century.

The pattern persists but its amplitude gradually declines from February 1998 until by May it has almost disappeared. However, June 1998 shows that SST along the Equator west of 100°E has started to reduce further to form a large negative anomaly which then persists through much of 1999. This is the characteristic thermal signature of La Niña. In this case there was a rapid flip from El Niño to La Niña without stabilizing for some time at the normal state in between.

The effectiveness of using SST anomalies is apparent from the graphic way in which Figure 11.7 reveals the unfolding of the El Niño drama. As in standard El Niño indicators based on *in situ* temperature anomalies (Figure 11.4), a persistent, very strong (3–6°C) positive anomaly along the Equator is the signature of an El Niño. A strong negative anomaly of magnitude up to 3.5°C indicates a La Niña. However, picking up the discussion from Section 6.2.1, if the anomaly is to give a clear distinction between El Niño, normal, and La Niña states, the climatology used as the baseline for the anomalies must be produced from as long as possible a time series, containing representative numbers of warm or cold events in proportion to their long-term occurrence statistics. Ideally a climatology based on several decades of satellite observations is needed; if shorter climatological baselines are used more care must be taken in interpreting resulting anomalies. In this case the Pathfinder climatology used was derived from AVHRR observations between 1985 and 2001, with gaps filled following the method of Casey and Cornillon (1999).

One drawback with using only anomalies to monitor an ENSO event is the misleading impression of air–sea interaction processes that it might create. Because atmospheric convection has a nonlinear dependence on SST, it is important to know its absolute value as well as the anomaly relative to climatology. For example, when SST rises above 28°C, convection is greatly enhanced (Graham and Barnett, 1987). Most of the warmest El Niño anomalies coincide with regions where climatological temperatures are low, and so affect the atmosphere less than might be expected. When upwelling switches off in the east, the surface temperature rises approximately to a magnitude that is commonplace farther west. What is more critical to ENSO evolution is the longitude where the highest absolute temperatures occur (i.e., the location of the Warm Pool), since this tends to drive atmospheric convection. Its location is not obvious from anomaly maps and so it is important to present absolute SST maps, such as Figure 11.6 as well as anomalies when analyzing air–sea interaction mechanisms in El Niño events. For example, comparing Figures 11.6b and 11.7, for November–December 1997, it appears that the Warm Pool which drives atmospheric convection is found at about 160°–170°W where the positive anomaly is no greater than around 2°C, whereas the use of anomalies alone might misleadingly focus most attention on the region between 90°W and 110°W where the

anomaly is over 5°C The main issue requiring caution when interpreting the evolution of temperature anomalies from satellites is the possibility of bias entering the reported SST due to either excessive undetected cloud contamination (cool bias) or increased numbers of diurnal warming events (warm bias), as discussed in more detail in sections 7.2.4 and 7.3.3, respectively, in *MTOFS*. The former is not a problem for *in situ* records, whereas satellite data dropout (gray areas in Figure 11.7) is evidence of a lot of cloud. Diurnal warming was avoided by using only night-time AVHRR passes for Figures 11.6 and 11.7. The potential for misinterpretation of data arises with a phenomenon like El Niño in which atmospheric changes are coupled to ocean variability. It is conceivable that cloud cover or cloud type could change with the ENSO cycle, as could the occurrence of diurnal warming that depends on wind speed and surface insolation. If this were the case then part of the SST anomaly correlated with ENSO could be an artifact of the measurement process. Although it would be small compared with the large amplitude of the equatorial thermal signature of ENSO, this could be more of an issue when looking for ENSO-correlated perturbations of SST in other parts of the world. The use of microwave radiometry for tracking the ENSO temperature signal can now largely eliminate the cloud problem, since microwave thermometry is largely insensitive to cloud, although a comparable problem is microwave sensitivity to heavy rain, which is another ENSO-dependent variable. It is therefore recommended to adopt the GHRSST approach (Donlon *et al.*, 2007), using SST analyses based on combining SST measurements from several sources (as discussed in Chapter 14). This mitigates against the risk of poor-quality temperature retrieval when a major volcanic eruption degrades the accuracy of infrared data, as happened when El Chichón (Mexico) erupted during the 1982–1983 El Niño. Work is therefore in hand to reanalyze SST data back to at least 1991, using ATSR data to provide a bias correction reference. When completed it will be interesting to discover whether it makes any significant changes to the global SST anomaly signatures of the 1997 El Niño which so far have been based almost entirely on AVHRR data.

It is arguable that the capacity for SST anomaly maps to be produced and published as soon as SST data are acquired provides a powerful visualization that helps not only scientists but the general public to see clearly how the El Niño phenomenon is evolving. SST anomaly maps are readily understood because they define the deviation from normal. They can be used with little ambiguity in newspapers and television news bulletins, so helping people to understand and cope with the impacts of an environmental phenomenon beyond their power to control.

11.2.4 Applying altimetry to the study of El Niño

Although data from Geosat were used to observe the El Niño during 1986–1987 (Delcroix *et al.*, 1994), monitoring of the phenomenon by altimetry has been transformed by the capacity to measure the sea surface height anomaly (SSHA) with an accuracy better than 3 cm. This has been available since the launch of TOPEX/Poseidon (T/P) in August 1992, continued by the Jason-1 mission from 2001 and Jason-2 since 2008. By the time of the major 1997–1998 ENSO event, mean sea

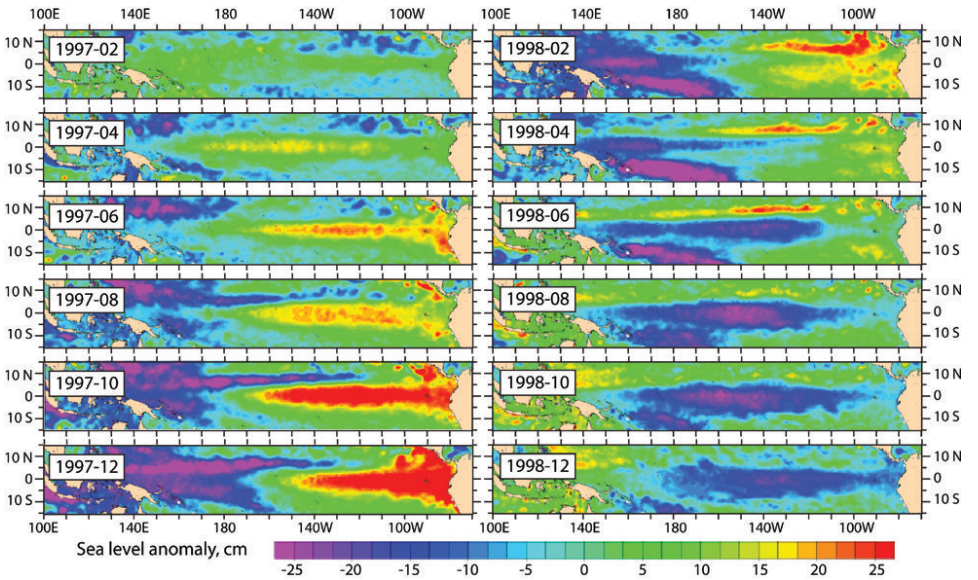


Figure 11.8. Monthly mean sea level anomaly maps (in centimeters) of the equatorial Pacific for every second month during 1997 and 1998. These monthly means have the seasonal cycle removed, referenced to 1993–2006. This figure was constructed by adapting altimeter products produced by SSALTO/DUACS and distributed by AVISO with support from CNES (data were accessed through *http://www.aviso.oceanobs.com/en/home/index.html*).

surface heights were well established for each of the orbit tracks, so that SSHA could be produced readily for each orbit. The SSHAs from every orbit over each 10-day repeat cycle are gridded to produce a time sequence of maps, sampled every 10 days.

Examples of such maps for the equatorial Pacific Ocean during 1997–1998 are presented in Figure 11.8 to show the altimetric signatures of El Niño and La Niña. During 1997 the sea level along the Equator rises towards the east, relative to its climatological level, in response to the reduction of the westward-blowing wind. By June the height in the east at 100°W is 20 cm above normal while it remains at 0 cm at 160°E. By December SSHA has increased to 32 cm at 100°W (not entirely clear in Figure 11.8 because it exceeds the range of the color scale) and has fallen in the west to −20 cm at 160°E, which corresponds to the maximum perturbation for this El Niño. By June 1998 the SSHA at the Equator has reverted to zero at both the east and west coastal margins, but now there is a large surface depression in the center, reaching −24 cm at 150°W.

What is striking about Figure 11.8 is the broad similarity of height anomaly patterns to the SST anomalies in Figure 11.7. The profile along the Equator is almost identical, and the same features are present in each of the maps. At the maximum of an El Niño there is a wedge of positive anomaly, highest at the eastern coast and tapering gradually westwards across the whole of the ocean. In La Niña events there is a negative hollow centered in the middle of the ocean. From an empirical stand-

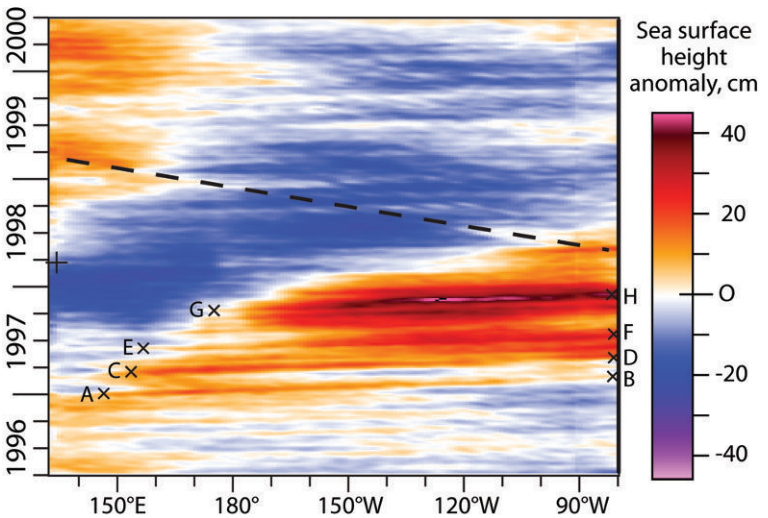


Figure 11.9. Time–longitude (Hovmöller) plot of the sea surface height anomaly (SSHA) measured by TOPEX/Poseidon along the Equator over the width of the Pacific Ocean, during the period April 1996 to June 2000. This highlights the very large positive anomaly of up to 45.5 cm at 110°–130°W during November 1997, corresponding to an El Niño. It is followed by a strong La Niña event shown by the depression of more than 20 cm in 1998 which continued to recur until early 2000 (figure constructed from gridded data at a resolution of 1° and 5 days, downloaded using the PO/DAAC POET facility at *http://poet.jpl.nasa.gov/*).

point, these plots suggest that the characteristic patterns of El Niño and La Niña events must be present in both the SST anomaly and the SSHA, if either phenomenon is to stabilize into a strong sustained perturbation from normal conditions. Finding the signature in both types of satellite image provides strong evidence that an El Niño or a La Niña is in progress.

Another way of presenting the same SSHA data is as a Hovmöller plot at 0°N (shown in Figure 11.9). This plot is based on a gridded SSHA field with a resolution of 1° latitude × 1° longitude and 5 days in time, combined from all altimeters available at that time. The El Niño event shows up as the anomalously high sea level that develops on the eastern side of the Pacific between about May 1997 and April 1998, followed by two successive lows corresponding to La Niña events in winter 1998/1999 and 1999/2000. However, close inspection shows that the high appears to originate from the western side, between 150°E and 170°E, as short individual bursts of high SSHA which last for 2–3 weeks. Some of these have been labeled on Figure 11.9 as A, C, E, and G at the points where they appear to start. Each of these then forms a narrow ridge sloping slightly upwards towards the right, to reach the Ecuador coast about 2–3 months later at the times labeled as B, D, F, and H, respectively.

It is suggested by Picaut *et al.* (2002) that these are individual responses to strong west-wind events which locally force a pulse in surface height and a corresponding

deepening of the thermocline. At a given longitude the SSHA subsides back to its previous level as the wind burst ends. However, the sloping ridge lines on the Hovmöller plot indicate that the sea level disturbance propagates eastwards as a wave front. If we assume that they are solitary baroclinic waves (a large depression of the thermocline), their speed of about 1.9 m/s (120° longitude in 80 days) is characteristic of a Kelvin wave (see Section 6.6.1). Through a more detailed spectral analysis of the Hovmöller record, highlighting signals corresponding to Kelvin and Rossby waves, Picaut *et al.* (2002) argue that the arrival of a Kelvin wave triggers a Rossby wave which travels more slowly back towards the west. The dashed line in Figure 11.9 indicates approximately the speed of westward-propagating signals, just discernible in some parts of the plot. It appears that the fairly rapid succession of Kelvin waves from several wind bursts many thousand kilometers to the west leads to the establishment of the high SSHA which persists for several months at the eastern side of the ocean during the El Niño phenomenon.

Transition to a La Niña event can be discussed along similar lines. The trough of negative SSHA which eventually destroyed the El Niño pattern and replaced it with the La Niña pattern can be seen propagating as a fairly steep front from the west, starting at 150°E in August 1997 and taking almost a year to reach the eastern coast. However, at a finer scale there are streaks of low SSHA spreading east more rapidly, probably Kelvin waves of depression which help to break down the quasistable El Niño state. These phenomena, analyzed in detail by Picaut *et al.* (2002), provide an example of how satellite data can offer new perspectives on the El Niño mechanism. If the start and end of different El Niño/La Niña phases can, in fact, be triggered by wavelike motions that have propagated long distances, Hovmöller plots of SSHA may contribute to improving predictive skills. Such observations, coupled with models, give hope that eventually it will be possible to forecast not only the occurrence but also the magnitude of future El Niño events.

Another way of using altimetry to monitor El Niño activity is to produce a time series of mean SSHA at a point or averaged over an area. Figure 11.10 shows an example produced from average SSHA over the Niño-3/4 region (between 5°S and 5°N and 120°W and 170°W) for the years since TOPEX or Jason data have been available. The corresponding temperature index (ONI, see Figure 11.5) is shown below it for comparison.

11.2.5 Satellite-observed wind fields and ocean surface currents

The ways in which wind fields over the sea can be mapped using satellite data are explained in Chapter 9. Scatterometry provides graphical snapshots of wind speed and direction, typically sampled twice per day. The output from QuikScat, for example, was made readily available from Remote Sensing Systems (*www.remss.com*) in graphical form within hours of acquisition as instantaneous plots (see, e.g., Figure 5.5) and as monthly averages (see Figure 11.17 in the section on observing the monsoon). Although these products were not available during the 1997 El Niño, such data have since allowed the detection of strong anomalous wind bursts in

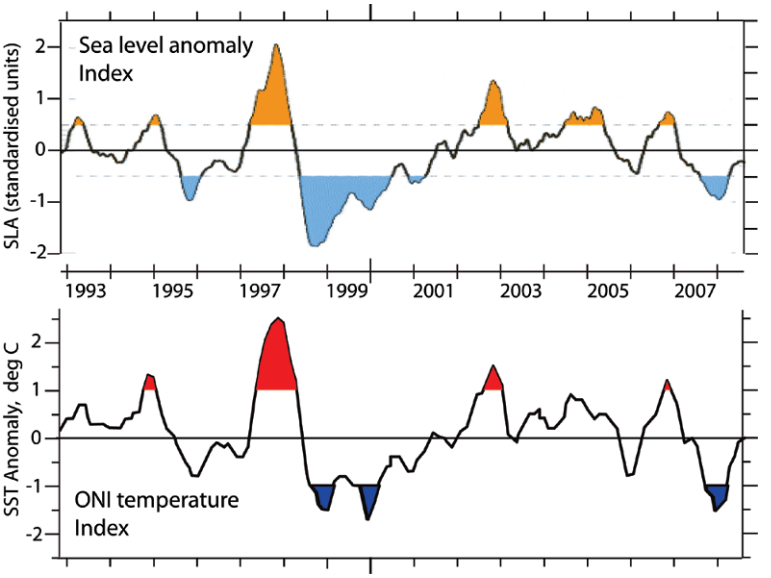


Figure 11.10. *Upper panel*: Time series of the sea level anomaly (in standardized units of sea surface height anomaly) averaged over the El Niño-3/4 region (between 5°N and 5°N and 120°W and 170°W). *Lower panel*: The corresponding part of the Ocean Niño Index (ONI)—see Figure 11.5, which is based on the temperature anomaly in the same region—is shown for comparison (the altimetry index was produced by CLS and obtained from the AVISO website at *http://www.aviso.oceanobs.com/en/home/index.html*).

the equatorial Pacific, as well as monitoring the prevailing wind effect over several weeks, both of which provide insights into the way El Niño events are triggered.

In 1997–1998 the ERS-2 scatterometer was operating, and Figure 11.11 shows a Hovmöller plot of the zonal wind component (i.e., the east–west component) constructed from the monthly mean, gridded, wind field record from ERS-2. This clearly shows that in the first half of 1997 there were anomalous westerly winds in the West Pacific reaching as far east as 170°E. As the El Niño developed in the second half of the year these spread as far east as 160°W–150°W. Ideally daily scatterometer data should also be monitored when trying to distinguish between cause and effect in the air–sea interaction process. While monthly mean values may represent broad conditions established by zonal atmospheric convection cells (see Figure 11.1) they may not reveal the strong but short westerly wind bursts that are implicated in disturbing the sea and triggering Kelvin waves (shown in Figure 11.9).

Given the combination of SSH from altimetry and the vector wind field from scatterometry it is possible to predict ocean surface currents. A simple diagnostic model approach uses altimetry to determine the geostrophic component of surface currents and the wind field to estimate the Ekman component (Lagerloef *et al.*, 1999). However, this has shortcomings in equatorial areas where strong shear may be encountered between surface waters and shallow undercurrents, causing errors in estimated currents especially in the equatorial cold tongue. It is essential to overcome

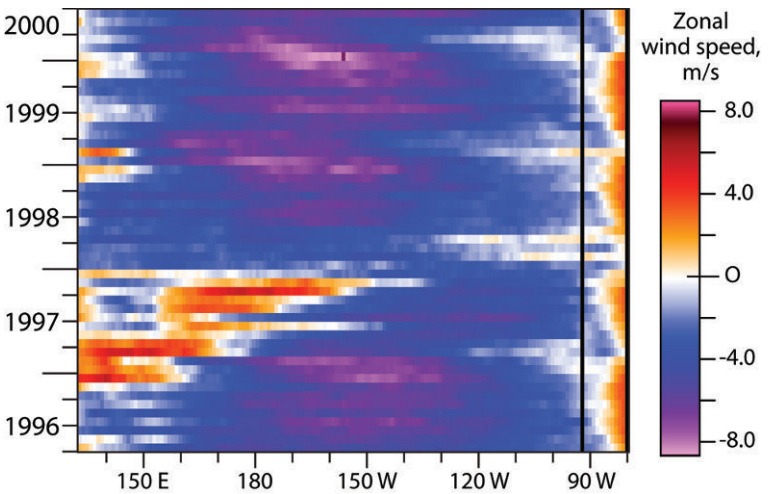


Figure 11.11. Hovmöller plot of monthly mean, zonal wind speed between 1°N and 12°S, at a latitude resolution of 1°, between April 1996 and June 2000. Positive values correspond to eastward-blowing (westerly) winds, and *vice versa*. These measurements are from the ERS-2 scatterometer. The plot was constructed by the author using ERS-2 wind data downloaded from Ifremer's Cersat facility (*ftp://ftp.ifremer.fr/ifremer/cersat/products/gridded/mwf-ers2/data/*).

these shortcomings if the evolution of an El Niño is to be monitored and predicted. Therefore a refined method of retrieving surface currents has been developed (Bonjean and Lagerloef, 2002), which now forms the basis of a data service provided by NOAA, called Ocean Surface Current Analysis–Real time (OSCAR).[2] Diagnostic equations in the revised model address the particular issues of vertical shear and the singularity at the Equator where the Coriolis parameter f is 0. The solution of these equations is approached globally, and requires temperature distribution (from satellites and *in situ* observations) as well as sea surface height and wind stress. A critical comparison (Johnson *et al.*, 2007) of the data products from the OSCAR service against independent current observations confirms that they provide accurate estimates of zonal and meridional time mean circulation. In the near-equatorial region they also provide reasonably accurate estimates of zonal current variability (correlations of 0.5 to 0.8) at periods as short as 40 days and at meridional wavelengths as short as 8°, although the variability of meridional velocity is in general poorly reproduced.

Figure 11.12 shows some examples of surface current fields in the equatorial Pacific Ocean before and during the 1997–1998 El Niño/La Niña event, as produced by the OSCAR service. Surface currents have an important role in advecting water properties. Anomalous eastward-flowing currents along the Equator can help to move the Warm Pool eastwards and contribute to setting up an El Niño event.

[2] See *http://www.oscar.noaa.gov/index.html*

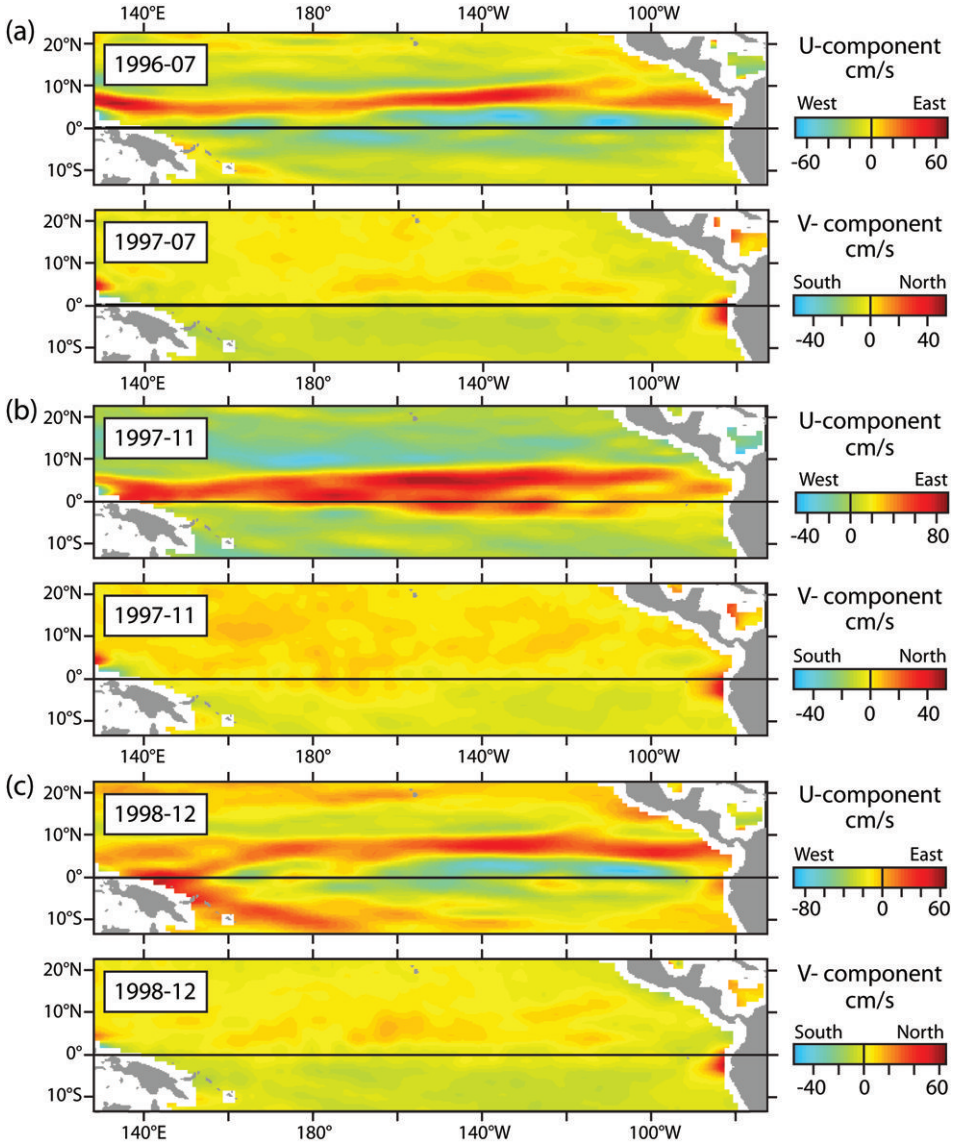


Figure 11.12. OSCAR surface current data products showing zonal and meridional components of monthly mean, retrieved current fields in the equatorial Pacific. (a) July 1996; normal conditions. (b) November 1997; El Niño conditions. (c) December 1998; La Niña conditions (figure based on mapped current fields obtained from *http://www.oscar.noaa.gov/index.html*).

The surface current maps in Figure 11.12 also show the meridional components which transport water between the Equator and higher latitudes, with the potential to change the temperature at the Equator. This is a reminder that, although the basic mechanisms of the El Niño phenomenon are described by equatorial sections such as

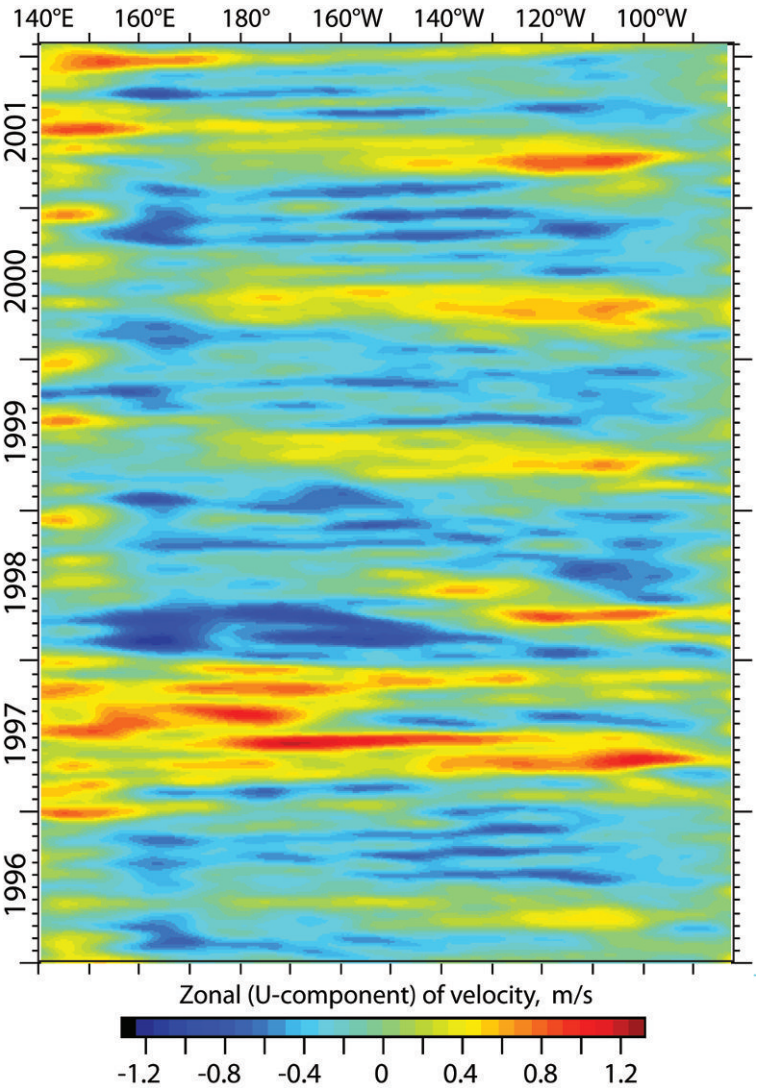


Figure 11.13. Hovmöller plot of monthly mean, zonal surface currents at the Equator over the Pacific Ocean, as retrieved by the OSCAR service, between 1996 and 2001 (adapted from a figure obtained from *http://www.oscar.noaa.gov/index.html*).

Figure 11.1, the full description from which future behavior can be forecast must be three-dimensional. Figure 11.13 presents a Hovmöller plot of these data showing anomalous behavior during the 1997–1998 EMSO event. This should be compared with Figures 11.9 and 11.11, facilitating the study of the link between zonal wind anomalies and advection of warm surface water.

While the OSCAR service aims to supply the needs of clients in a diversity of operational situations, there is no doubt that in this case it makes a useful contribution to monitoring and forecasting El Niño events (Lagerloef *et al.*, 2003). It is an excellent example of how careful processing of data from several satellite types (in

this case altimetry and scatterometry) measuring different variables can deliver near real–time estimates of another variable (the ocean surface current field) which is not directly measured from space. Such an approach exploits dense, two-dimensional spatial sampling of the surface by satellites and makes it more accessible to the wider oceanographic community. In the context of studying the complex behavior of the ocean during the development of El Niño events, it complements the horizontally sparse array of moored buoys which provide necessary vertical sampling of the phenomenon.

11.2.6 Chlorophyll

As explained in Chapter 7, chlorophyll concentration is readily mapped by ocean color sensors. However, individual images of chlorophyll tend to be patchy because of the natural variability of biological populations. Moreover, as discussed in Chapter 5, a mature upwelling system in which primary production is in equilibrium with zooplankton grazers may not display the high chlorophyll concentrations evident in blooms that occur in episodic upwelling, or in nutrient-rich shelf seas, even though they sustain a large fishery. It is therefore interesting to consider how well ocean color sensors are able to detect the reduction in primary production which occurs when the normal coastal and equatorial upwelling system is switched off during an El Niño event, and which can lead to the collapse of fisheries. Monthly averages from global composite images serve as the most effective way of integrating observed chlorophyll in space and time, overcoming short-scale patchiness while still preserving seasonal variability.

When assessing the capacity of satellites to detect the impact of an El Niño, the crucial issue is to be able to compare an El Niño year with a non-El Niño year, and then to determine whether any measured difference is significantly greater than the natural year-to-year variability of non-El Niño years. The 10-year calibrated archive of SeaWiFS data, monitored for long-term stability in comparison with *in situ* data from validation sites, provides an ideal tool for this purpose.

Unfortunately SeaWiFS was not launched until September 1997 by which time the El Niño was well under way. However, initial phases were studied using data available from the OCTS instrument on ADEOS until June 1997 (Murakami *et al.*, 2000), in which a 40% reduction of chlorophyll-*a* was reported for the central equatorial Pacific. Two studies (Chavez *et al.*, 1999; Murtugudde *et al.*, 1999) report the first views of the evolving El Niño using data from the first year of SeaWiFS (1997–1998). Soon further papers followed to report the massive bloom that emerged between 160°W and 130°W, probably responding to iron enrichment from the shallower-than-normal undercurrent, and migrated eastward to the coast during the subsequent La Niña (McClain *et al.*, 2002; Ryan *et al.*, 2002). Strutton *et al.* (2001) also noted that these patches of enhanced productivity were linked to tropical instability waves that are associated with La Niña events. Figure 11.14 shows monthly-averaged chlorophyll measured in the east tropical Pacific for three characteristic stages between 1996 and 1998, corresponding to a normal year, an El Niño, and a La Niña.

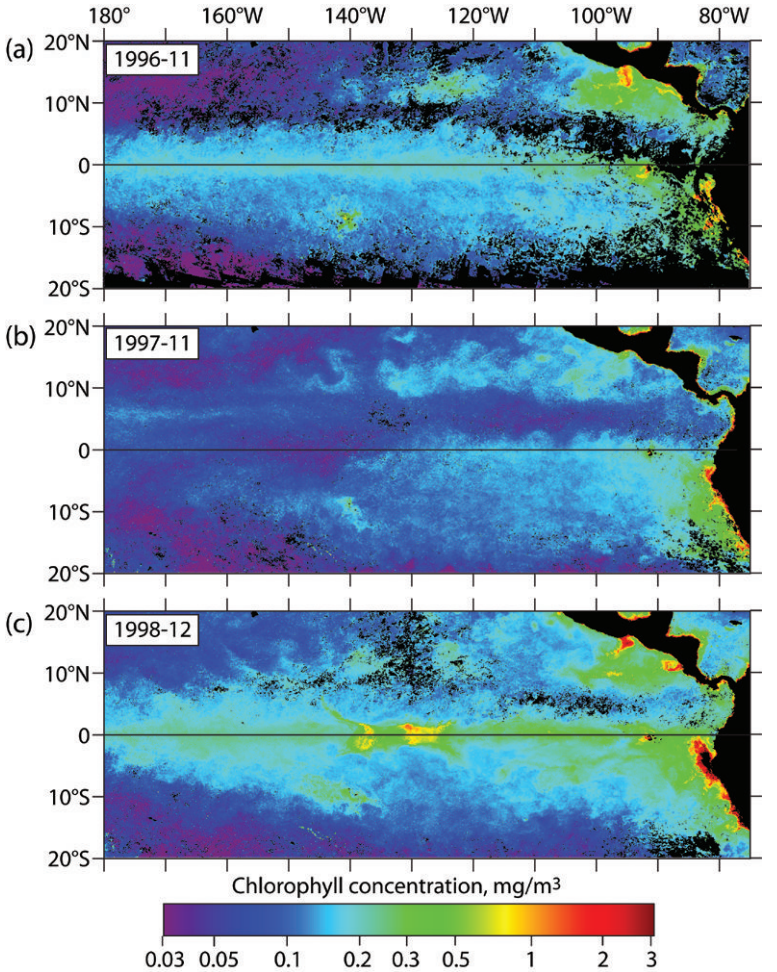


Figure 11.14. Maps of chlorophyll monthly mean concentrations in the eastern equatorial Pacific Ocean, derived from satellite ocean color sensors. (a) November 1996 representing normal conditions. These data were acquired by the OCTS sensor. (b) November 1997 representing El Niño conditions, derived from SeaWiFS. (c) December 1998, representing La Niña conditions, also from SeaWiFS (maps constructed from digital datasets downloaded from the NASA Ocean Color website).

Another useful way of evaluating the El Niño impact on biology is to prepare anomaly images that show the difference between a given month or season or year and the climatological 10-year average for that month or season (11 years for October to December). Figure 11.15 presents anomalies corresponding to the same monthly means as Figure 11.14. Note that these are evaluated as the difference in $\log_{10}$ values of chlorophyll in milligrams per meter and therefore correspond to the multiplicative factor by which the actual month is greater or less than climatology.

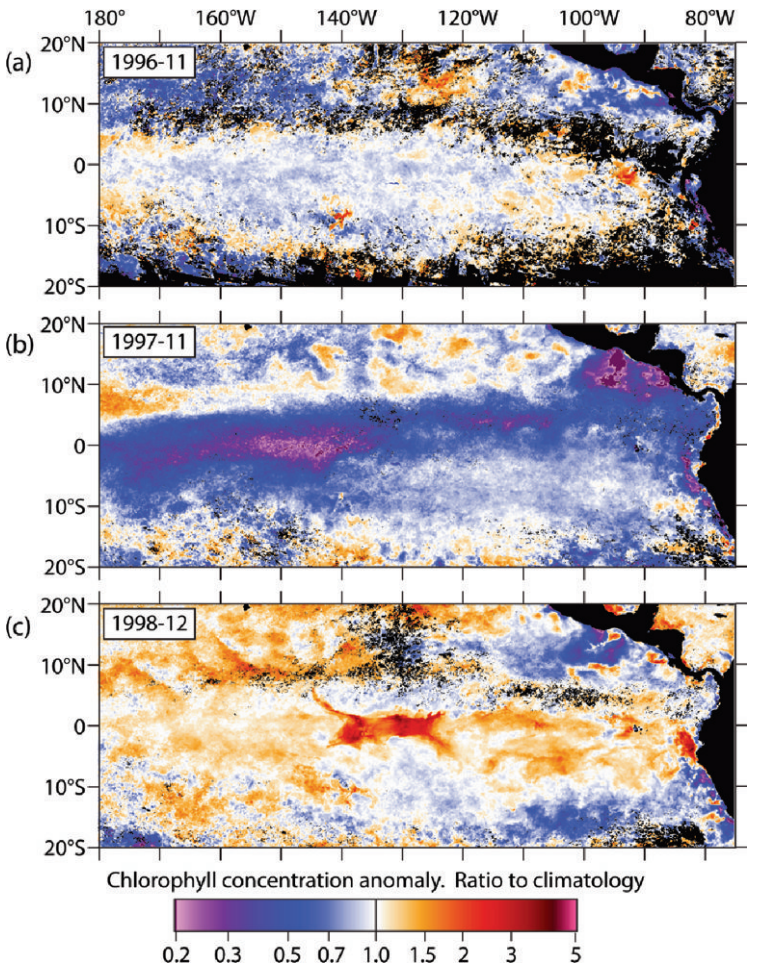


Figure 11.15. Maps of the chlorophyll concentration anomaly in the eastern equatorial Pacific Ocean, derived from satellite ocean color sensors. (a) November 1996 representing normal conditions. (b) November 1997 representing El Niño conditions. (c) December 1998, representing La Niña conditions. These maps were produced using the data presented in Figure 11.14 and the monthly mean climatology of chlorophyll based on 10 years of SeaWiFS data. The anomaly is obtained as the difference between the $\log_{10}$ of chlorophyll in milligrams per meter of the actual month and the climatology of the month. The scale therefore represents at each pixel the multiplying factor of the actual chlorophyll for the given month relative to the climatology at that pixel (constructed from digital datasets downloaded from the NASA Ocean Color website).

Rather than showing the actual structure of chlorophyll highs or lows, they highlight the location and spatial structure of differences.

Given the ready availability of data from SeaWiFS, MODIS, and MERIS the way is open for further exploration of the time variability of chlorophyll distributions in relation to El Niño indices.

11.2.7 Rainfall over the ocean

One of the ways in which the human impact of ENSO events is felt strongly is through changed patterns of rainfall. Over land these are defined quite well. However, it is harder to estimate rainfall over the sea using *in situ* sampling because of gross undersampling by the sparse measurements available, given the intermittent character of rainfall and the small-scale spatial patchiness of rain events, leaving a gap in our knowledge and understanding of the El Niño phenomenon. The development of satellite sensors for rain over the ocean has helped to remedy this problem and it is worth mentioning them here even though strictly this lies outside the scope of a text on satellite oceanography. Measurements have been available from the SSM/I for about three decades and, along with various other sensors from time to time, plus rain gauges over land, these have been compiled into a global, monthly time series of rainfall produced by the Global Precipitation Climatology Project (GPCP) (Adler *et al*., 2003). Since 1997 there has been a satellite dedicated to measuring rain, the Tropical Rainfall Measuring Mission (TRMM) which carries an active rain radar as well as the TRMM microwave imager. The AMSR-E has also produced daily maps of rainfall over the sea since 2002. More information about these microwave radiometers can be found in chapter 8 of *MTOFS* (Robinson, 2004).

Although the most effective rainfall measurements over the sea do not stretch back before 1997, an alternative method was developed using radar altimetry (Quartly *et al*., 1996) which allows statistics on approximate rainfall distribution to be compiled from 1992 onwards. The method for producing these (Quartly *et al*., 1999) does not represent rainfall directly, but indicates the percentage of times when rainfall above the detection threshold was recorded. Following this procedure Quartly *et al*. (2000) calculated bi-monthly estimates of rainfall distribution over the equatorial Pacific from 1993 to 1999, deduced from the TOPEX dual-frequency altimeter, to reveal changes between normal, El Niño, and La Niña years. As an example of their results, Figure 11.16a shows the likelihood of rainfall during November and December averaged over the years 1993–1996, which were normal years, whereas Figure 11.6b shows the same months for 1997 under El Niño conditions. The contrast between the two cases is striking. The main band of rain straddling the Equator in (a) between 140° and 180°E has spread eastwards in (b) for about 40° to reach 220°E, following the movement of the Warm Pool. Moreover, the narrow band of rain at about 10°N spanning the whole width of the Pacific in (a) has spread south to reach the Equator in (b).

As an alternative to mapping rainfall conditions month by month, the availability of the long GPCP time series has made it possible to evaluate the correlation in the time domain between rainfall and regional climatological indices (Kyte *et al*., 2006). The result is a map of the sensitivity of rainfall to the index. The example shown in Figure 11.17 shows the global dependence of rainfall on the Southern Oscillation Index (SOI). When the SOI is positive, corresponding to La Niña conditions, rainfall is enhanced in the regions with red shading and reduced where it is blue. When an El Niño event occurs (negative SOI), rainfall is enhanced

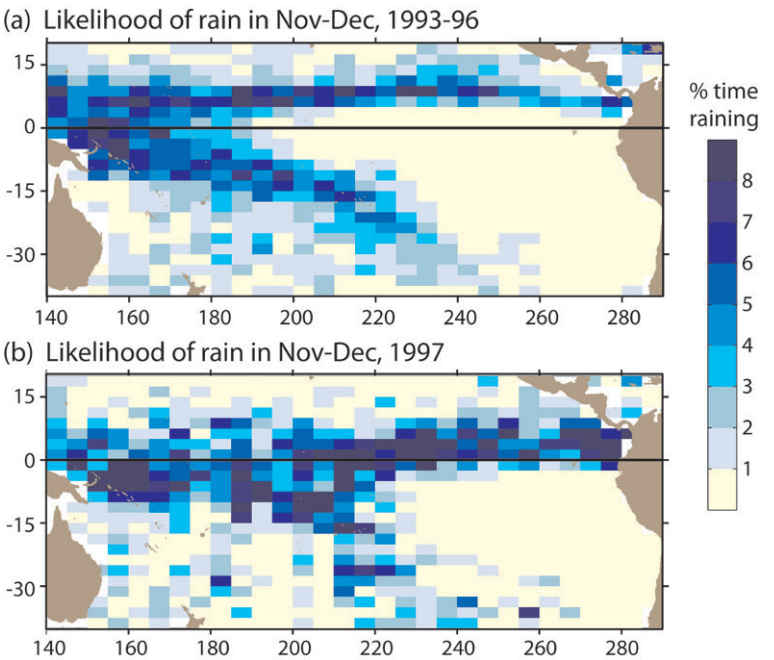


Figure 11.16. Rainfall patterns over the tropical Pacific Ocean associated with an El Niño. (a) The likelihood of rain during November and December of normal years 1993–1996 as measured using the TOPEX dual-frequency altimeter. (b) As (a) but for November–December 1997, an El Niño year (figure courtesy of Graham Quartly, and based on figure 1 of Quartly *et al.*, 2000).

where the map has blue shading and reduced where it is red. Sensitivity maps like this are a very effective way of demonstrating the response of a particular environmental variable to regional fluctuations of climate over interannual and longer timescales. To be effective a time series of one or more decades is desirable, and so far the technique has been applied mainly to variables like rainfall or sea state where long time series exist. As longer series of climate quality datasets are accumulated for satellite-derived SST, SSHA, and ocean color products it will be appropriate to use similar techniques to explore the sensitivity of these variables also to regional climate indices.

11.2.8 Synergy

It should already be evident to readers of the previous sections that the greatest impact of remote sensing on our understanding of El Niño comes from the capacity it provides for intercomparison between similarly sampled measurements of different ocean variables. So far we have considered each remotely sensed variable individually. When the different time series of mapped variables are viewed and analyzed together they point to relationships and open up new ideas about causality. For

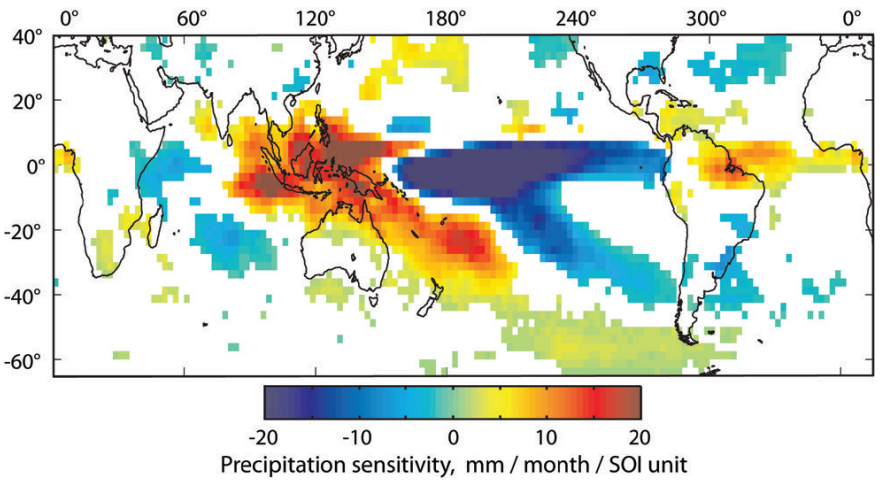


Figure 11.17. Sensitivity of satellite-derived rainfall over the sea—from the Global Precipitation Climatology Project (GPCP) to the Southern Oscillation Index—evaluated over the period 1979–2000. Note that during an El Niño the Southern Oscillation is negative and during a La Niña it is positive, so the blue color shows where rainfall is enhanced during the El Niño and red where it is reduced (figure provided by Graham Quartly, based on part of figure 1 of Kyte *et al*., 2006).

example, the studies by Strutton *et al*. (2001), McClain *et al*. (2002), and Picaut *et al*. (2002) in different ways provide much evidence of this.

By the simple expedient of lining up Hovmöller plots of wind stress, SSHA, and SST (absolute, not the anomaly), Picaut *et al*. (2002) are able to reveal and illustrate interaction mechanisms which offer a much more complete understanding of how these variables interacted during 1997–1998. Alignment of Hovmöller plots shows where there are correlations between perturbations of different variables, and any lags between one or another. By itself a correlation does not confirm causality, but may often suggest possible mechanisms which link different variables together, such as wind bursts, pulses of zonal velocity, and perturbations of sea level anomaly. When sea level pulses are shown to propagate as Kelvin waves, detectable by sloping signatures in Hovmöller plots, this can offer an explanation for the delays that elapse between cause and effect separated by many degrees of longitude.

In a similar way, Hovmöller plots relating enhanced chlorophyll concentration to SST and information about surface layer depth can point to explanations for the source of nutrients required to maintain the blooms. The ultimate goal is to develop modeling tools that will allow ENSO events to be predicted with confidence. The insights gained from continuously collected satellite data, analyzed in a variety of ways, can point to what are the most important interaction mechanisms between different ocean and atmosphere variables that need to be incorporated into predictive models. Following each success or failure of a model to predict, or falsely predict, an event, further analysis of the data allows refinement of models.

In summary, we conclude that the availability of several streams of different ocean variables from satellites has undoubtedly contributed to improved understanding of the El Niño phenomenon. Along with *in situ* observations and exploited by assimilation into numerical ocean–atmosphere models, satellite data are improving the capacity to forecast future events. Moreover, while an episode is in progress, satellite data can readily be used to monitor changes taking place in the ocean and portray them in a clear visualization that can be grasped by the general public.

11.3 MONSOONS

11.3.1 Introduction

The term "monsoon" in its most general sense is used to describe the local climate in the tropics wherever there is a marked shift in wind direction between one part of the year and another, causing rainfall to be strongly seasonal. Its fundamental cause is the change in temperature between land and the adjacent ocean. Whereas SST changes only gradually through the year, land rapidly warms during the summer and cools in the winter relative to the ocean, creating atmospheric pressure gradients that drive seasonal winds from sea to land in summer and the reverse in winter. Regular annual cycles of change in both local weather and local sea conditions have no doubt for thousands of years been part of the life knowledge of indigenous populations of the tropical regions affected, which are mostly in South East Asia and the Indian subcontinent. However, in the 20th century meteorologists were able to show that many different local characteristics of seasonal variability of the weather are part of the wider and, to some extent, globally connected phenomenon of the monsoon.

It is only relatively recently that oceanographers have begun to explore the ocean's role in, and response to, the monsoon (e.g., Fischer *et al.*, 2002; Weller *et al.*, 2002). They have begun to recognize the different behavior of the ocean between years when monsoon winds are strong or weak. It is evident that the response of the ocean to some extent feeds back to affect the transfer of water and heat from the ocean to the atmosphere. An understanding of the ocean's response to monsoon winds is thus necessary for improved seasonal forecasting of monsoons. This is of considerable human importance given the impact of rainfall on the wellbeing of hundreds of millions of people living in tropical countries, especially when rainfall is significantly greater or less than expected.

As in most regional air–sea interaction processes, satellite data have an important role to play alongside the deployment of moorings and other *in situ* oceanographic measurements, bringing the broader spatial and temporal overview which relates localized experiments to the regional and global context. Satellite data can provide regular measurements of the spatial distribution of ocean properties such as currents, turbulent eddy mixing, temperature, upwelling, and primary production, as well as of surface winds and waves which are important for shipping.

This section of the chapter explores the capacity of remote sensing to illustrate the typical response of the ocean to monsoon winds. It focuses on the Indian Ocean where monsoon behavior relates to the Indian subcontinent. A similar approach could be applied to the South East Asian monsoon and the response of the South China Sea, although detailed processes will vary from place to place depending on local geographic circumstances.

11.3.2 Illustrating the Indian monsoon using satellite data

Scatterometer data offer a simple way to characterize the typical monsoon wind forcing that occurs in the Indian Ocean. Figure 11.18 shows monthly mean winds for April, July, October, and December 2005. The southwest monsoon occurs during the boreal summer months of June to September when winds blow from the southwest off the Indian Ocean onto the Indian subcontinent, as shown in Figure 11.18 for July (b) in contrast with April (a) and October (c) when winds are weak and lack a dominant direction. In the winter months of December to January the winds blow strongly from the northeast off the land onto the sea, as illustrated for December in Figure 11.18d. Comparable seasonal differences occur over the East Indian Ocean and South East Asia. Rainfall occurs wherever the wind has traveled for some distance over the sea, gaining higher water vapor content, and then blows over higher land (e.g., over northwest India during the southwest monsoon in July). Each year there are slightly different patterns of wind and the timing of the strongest winds also varies from year to year. Readers can explore for themselves the interannual variability using the Remote Sensing Systems website (*www.remss.com*) where these images were obtained.

The SST distributions corresponding to these wind data in April, July, October, and December 2005 are shown in Figure 11.19. These monthly composites are derived from the Pathfinder dataset version 5.0, based on acquisitions from the AVHRR infrared sensor. During the southwest monsoon, upwelling occurs off the Somali and Arabian coasts. It is strongest between 5°N and 11°N, where upwelling water has a temperature of about 14°C. During the northeast monsoon, strong upwelling occurs along the northwest coast of India and in the Bay of Bengal.

Complete reversal of the wind direction between the southwest and northeast monsoons in summer and winter results in major changes of circulation in both basins of the North Indian Ocean, the Arabian Sea, and the Bengal Sea. This makes it different from all other oceans where major currents and gyres may modulate seasonally but do not reverse. Figure 11.20 illustrates this. It represents mean SSHA over several years for July and January. Geostrophic currents associated with the SSHA flows cyclonically (counterclockwise in the northern hemisphere) around the lows and anticyclonically around the highs. Although the SSHA can only be interpreted in terms of variable currents relative to steady flows, monsoon behavior causing seasonal flow reversals means that over much of the area shown the mean flow is small and so for this special case currents retrieved from the sea level anomaly are quite similar to absolute currents.

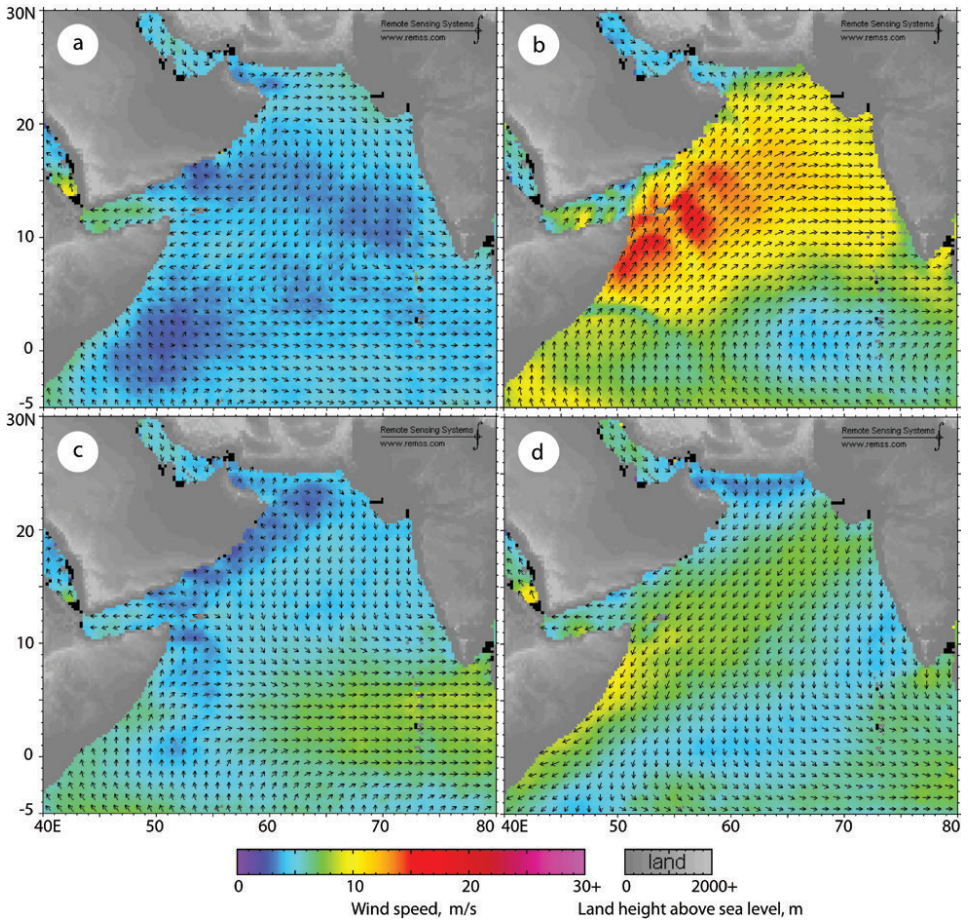


Figure 11.18. Monthly mean wind vectors retrieved from QuikScat over the Arabian Sea for: (a) April 2005; (b) July 2005 during the southwest monsoon; (c) October 2005; and (d) December 2005 during the northeast monsoon. Data obtained from the website of Remote Sensing Systems (QuikScat data are produced by Remote Sensing Systems and sponsored by the NASA Ocean Vector Winds Science Team) (image adapted from a graphic image map acquired from *www.remss.com*).

Strong seasonal upwelling evident in SST maps in Figure 11.19 has a controlling effect on primary production in the region, as is evident in maps of chlorophyll derived from ocean color sensors (as shown in Figure 11.21). In May, when there is little wind-induced upwelling over the Indian Ocean, chlorophyll concentration is at its lowest point during the year, with primary production found only close to the Arabian coast and in the gulfs of northwest India. During the southwest monsoon, strong upwelling along the Somali and Arabian coasts leads to strong phytoplankton growth not only off those coasts but spread across the Arabian Sea and around the south coast of India. This is shown in Figure 11.21b, which is for September, towards

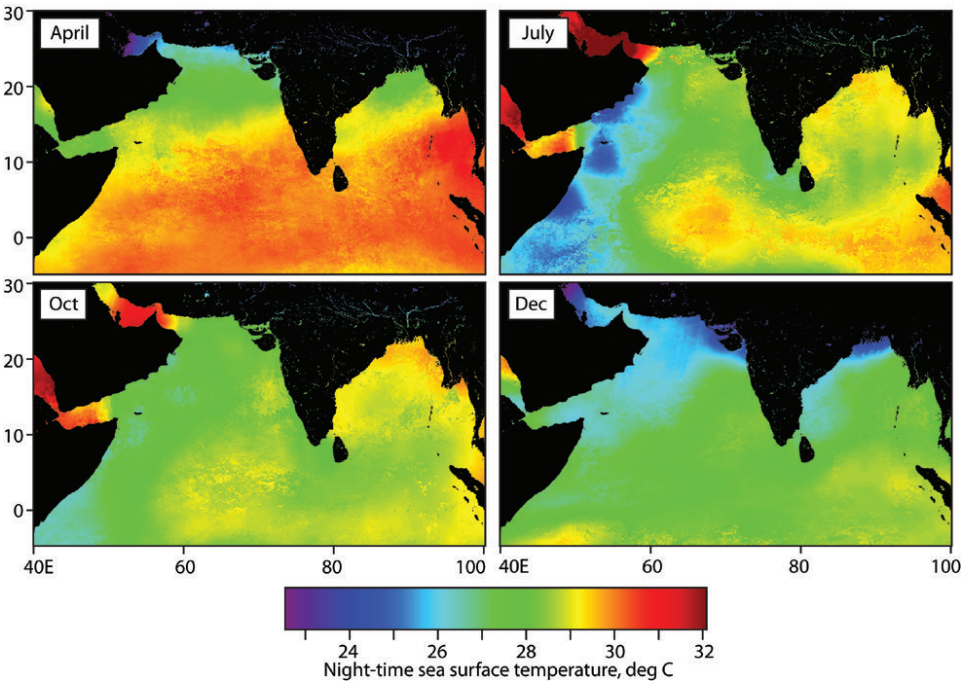


Figure 11.19. Monthly composite SST images from Pathfinder version 5 processing of AVHRR infrared data over the North Indian Ocean for April, July (southwest monsoon), October,and December (northeast monsoon), 2005 (image maps produced by the author from digital data obtained from the NODC website).

the end of the southwest monsoon. It must be noted that such is the density of cloud cover during July and August that it is very difficult to obtain cloud-clear monthly composites from ocean color sensors for those months over the north parts of the Indian Ocean, one of the limitations of satellite observations of monsoon behavior. By November, chlorophyll concentration has reduced considerably between monsoon periods, although not as much as in April–May. The pattern of enhanced production associated with the northeast monsoon (shown in Figure 11.21d), is very different from that of the southwest monsoon, being found mainly in the Arabian Gulf.

11.3.3 Interannual variability of the Indian monsoon

While the annual cycle of the ocean response to monsoon winds occurs every year approximately as shown by the examples in the previous section, it does not repeat itself precisely. Because monsoon winds vary in strength from year to year there is also interannual variability in the ocean's response. However, this raises the question of whether it is entirely a case of the ocean following the atmosphere, or is the atmosphere affected by the ocean, bearing in mind it is ocean–land temperature

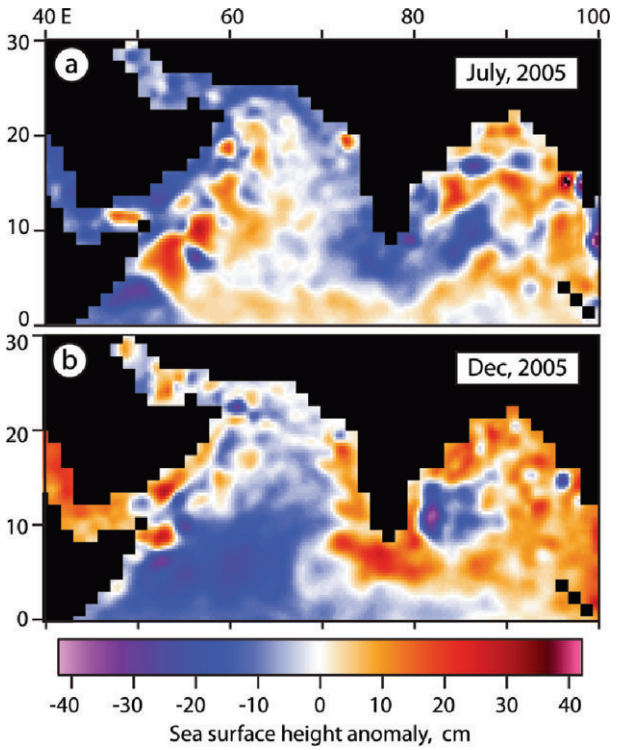


Figure 11.20. Sea surface height anomaly maps over the Indian Ocean from the multimission SSHA merged product from AVISO. (a) Final week of July 2005 showing southwest monsoon conditions. (b) Final week of December 2005 showing northeast monsoon conditions (images produced by the author from digital data downloaded from the AVISO website).

contrast that tends to drive winds in the first place. If the southwest monsoon winds are weaker than usual, as in 2002, then rainfall is reduced and droughts may seriously reduce the agricultural harvest. Conversely a stronger-than-usual monsoon as in 2004 can cause flooding with consequences that economically may be equally devastating.

The amount of monsoon rainfall is recorded in the Monsoon Index based on annual measures of all-India summer rainfall (Parthasarathy *et al.*, 1992) although several other ways have been proposed for characterizing monsoon strength (Wang and Fan, 1999) including those based on wind circulation (Webster and Yang, 1992). This is of more than academic interest since variability of the Indian monsoon is one of the strongest climate signals after the ENSO and can be related to other monsoon systems such as that in South East Asia and to the ENSO state (Gadgil, *et al.*, 2004).

In recent years the availability of satellite-derived measurements of ocean currents, SST, and the satellite-derived heat fluxes discussed in Chapter 10, have enabled a more detailed study of air–sea interaction processes (Liu and Xie, 1999). For example, information over the ocean about heat flux and SST has helped to account for the strong differences between the southwest monsoon in 2002 and 2003 (Ramesh Kumar *et al.*, 2005). Altimetry has allowed the effect of monsoon strength on mesoscale variability in the Arabian Sea to be observed (Subrahmanyam *et al.*, 1996; Subrahmanyam and Robinson, 2000). It is to be expected that satellite ocean

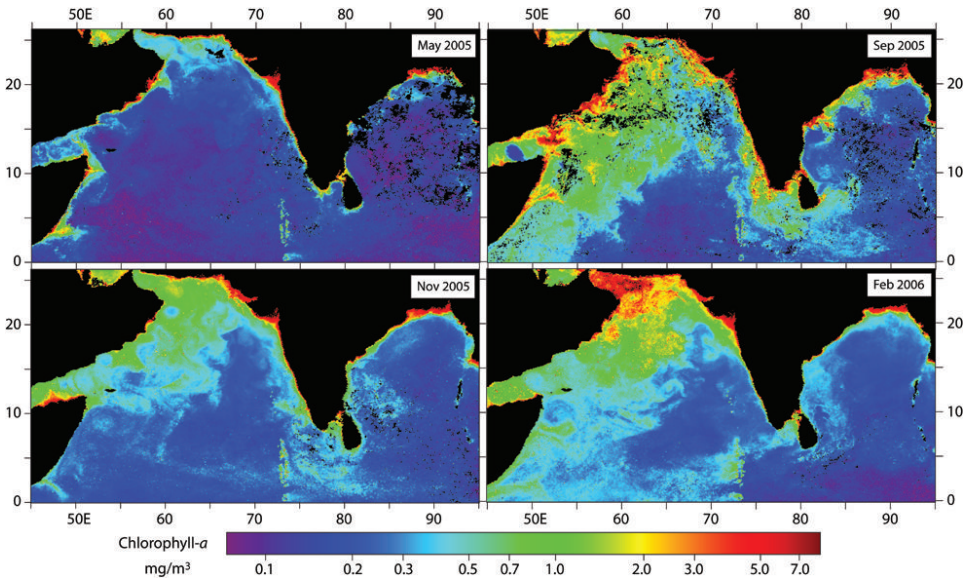


Figure 11.21. Satellite-derived maps of chlorophyll concentration at different stages of the Indian Monsoon. (a) May 2005. (b) September 2005 towards the end of the southwest monsoon. (c) November 2005. (d) February 2006 towards the end of the northeast monsoon (image maps produced by the author from digital datasets of monthly averaged, SeaWiFS chlorophyll datasets downloaded from the NASA Ocean Color website.

observations will be increasingly exploited for the study of air–sea interaction within monsoon systems around the world, and will help to contribute towards improved forecasting capability.

11.4 SEA ICE DISTRIBUTION

11.4.1 Introduction

Remote sensing of polar regions is an extensive field of study to which whole volumes have been devoted (e.g., Carsey, 1992; Haykin *et al.*, 1994; Wadhams, 2000; Rees, 2005; Lubin and Massom, 2006). Most such work lies outside the scope of this volume, but the monitoring of sea ice and its extent is of equal interest to oceanographers as to polar scientists. The annual advance and retreat of the edge of sea ice, releasing excess salt as ice is formed in autumn and releasing cold fresh water as it melts in spring is an important element in understanding the oceanography of high-latitude seas. Interannual or decadal variability and secular trends in the seasonal cycle of ice concentration in both polar regions is likely to have an impact on ocean circulation more widely, up to the global scale. It is also an important

component of the Earth's climate and potentially a key factor in understanding and predicting climate change.

For example, in the North Atlantic the advance and retreat of sea ice affects water mass characteristics (T, S) in the Greenland and Norwegian Seas (Peterson *et al.*, 2006). Here the formation of bottom water takes place as a stage in Atlantic meridional overturning, part of global thermohaline ocean circulation (Rahmstorf, 2006). It is not yet clear to what extent ice behavior affects Atlantic meridional overturning circulation. However, as we see the north polar ice cap apparently retreating rapidly it is of more than academic interest to consider how this may have consequences for the climate of northwest Europe and ultimately global, deep-water ocean circulation. Similarly in the south, variability of the Antarctic sea ice front is related to circulation and hydrography of the Southern Ocean. For example, see the fascinating observations made by a different approach to "remote sensing" by Charrassin *et al.* (2008) who analyze data recorded by *in situ* sensors carried by elephant seals under and around the ice where neither satellite nor conventional ocean sampling can presently reach.

From an even wider Earth system perspective, the polar ice caps have a significant impact on Earth's albedo. Reduction of the extent of north polar sea ice in summer reduces the reflection of sunlight, changing the amount of solar heating absorbed by the Arctic Ocean and the polar atmosphere. Meteorologists have already seen the impact of this on northern hemisphere weather systems. It is no longer a purely hypothetical question to ask how the absence of a summer ice sheet in the Arctic Ocean will affect climates around the world. This is why the topic of sea ice distribution finds a place in this chapter where human impact is the common factor. That is reinforced by the fact that remote sensing has been so instrumental in revealing clearly how the long-established patterns of sea ice are changing.

11.4.2 Measuring sea ice from space

The satellite sensor we shall focus on is the microwave radiometer which provides coarse-resolution climatological records of the annual advance and retreat of sea ice. Microwave radiometry is not the only way of mapping sea ice. In fact, for the crucial operational task of mapping the ever-changing landscape of sea ice as encountered by mariners in polar seas, the coarse-resolution microwave radiometer is not appropriate because ships navigating ice-infested water need to know precisely where the leads of open water can be found within the pack ice. While visible and infrared medium-resolution sensors have some part to play too, the key sensor for this activity is the synthetic aperture radar (SAR). In the last decade there have been advances made in sea ice monitoring at the sub-100 m resolution scale. These techniques and scientific understanding of the processes are now applied operationally (Onstott and Shuchman, 2005; Askne and Direking, 2008) using a combination of the ASAR on Envisat and the Radarsat-2 SAR, although there is no space in this volume to elaborate further.

Microwave radiometry distinguishes the open sea surface from floating ice because ice, with a higher microwave emissivity than water but only a slightly cooler temperature, appears brighter. Thus the relative proportions of open water and ice within a single field of view of a radiometer can be estimated from the measured brightness temperature. Of course other factors affect the brightness, such as atmospheric moisture, actual temperatures of the sea and the ice, roughness of the sea, and texture of the ice, but these need not be known explicitly. As is typical for retrieving environmental variables from microwave radiometry (see Section 2.4.4 in this volume and chapter 8 in *MTOFS*) the use of multifrequency radiometers with different polarization detectors has allowed empirical algorithms to be developed that directly predict sea ice concentration. Section 8.3.7 of *MTOFS* outlines these algorithms and their limitations. Specific algorithms were developed for use with the Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR) operating between 1979 and 1987 (Cavalieri *et al.*, 1984), and its successor the SSM/I on the U.S. Defence Meteorological Satellite Program (DMSP) satellite series (Cavalieri *et al.*, 1991, 1995).

The U.S. National Snow and Ice Data Center (NSIDC) has developed a Sea Ice Index (Fetterer *et al.*, 2002, updated 2008) which makes sea ice data publicly available through a web interface. It supplies daily updates of sea ice concentration produced by the NASA Goddard Space Flight Center (GSFC), and places these in the climatological context. It also provides a variety of ways to present data online, including animations of time series. The main product is a sea ice concentration map for each polar region, at a nominal resolution of 25 km. Provisional versions are available within one day of acquisition and these are replaced by final versions within about one month. A service providing similar products but from an independent data analysis system is available from the Oceans and Sea Ice Satellite Applications Facility (OSI-SAF) of Eumetsat.[3] Figure 11.22 shows an example of an OSI-SAF daily map of sea ice concentration in the southern hemisphere on August 22, 2009. Towards the end of the austral winter south polar sea ice has nearly reached its maximum coverage, although the extent of colored regions on the image indicates that a lot of the ice-dominated area is not completely covered by sea ice (only where the concentration is 100% is the image purely white in the color palette used to scale this dataset). Where concentration is much less than 100% it implies that there are polynyas and leads of open water present in places, although the characteristics of these is unresolved by the microwave radiometer, and SAR would be needed to map them. However, it is only in a narrow fringe adjacent to the open sea that ice concentration falls below 50%.

Examples of NSIDC sea ice extent maps are shown in Figure 11.23 for Antarctica and Figure 11.24 for the Arctic. These maps are derived from ice concentration data by defining ice extent as follows. Any 25 km pixel in ice concentration images where concentration is less than 15% is treated as open water. Where ice concentration is 15% or more it is treated as ice-covered, the 15% contour being treated as the ice edge which determines ice extent. The advan-

[3] See *http://www.osi-saf.org/*

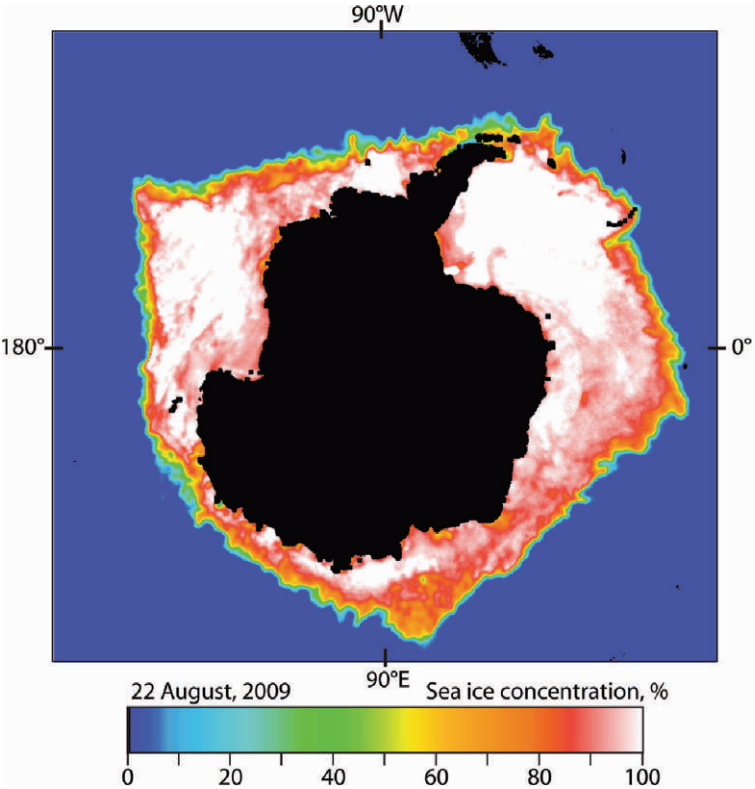


Figure 11.22. Daily map of sea ice concentration around the Antarctic on August 22, 2009. Ice concentration is the percentage of full ice cover within a pixel (image produced by the author from a NetCDF file obtained from the OSI-SAF website).

tage of this approach is to highlight the advance and retreat of the main coverage of ice each year, as it affects navigation. Ice extent may also respond to the wind, increasing if the wind promotes divergence in surface currents that opens up leads and reduces the concentration. Note that for scientific analysis of the volume of ice, it is more appropriate to use ice concentration maps than ice extent. The thick, gray line in Figures 11.23 and 11.24 shows the median position of the 15% ice edge for the same stage in the seasonal cycle. It is based on 22 years of archived data from January 1979 to December 2000. For the given stage during the year, data from all years are used to evaluate the probability that a pixel has more than 15% ice cover. The climatological ice edge for that stage of the year is defined by the contour along which that probability is 50%.

Monthly composite maps of sea ice concentration are produced by averaging, for each pixel, the concentration from every daily dataset that contains valid data during the month. To ensure quality, there must be at least 20 valid days per month or the pixel is flagged as no data. Monthly ice concentrations have a meaning that is subtly different from daily maps. Whereas for daily data the value represents a true measure of the instantaneous spatial concentration of ice, monthly data also averages over time, confusing the meaning of concentration. For example, if a particular pixel reports 50% concentration for the month, it could be that for every day

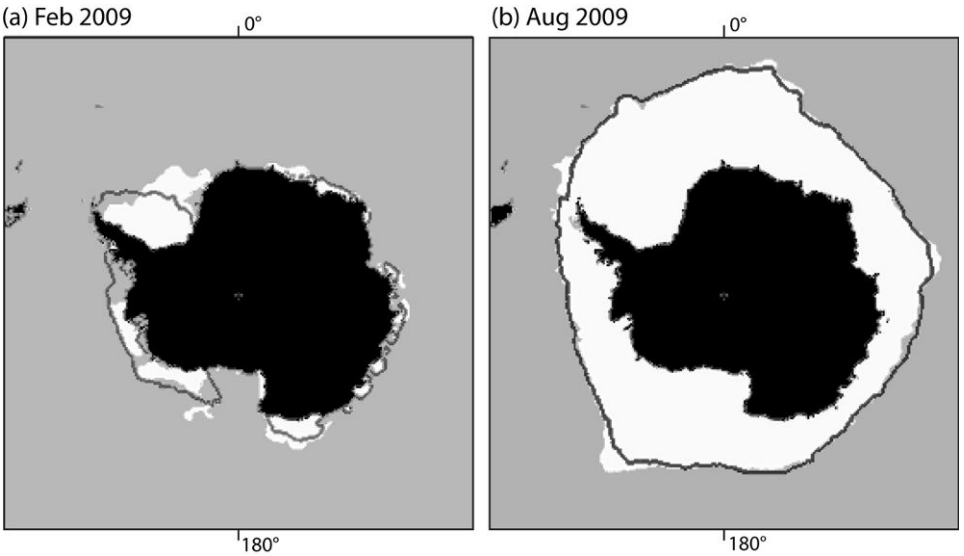


Figure 11.23. Monthly sea ice extent in the Antarctic Ocean for (a) February 2009, corresponding to the austral summer minimum; (b) August 2009, austral winter maximum. The extent, shown in white, is the area where concentration is greater than 15%. The thick, gray line is the median ice edge for that month for the period 1979–2000 (figure based on images obtained from the Ice Index Archive available through the NSIDC website).

the spatial concentration was 50% or, at the other extreme, it could be that for half the days the concentration was 100% and for the other half there was no ice at all.

From monthly averages, seasonal climatologies have been established for the period 1979–2000 and anomaly maps are produced by comparing individual months with the climatology for that month. Trend maps are also produced by fitting, for each pixel individually, a trendline over values for the same month in previous years.

Note that Figures 11.23 and 11.24 are in fact monthly maps of sea ice extent, produced from monthly composite concentration maps. In each case they show the month of maximum extent which is August for Antarctica (Figure 11.23b) and March for the Arctic (Figure 11.24a), and also the month of least ice coverage which is February for the Antarctic and September for the Arctic. Within each figure the comparison between (a) and (b) shows clearly the large difference in ice extent between winter and summer conditions. The greatest difference is in the southern hemisphere where the mean (1979–2000) sea ice extent is $18.1 \times 10^6\,\mathrm{km}^2$ in winter and drops to $2.9 \times 10^6\,\mathrm{km}^2$ in summer. In the northern hemisphere, the winter mean extent is $15.7 \times 10^6\,\mathrm{km}^2$ and drops to $7.0 \times 10^6\,\mathrm{km}^2$ in summer although that has changed considerably since the mean was established in 2000, as discussed below.

There is a wealth of detailed information to be found in daily and monthly time series of sea ice concentration maps from the last 30 years. For example, it is

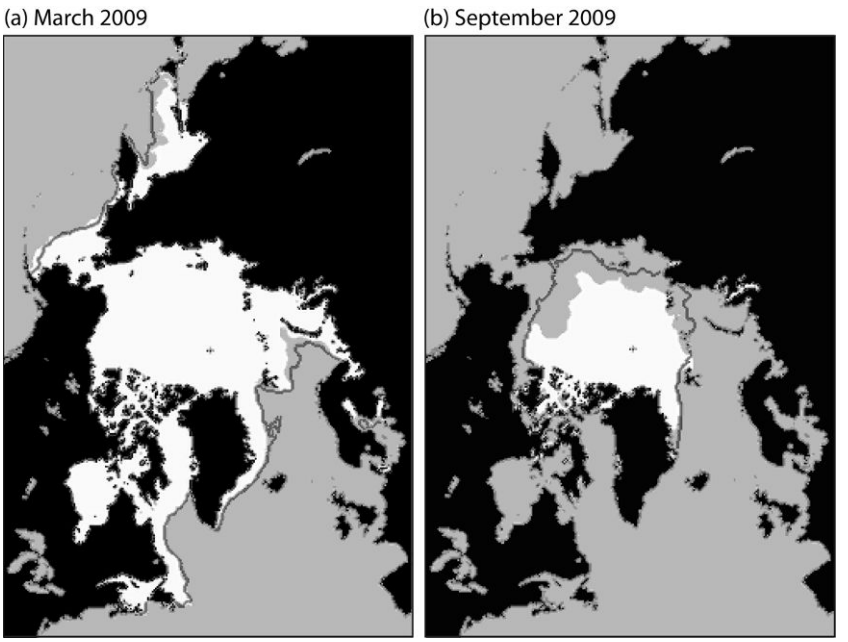


Figure 11.24. Monthly sea ice extent in the Arctic Ocean for (a) March 2009, corresponding to the boreal winter minimum, (b) September 2009, boreal summer minimum. Ice extent, shown in white, is the region where concentration is greater than 15%. The thick, gray line is the median ice edge for that month for the period 1979–2000 (figure based on images obtained from the Ice Index Archive available through the NSIDC website).

fascinating to examine the way ice returns to the Arctic Ocean following the end of summer. Figure 11.25 shows this as a series of images spaced at 2-week and 4-week intervals. In 2009 the northeast navigation passage opened for the first time to allow a few large container ships to reach European ports directly from East Asia. Clear open sea can be seen all along the Siberian coast in (a), (b), (c), and (d) but rapidly closes in after that. Ice gradually spreads down the east coast of Greenland over the period shown. The Bering Strait remains open until November (g) but has closed completely by (h) and (i).

11.4.3 How is the distribution of sea ice changing?

Figure 11.26 shows the 30-year time series of monthly mean ice extent in minimum and maximum months in each hemisphere. These data come from the NSIDC Sea Ice Index, but as plotted here they are not scaled in relation to the percentage anomaly from the 1979–2000 climatology (the percent change values shown on the left axis). Instead the y-axis is scaled according to the anomaly of sea ice extent given in absolute units of actual area (labeled on the right axis). Thus variability of the actual area covered can be compared between north and south and across different

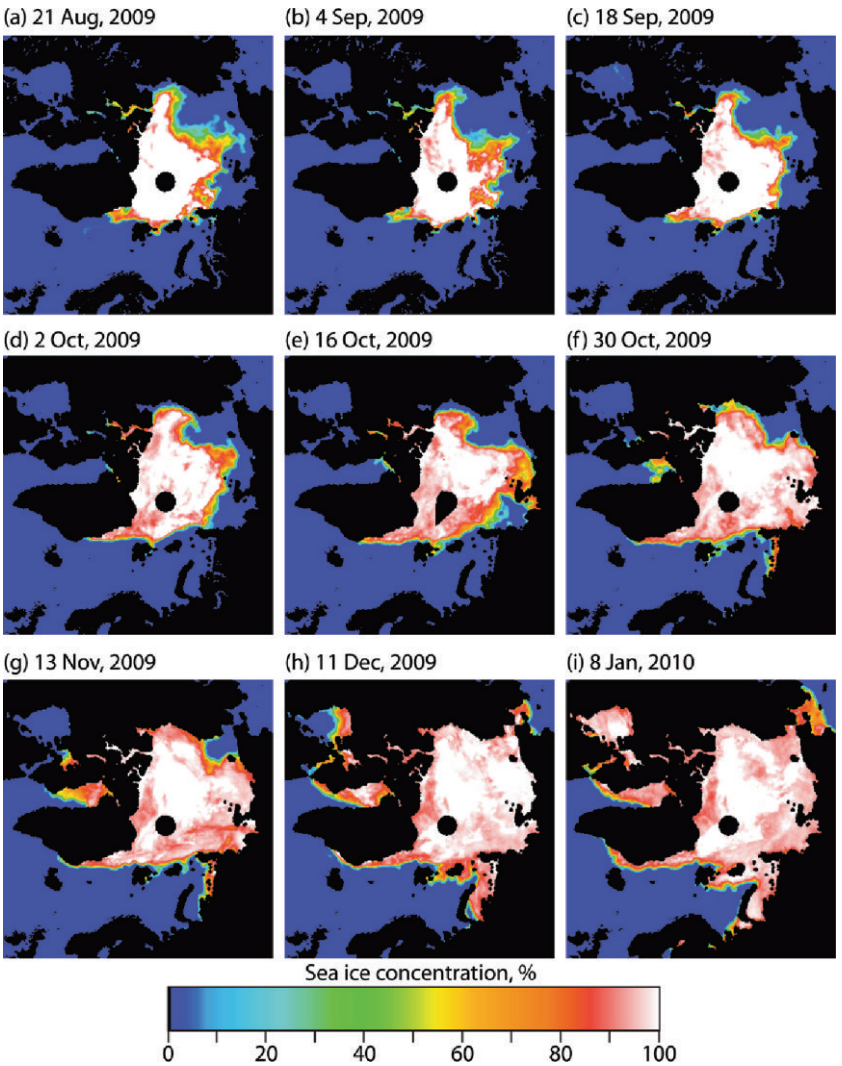


Figure 11.25. Time series of OSI-SAF sea ice concentration maps for the Arctic Ocean at 2-week to 4-week intervals showing the pattern of growth of sea ice coverage following the end of summer in 2009: (a) August 21, (b) September 4, (c) September 18, (d) October 2, (e) October 16, (f) October 30, (g) November 13, (h) December 11, (i) January 8, 2010 (figure produced by the author from the NetCDF files of digital image data obtained from the OSI-SAF website).

months. Note that the origins of area axes are different in each case although incremental scales are the same.

Sea ice around Antarctica varies from year to year but there are no strong trends apparent. Sea ice extent is greatest in August or September (Figure 11.26c) and least in February (d). Interannual variability has a similar amplitude for both maximum

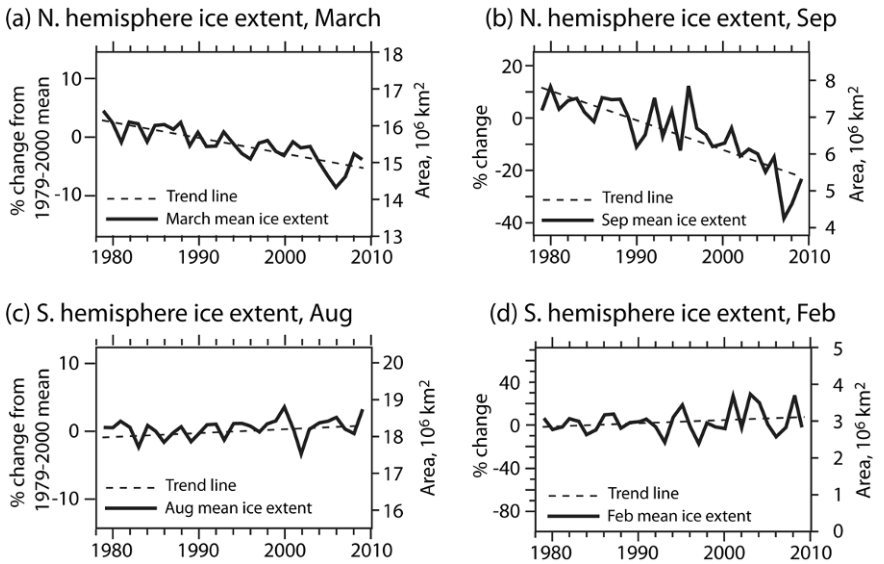


Figure 11.26. Annual time series (solid line) and trendline (dashed) between 1979 and 2008 of sea ice extent averaged over a month, showing (a) northern hemisphere in March, maximum ice extent in the Arctic; (b) northern hemisphere in September, minimum ice extent in the Arctic; (c) southern hemisphere in August, maximum ice extent in the Antarctic; (d) southern hemisphere in February, minimum ice extent in the Antarctic. These graphs were plotted using data obtained from the NSIDC Sea Ice Index. The data are provided by NSIDC as percentage anomalies relative to the climatological (1979–2000) monthly mean (left axis). However, in these graphs the vertical axis has been scaled individually by the climatological mean extent for that month so that the absolute anomaly in units of area (right axis) is scaled the same for all four plots.

and minimum months although in percentage terms it is much greater for the minimum month. Although Figure 11.23 shows only a single year, the difference between the position of the climatological February ice edge and the ice edge location in February 2009 is evidence that in local regions of the Antarctic there is greater interannual variability of sea ice extent although the overall total plotted in Figure 11.26 is steadier. This implies that in years when some regions have less ice, others have more. Zwally *et al.* (2002) discuss the issues associated with observed variability of Antarctic sea ice.

Trends in the Arctic Ocean are very different (as shown in Figure 11.26a, b). Here winter maximum extent is reducing at a rate of 2.8% per decade using the 1979–2000 mean as a baseline, and this is significant in relation to the standard deviation of interannual variability. However, the trend in late summer extent, minimum amount of sea ice left after the spring and summer melt, shows an even greater decrease than the winter in absolute terms. In percentage terms the trend was calculated in 2009 as $-11.2 \pm 3.1\%$ per decade. Until 1997 the trend was much weaker than this but during the next decade the decline accelerated, leading scientists

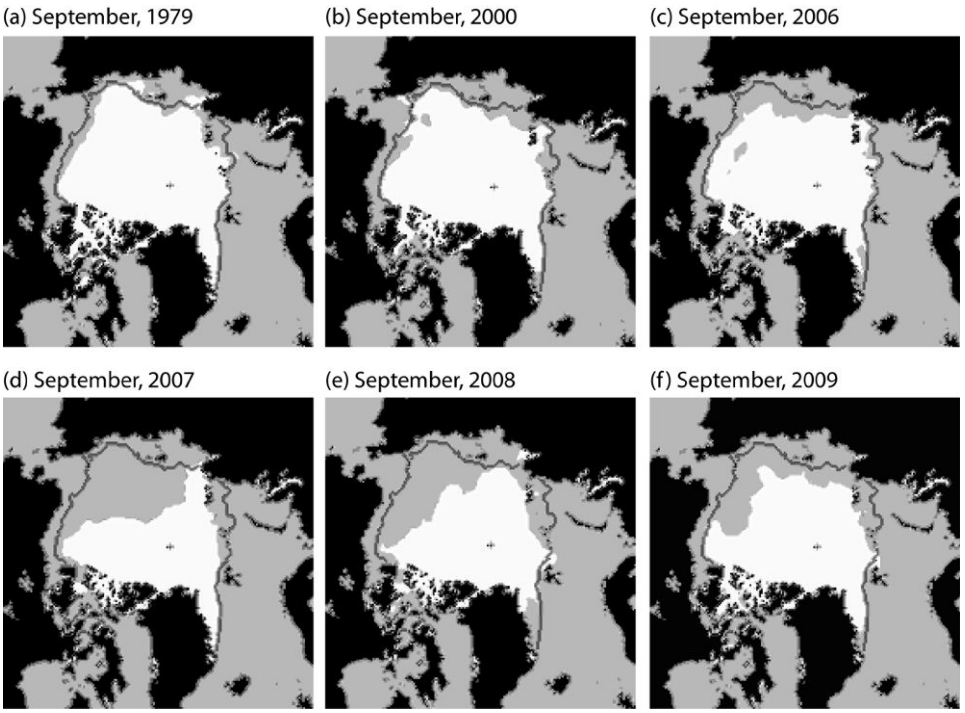


Figure 11.27. September monthly sea ice extent in the Arctic Ocean for (a) 1979, (b) 2000, (c) 2006, (d) 2007, (e) 2008, and (f) 2009, showing the reduction of Arctic sea ice in summer over three decades. Ice extent, shown in white, is the region where concentration is greater than 15%. The thick, gray line is the median ice edge for September evaluated for the period 1979–2000 (figure based on images obtained from the Ice Index Archive available through the NSIDC website).

to speculate in the early 2000s that the Arctic Ocean could be ice-free by 2050. What happened in 2007 took almost all polar and climate scientists by surprise as the area fell to 2.8×10^6 km^2, little more than half its value as recently as 1996. Figure 11.27 shows September ice extent maps for 1979, 2000, and then every year between 2006 and 2009. Compared with all previous years, in 2007 wide areas of Arctic Ocean north of the Bering Strait opened to the atmosphere for the first time in living memory, reaching about 500 km from the North Pole. This is what caused the deep spike in Figure 11.26b when the area of remaining ice dropped to just over 4 million square kilometers. In 2007 the ice disappeared from the Canadian islands and the North-West Passage opened up for shipping although in that year the northeast passage remained closed.

The year 2008 witnessed the second lowest recorded summer minimum, and in 2009 the minimum value rose again, returning closer to the long-term trendline, implying that this was not the start of a runaway melting of Arctic ice. Nonetheless some different new parts of the Arctic Ocean opened up each year,

such as the northeast passage route mentioned earlier. Once multiyear ice has melted from a region, the first-year ice that replaces it the next year is probably less stable to resist melting. We might expect this to accelerate the melting trend, although at present the unexpected severity of what has occurred in the last 3 years leaves sea ice science uncertain about what will happen next. Those reading this in subsequent years are encouraged to explore Sea Ice Index online resources for themselves to find out how this trend develops. Nonetheless we can be certain of one thing; without satellite data it could have taken years to discover the large changes in Arctic Ocean ice patterns and even longer to gain detailed information on which to base improved understanding. As it is, the knowledge accumulating in the Sea Ice Index and other remotely sensed datasets can be expected to lead more quickly to a more confident understanding of the processes that seem to be heading towards removing the summer ice cap of the northern hemisphere. Here is a clear demonstration of how the availability of satellite data has transformed our access to immediate knowledge about how our planet is changing.

It is too early to fully understand what difference the change in summer ice extent will make to the wider oceanography of the Arctic Ocean, although already a scientific review of the issues has been published (Perovich and Richter-Menge, 2009). However, other ocean sensors on satellites can explore different aspects of the newly revealed Arctic Ocean, without the need to plan cruises or launch buoys. For example, SST sensors were able to measure temperatures in the Arctic Ocean which reached at least 10°C during 2007. Satellites are well suited to monitor the unexpected changes taking place in our ocean environment as a consequence of global warming.

11.5 TIDES, SEA LEVEL, SURGES, AND TSUNAMIS

11.5.1 A surveyor's benchmark in the sky

Satellite radar altimeters are versatile instruments with a wide range of applications that have discovered interesting and important things about the ocean. Previous chapters have shown their capacity to observe and track ocean eddies and fronts, or to reveal large-scale planetary waves and thus contribute uniquely to the science of dynamical oceanography, Their ability to measure significant wave height and wind speed is shown in Chapters 8 and 9 to have more practical and immediate significance for mariners. Given the emphasis in this chapter on ocean phenomena with a human impact, we now focus attention on the place where the ocean has the potential to affect human civilization most of all; that is, along the coastlines of the world. Most countries with a shoreline can point to events in fairly recent history when sea level has risen above its normal tidal range to flood low-lying land, leaving devastation and human tragedy in its wake.

The factors that directly cause such raised sea level events are storm surges combined with high tides, or tsunamis, while more gentle but longer lasting fluctuations in mean sea level over months to years can subtly exacerbate these extreme

events. Secular changes in sea level over decades or centuries of recorded history have inexorably changed the position of coastlines, flooding towns or leaving once-busy ports stranded kilometers inland. Now that we recognize our planet has been set on course to warm rapidly, we expect worldwide sea levels to continue rising, although there is uncertainty about how rapidly. Will the 3.1 ± 0.7 mm/yr increase recorded between 1993 and 2003 (Bindoff *et al.*, 2007) be maintained, reduced, or increased, and how does it vary geographically? The challenge to understand the processes behind both sudden sea level events and slow change has grown beyond mere scientific curiosity to an essential requirement that will allow civilization to prepare for what is to come.

Why should satellite altimetry have a part to play in this? After all, it is sea level at the coast with which we are most concerned, and altimetry is less reliable in shallow seas close to land (although Chapter 13 notes it is likely to improve). The reason altimetry is not only useful but has revolutionized the way we observe and understand these fast and slow sea level processes is that it measures reliably in the open sea and throughout the global ocean. Most coastal measurements of sea level, however accurate and carefully leveled-in to regional geodetic networks, are representative of only their immediate surroundings. If we are to understand why sea level suddenly or gradually changes at the coast we need to know what is happening offshore and beyond, into the deep ocean. Is a change detected by a tide gauge a purely local phenomenon or part of a wider pattern? For decades coastal geodesists have grappled with such questions, struggling to compare levels at widely isolated sea level monitoring stations which are not leveled to the same geodetic benchmarks. Satellite altimetry using precise orbit tracking effectively provides a "benchmark in the sky". This changed the face of tidal science and marine geodesy in the decade following the launch of the T/P mission in August 1992.

A general introduction to the scientific principles of measuring sea surface height using a satellite altimeter can be found in chapter 11 of *MTOFS*, while Fu and Cazenave (2001) provide a comprehensive account of satellite altimetry in general, with particular chapters devoted to ocean tides (Le Provost, 2001), sea level change (Nerem and Mitchum, 2001), and geodesy (Tapley and Kim, 2001). Important information for understanding the applications of altimetry discussed here is that instruments such as T/P and its successors the Jason series, circling in an altimetry-optimized, non-Sun-synchronous orbit, can measure height relative to a reference ellipsoid with an absolute accuracy of 4.2 cm for an individual record (integrated over 1 s) and approaching 2 cm when averaged over a few hundred kilometers along-track. The ground track has been exactly repeated every 9.9156 days since 1992 so that precise levels can be derived from long-term averages.

After several years of T/P operation it became possible to perform tidal analysis on the altimeter sea surface height (SSH) record in order to obtain detailed information about the global spatial distribution of harmonic constituents of astronomical tide (Le Provost, 2001). This in itself was a remarkable achievement since previously tides over the deep ocean were rather poorly known. Now, astronomical tide height (i.e., the change in height caused only by the response of the sea and the solid Earth to the gravitational attraction of the Sun and Moon) can be predicted to 1 cm or 2 cm

with confidence almost anywhere in the world. This is a necessary prerequisite to distinguishing sea level displacements caused by meteorological forcing agents such as pressure and winds from the astronomical tide. It is also essential when averaging over longer periods to remove the tide-induced signal when determining the mean sea level about which tides oscillate. If the predicted tidal signal could not be confidently subtracted from the altimeter record the residual tidal signal would tend to dominate the record and make it harder to detect any long-term temporal variability of mean sea level and its spatial distribution.

11.5.2 Mean sea level

It took a number of years after the launch of T/P before any results of what it could tell us about mean sea level (MSL) began to emerge. That is because it takes time and patience to calibrate the sensor and to apply various corrections in the processing chain to produce the altimeter distance measurement. Similarly the model for precise orbit calculation takes time to be refined, and several years were needed to derive new tidal models. An additional requirement is to test for any drift in the bias of the altimeter signal, which might not affect its application to ocean dynamics very much, but is problematic for MSL changes. This was done partly by comparison with tide gauge records (Mitchum, 1994, 1998). Here comparison delivered new information about land movements affecting tide gauges, which needed to be fed back to the calibration process (Mitchum, 2000). It was very gratifying for the T/P project and its science team that T/P performed so well and also continued in operation well beyond its design life, allowing satisfactory overlap with the follow-on mission, Jason-1.

A more fundamental reason for not expecting information about global MSL trends from the first few years of the T/P mission is that they would not be meaningful. There are seasonal and interannual variations in MSL, associated with variability in ocean circulation and the distribution of water masses. Consequently the values of MSL evaluated from every T/P 10-day cycle produce quite a noisy record. Additionally ENSO events are found to make a big impact on global MSL, implying that multidecadal records are needed to confidently detect climate trends. It was therefore very important for Jason-1, launched in 2002, to be intercalibrated with T/P during a 6-month period when they each flew over the same ground track with an overpass time difference of just 70 s. After 10 months of calibration Jason-1 was confirmed to be performing equally as well as T/P (Ménard *et al.*, 2003). Improvements in orbit tracking mean that Jason-1's orbit is now believed to have a radial accuracy of 1 cm (Lutchke, 2003).

Global trends of MSL

After much careful analysis it was demonstrated that the MSL records of T/P and Jason-1 could be combined almost seamlessly (Leuliette *et al.*, 2004). This paper became the primary source for satellite-derived information about MSL trends in the *IPCC Fourth Assessment Report* (Bindoff *et al.*, 2007). Figure 11.28 shows the

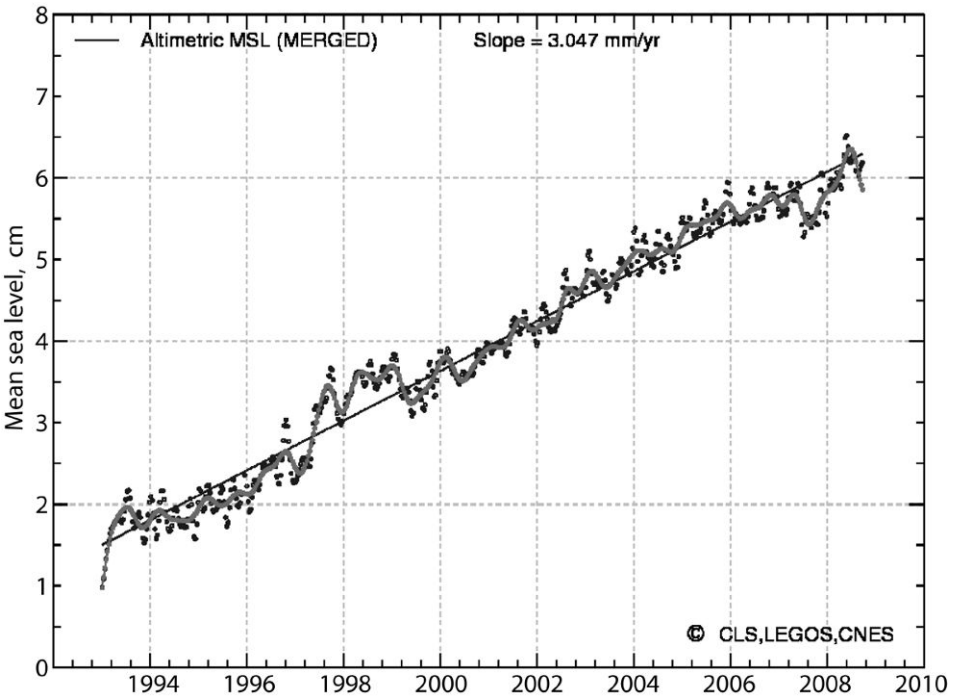


Figure 11.28. Global mean sea level from the multimission SSALTO-DUACS data altimetry dataset. Seasonal variations have been removed and corrections for inverse barometer effects have been applied (figure copyright CLS/LEGOS/CNES and was obtained from the AVISO website at *http://www.aviso.oceanobs.com/en/news/ocean-indicators/mean-sea-level/index.html*).

trend of global MSL from 1993 to 2008, extended a further 4 years beyond that available to the IPCC (Nerem *et al.*, 2007). The dots are from individual 10-day repeat cycles, and the solid line is a smoothed version of this record. The straight line is the average trend over the span of data. This presents an unambiguous picture of global MSL increasing at an average rate of 3.0 ± 0.4 mm/yr. There is short-term noisy behavior, and the El Niño of 1997–1998 stands out clearly as a positive MSL event, but over the 15-year data span it can be seen in context.

Nonetheless, care must be exercised in using a figure like this. We should resist the temptation to extrapolate it forward in time to make a forecast. As it represents nothing more than observations, only time will tell how the curve will extend through the next 15 years. Forecasting a rise in MSL should instead be based on models of the ocean–atmosphere system in which hypotheses are made about factors driving global warming, and their consequences for the future of MSL are evaluated. The strength of altimetry is that it provides a definite statement about present and recent sea level trends against which the predictions of climate models can be tested (see, e.g., Leuliette *et al.*, 2006). Moreover, since the geographical distribution of the trend is also readily available from the altimetry record (as shown in Figure 11.29),

this provides a richer set of information with which to confront the models, enabling them to be refined. Note in this figure how there is considerable spatial variability of MSL change, including some regions where it is falling.

Attempts to account for the observed trend in terms of the factors expected to control sea level provoked some interesting scientific debate (see, e.g., Lombard *et al*., 2006), leading to the eventual consensus that, over the decade 1993–2003, thermosteric sea level rise (caused by the expansion of warming sea water in the upper ocean) was about 50% that measured by the altimeter and the rest must be accounted for by additional water mass entering the ocean, mainly from land ice melt. Garcia *et al*. (2007) approached this issue from a different direction, using a combination of T/P and GRACE data. Whereas T/P data measure the distribution of absolute height change, GRACE data, by tracking small changes in gravity, measure the change in mass of the ocean water column (Nerem *et al*., 2003). By this means it provides information about additional water that enters the sea from other sources, presumably mainly melting land ice. The difference is assumed to be the thermosteric effect.

The altimeter record of global MSL change has also been joined to the historic tide gauge record, providing a longer observational record running back many decades against which predictions of climate models can be tested. This is beneficial since it includes periods when the trend of MSL was less than today (Church and White, 2006). This evidence of accelerating MSL offers a more challenging test for predictive models.

Regional changes in MSL

While headline-writers tend to focus on the global mean sea level trend, it is evident from Figure 11.29 that changes of MSL are highly variable geographically. It is therefore of considerable public importance that ocean scientists should determine regional patterns of recent and current MSL trends, which can inform planning decisions about local risks of future changing sea level and flooding. Altimetry can provide such information at a spatial resolution of 100 km to 200 km. The Mediterranean Sea provides an example of a region where the patterns of MSL perturbations have been mapped, either using altimeter data alone (Larnicol *et al*., 2002) or in conjunction with tide gauge records (Fenoglio-Marc, 2002; Tsimplis *et al*., 2008).

It is not only regional, long-term trends of MSL, but also patterns of rise and fall related to climatically varying factors that can be applied to local management of flooding threats. For example, Woolf *et al*. (2003) analyzed 9 years of monthly mean T/P sea level data in 1° grid squares over the North Atlantic, finding a seasonal signal with a range of over 120 mm in some places and which varied spatially at lengthscales down to 200 km. When the seasonal signal is removed, the residuals represent interannual variability, some of which may be a result of limited sampling, but most of which represents actual conditions, with standard deviations of up to 100 mm in some places. Perhaps the most interesting result is that in winter months and in the northeast part of the region there is a high correlation between

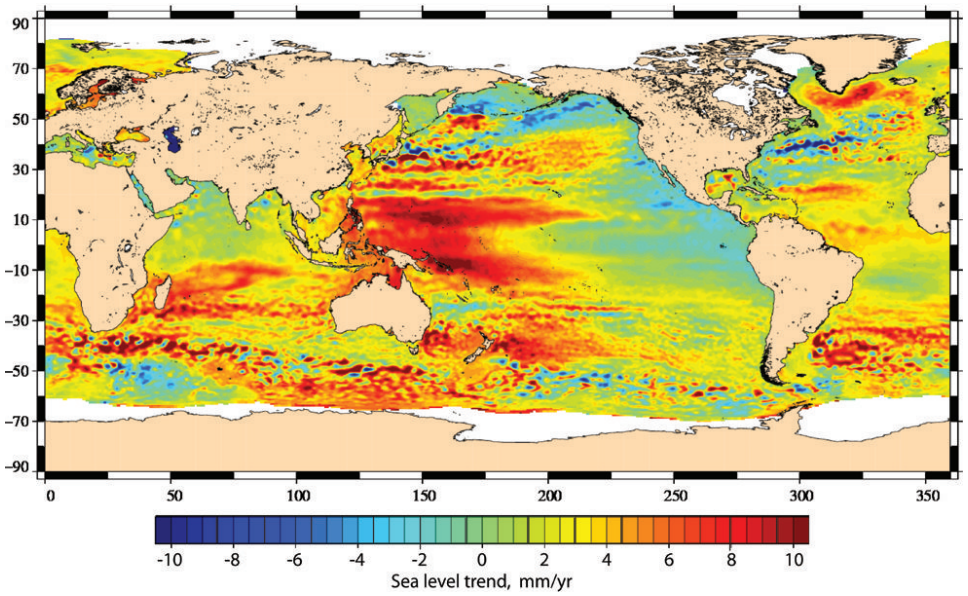


Figure 11.29. Local trends of global mean sea level from the multimission SSALTO-DUACS data altimetry dataset for the period October 1992 to January 2008. Seasonal variations have been removed and corrections for inverse barometer effects have been applied (image copyright CLS/LEGOS/CNES and was obtained from the AVISO website at *http://www.aviso.oceanobs.com/en/news/ocean-indicators/mean-sea-level/index.html*).

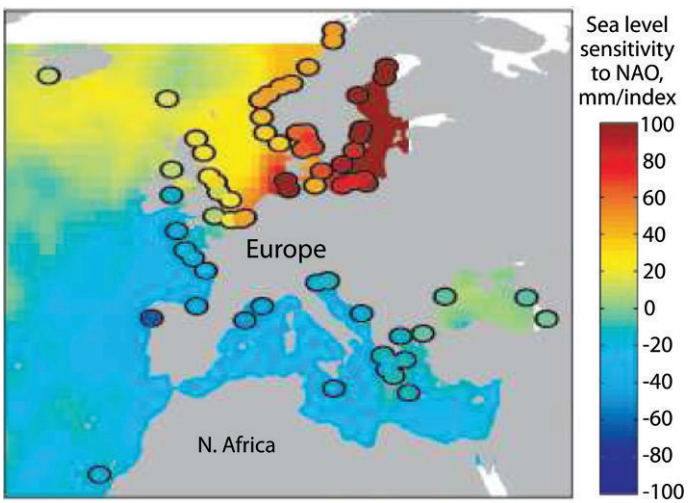


Figure 11.30. Sensitivity (rate of change per index) of wintertime sea level to the North Atlantic Oscillation. Values in circles are calculated from tide gauges at those locations. The remaining values (on an identical scale) are derived from a $1° \times 1°$ climatology of sea level from TOPEX (image provided by David Woolf after Woolf and Gommenginger, 2008).

interannual variations and the North Atlantic Oscillation (NAO) index. Figure 11.30 illustrates this (Woolf and Gommenginger, 2008). Note, for example, that in the Baltic Sea there is a strong response of MSL to changes in the NAO index. A similar analysis of tide gauge records produced the results in the small circles around the coast. These agree well with altimetry, which shows how the relationship varies offshore. Analyses such as this can be very helpful in explaining why sea level may be higher or lower than normal for several months, even years, at a time, and are a reminder that not all nonseasonal sea level change is part of the secular trend associated with global warming.

11.5.3 Storm surges

Storm surges occur when strong meteorological events cause the sea level to rise or fall sufficiently to exceed the normal bounds of highest or lowest astronomical tide to which a particular coastline, and the human infrastructure associated with it, are adapted. Extreme low levels, negative surges, can be problematic for shipping, especially deep-draught vessels in shallow passages such as the Strait of Dover or the Torres Strait. Extreme high levels, or positive surges, can cause widespread flooding and loss of life; their prediction is a major activity in the many parts of the world where they pose a threat.

Surges are caused by weather systems through two mechanisms. First, the sea acts as an inverted barometer; sea level rises 1 cm for every reduction of 1 mbar in atmospheric pressure so the effect is greatest at the center of a moving depression or tropical storm. Second, there is a dynamic response of the ocean to movement of the low-pressure center and also to moving wind stress patterns associated with the storm. The dynamic response is difficult to predict. For example, in shallow shelf seas there can be a resonant response of the ocean depending on the track and speed of meteorological forcing relative to the free propagation speed of longwaves in that particular depth of water. For tropical cyclones there are different dynamic responses, depending partly on the bathymetry of the continental slope and shelf as land is approached.

Thus while it is difficult enough for meteorologists to predict the path of a major midlatitude depression or a tropical cyclone, it is even harder to predict what effect the storm will have on sea level along the coast. Therefore any opportunity to monitor the sea level perturbation of a storm surge before it reaches land delivers valuable data for predictive models on which public warnings are based. Altimetry can provide such data if the altimeter track crosses fairly close to the center of the storm. The timescale for sea level at a point to rise to its maximum and fall again as a storm surge passes may be quite short if the storm is fast-moving, typically measured in hours rather than days. This means that the sparse space-time sampling of a single altimeter is unlikely to encounter the full magnitude of a particular surge, even if it manages to observe it at all, and so it is not used as an operational tool at present, although there are examples reported where altimetry has detected storm surges (Woolf and Gommenginger, 2008) and, most notably, Hurricane Katrina (Scharroo *et al.*, 2005), where three different altimeters showed the surge height at different stages of development.

If altimetry is to become an operational tool for tropical cyclone surge monitoring then a finer grid of sampling is needed in space and time. This could be provided partly by the use of a swath altimeter (see section 11.5.5 in *MTOFS*), which would be able on occasions to map sea level distribution in two dimensions across the storm. To be sure of encountering the surge at a useful stage for assimilation into models it would need a constellation of altimeters to be in place, as discussed in Section 8.7 (Allan, 2006). However, for the type of storm surge whose behavior is constrained as much by topography and bathymetry as by the storm position, which is the case for North Sea surges, altimetry has already been used in the development and validation of a regression-type, surge-forecasting model (Høyer and Andersen, 2003). Here there are sufficient altimeter tracks over the sea, especially from the long repeat of the ERS orbit, for many surge events to be found in the historical record. This provides information about surge behavior offshore, supplementing the coastal tide gauge record for tuning the regression model. Thus although the forecasting model is based on real-time tide gauge records, it can make predictions offshore because of the information already supplied from archived altimetry data.

11.5.4 Tsunamis

As this chapter is written, memories are still fresh of the tragic events of December 26, 2004 when a tsunami, generated by an earthquake of magnitude 9.3 off Sumatra, devastated a widespread number of locations in South East Asia and across the Indian Ocean, resulting in the loss of more than 200,000 lives in at least eight nations. Ocean scientists are asked what can be done to provide a reliable and practical warning system against future tsunamis. Is there a role for ocean-observing satellites?

Tsunamis are caused when a sea bed disturbance creates a perturbation of the overlying water column. For example, if the sea bed is displaced upwards or downwards over an extended region, the water column above the disturbance must move with it, raising or lowering the sea surface. Immediately a pressure imbalance with the surrounding water is created, causing the disturbance to radiate out as a largely barotropic propagating wave. Outside the source region this typically becomes a surface displacement of between 0.1 m and 1 m depending on the severity of the seismic event, which reduces with radial spreading of the energy away from the earthquake zone. The first displacement may be positive or negative and is followed by a series of undulations with a wavelength of order 100 km. The speed of propagation of this wave front will be close to the barotropic longwave speed of $\sqrt{(gh)}$, where h is ocean depth, which equates to 720 km/h in water of depth 4,000 m. In the deep ocean the wave rises and falls very gently over several minutes and would be barely detectable by someone on a ship. It is when the wave front arrives at the continental slope that its amplitude grows enormously by up to two orders of magnitude as its speed slows and the energy it carries is compressed.

The possibility of detection and warning depends first on the monitoring of seismic events of a sufficient magnitude to create a problematic tsunami. Being

aware of such an occurrence is not enough to issue a warning since not all seismic events disturb the ocean. Before alerting the population it is important to know that a tsunami is genuinely propagating in a known direction with sufficient magnitude to cause flooding, since false warnings can create unnecessary havoc and lead to complacency when a warning is justified. Models are available to describe the timing of arrival after the seismic event, but to become fully reliable there needs to be a way of detecting the tsunami while it is still out at sea and perhaps hours away. Bottom-mounted pressure gauges have the sensitivity to detect small disturbances of a few centimeters, and must surely form the core of any operational tsunami-warning system. However, satellite altimetry can also contribute, with the potential to provide additional data that will help to refine inputs to tsunami propagation models, such as the location and extent of the source region.

Large tsunamis had previously been detected by altimetry (Okal *et al*., 1999) although it was evident from that study that with few altimeters in orbit there is an element of luck attached to whether an altimeter track will intersect with a tsunami wave front. Satellites did pass over the tsunami of December 26, 2004, as reported to the scientific community within less than a month (Gower, 2005). In fact there were five different overpasses by four different satellites. After further analysis Gower (2007) showed that the tsunami was revealed clearly on the Jason-1, T/P, and Envisat RA-2 overpasses, within 1:53 h, 2:0 h, and 3:15 h, respectively, of the earthquake, by simply comparing the sea surface height anomaly (SSHA) from each sensor with previous overpasses along that orbit. The tsunami profile appeared close to the location in which numerical models of tsunami propagation placed it. This approach was not able to identify the tsunami on two GFO overpasses 7:00 h and 9:00 h after the earthquake by which times the main wave front had reached 40°S and 50°S, respectively, and was a lot weaker. However, after further analysis by Ablain *et al*. (2006) in which mesoscale eddy perturbations of the SSHA for a period of ±20 days either side of December 26 were mapped objectively (Le Traon *et al*., 1998) and subtracted from tsunami orbits, its signature was found on all five overpasses.

The latter approach would appear to be more effective to use in an operational monitoring system because it is able to work in regions where eddy energy is greater and for tsunamis that are weaker than was the case for the signatures acquired within 4 hours of the December 26 Earthquake. Figure 11.31 shows the Jason-1 signature in the SSHA residual after applying the Ablain *et al*. (2006) processing. In the *upper panel* is the orbit track relative to the height field predicted by the tsunami model of the French Commissariat à l'Energie Atomique (CEA). In the *lower panel*, altimeter height is compared with model height predictions before and after altimeter signatures had been used to refine initial conditions for the model. The improvement obtained shows that if this could be done in a future monitoring system it could improve forecast of the detailed profile of an advancing tsunami before it reaches the coast. There is therefore evidence that altimetry can play a very useful, though not pivotal, role in future tsunami-warning systems, but to be effective there would have to be more frequent coverage of the globe by a constellation of satellites (Allan, 2006), as discussed in Section 8.7.

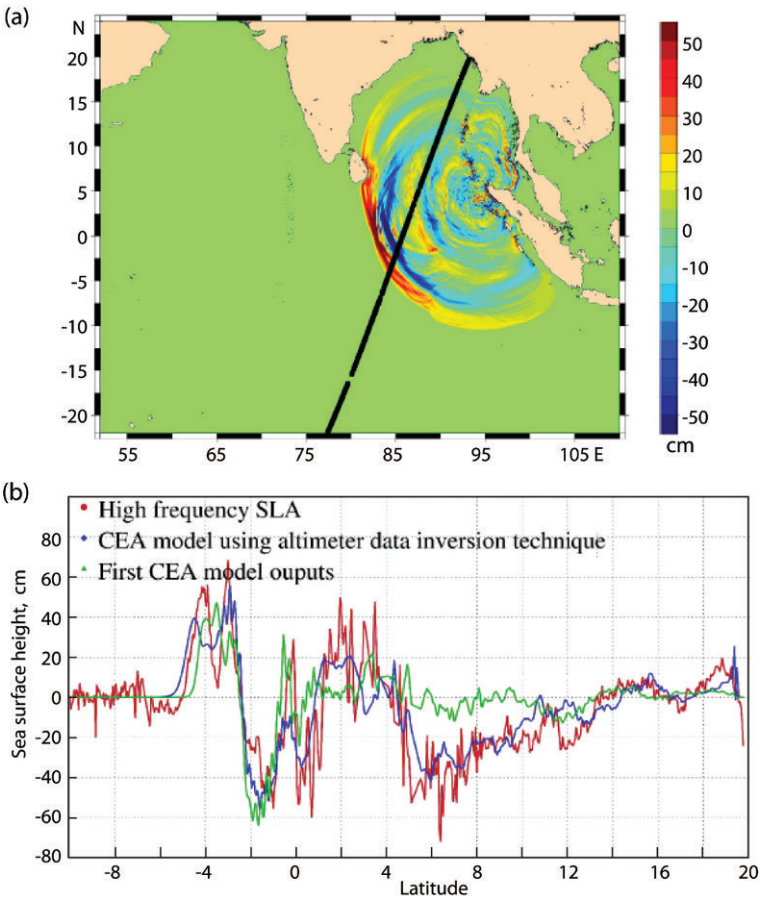


Figure 11.31. *Upper panel*: Tsunami wave heights, as computed by the CEA model 1:53 h after the earthquake with the Jason-1 pass superimposed. *Lower panel*: 20 Hz sea level anomaly (red), and CEA model output from the initial (green) and revised (blue) run (figure copied from Ablain *et al.*, 2006.)

11.6 CONCLUSION

The topics discussed in this chapter represent only some of the most prominent ways in which satellite data can assist mankind in the challenge of learning to live in harmony with the unpredictable and sometimes extremely variable natural marine environment. It is instructive to review the particular capabilities of satellite ocean data which make them effective for observing these various phenomena, and to identify ways of presenting them that will facilitate their use in operational monitoring and forecasting systems.

Of fundamental importance is the regular acquisition of raw data from sensors at a spatial and temporal resolution appropriate for the ocean process being observed, and their processing into a time series of gridded sea surface data products. If these are to be useful to mainstream ocean scientists they need to be well-calibrated, with error estimates that are validated throughout the lifetime of the sensor. Depending on the sensor, the character of the data, and whether it depends on

cloud-free conditions, it is normally important to aggregate individual synoptic views into composites over longer spans of time. When such records have been obtained over several years then climatologies should be evaluated, which then allows anomalies to be produced along with the primary variable.

It is evident from hindsight that many applications of processed datasets such as SST, SSHA, and chlorophyll were not envisaged when datasets first started being produced. In the early days of satellite oceanography specialist data products for particular applications (such as the study of ENSO) had to be generated specially by remote-sensing experts, or they would not have been available. Now the situation is different since many derived data products are being routinely produced, validated, archived, and when necessary reprocessed. They are therefore immediately available in a Web-accessible archive to be consulted when the need arises. This includes the whole suite of optical and biogeochemical products that can be derived from ocean color sensors and surface currents or kinetic energy derived from SSHA.

Another common element in the applications discussed in this chapter is the benefit of synergy derived from using different types of data products in a complementary way. For example, comparing the wind field with altimeter-derived products and SST allows the causes and effects of ocean-dynamical interaction with the atmosphere to be understood. The use of common resampling grids for composite products from different sensor types facilitates intercomparison between datasets (e.g., by aligning Hovmöller plots). If datasets of different products have different grid sizes and different integration time intervals, the task of correlating, say, 20 years of archived data from two different sources may be prohibitively time-consuming. However, if they can readily be matched on the same grid they will be widely used for research and will encourage the development of operational applications, especially if they can be accessed in similar common formats such as NetCDF or HDF.

A fourth element for enhancing the applicability of satellite ocean datasets has also emerged fairly recently. This is the important benefit of producing merged data products of a particular type, derived from several different sensors. Examples are multimission, altimetry, SSH products and derivatives compiled by merging data from several different altimeters (as produced by SSALTO-DUACS), and analyzed SST fields produced by combining SST products from infrared and microwave radiometers.

The systems needed to generate such merged products in near-real time for operational applications are discussed further in Section 14.4 in the context of operational ocean-forecasting systems now becoming available. These incorporate satellite data and *in situ* measurements into numerical models of ocean-dynamical processes. While it might seem that the unique contribution of satellite data will be hidden by assimilating them into ocean models, this is to miss the point. It is time to move on from the stage where the ambition of a space agency was limited to producing their own individual "branded" data products from their particular sensor. While there is nothing wrong with such competition to raise the standard of data quality, it is even more important that agencies collaborate rather than compete. For example, it is globally more efficient for agencies to reach agreement to launch

satellites and sensors that complement each other, thus ensuring that a full spectrum of sensor types is maintained in orbit and delivering data without data gaps.

No matter how successful ocean scientists become in detecting the onset of potentially catastrophic conditions of El Niño, monsoon, ice, or tsunami, if such knowledge is to translate to benefits in people's lives it has to be linked to a responsible structure of governance that makes effective and safe use of the forecasts. This is already established in some cases (e.g., ice warnings in certain parts of the world). But what should be the response if a confident forecast is made of an El Niño? Who will advise farmers on mitigating crop damage? On a more rapid timescale, how should local fishing villages or beach resorts be warned of an approaching tsunami? These are tasks for governments, but the global reach of remote sensing means that global-scale collaboration is needed. It is gratifying to realize that satellite oceanography has reached a stage where it can offer positive benefits in the lives of the general population, a kind of payback to those who indirectly have funded research and development that have brought us to this point. But those benefits will only be delivered when local government agencies are fully connected into ocean-forecasting systems.

Within the science community agencies such as the World Climate Research Program (WCRP) and the Joint Committee for Oceanography and Marine Meteorology (JCOMM) have established administrative structures for collaboration and agreements on common standards for data and practices that will facilitate different agencies working together. In the case of the WCRP the aim is to generate a reliable archive of climate records (see Section 14.6), and for JCOMM it is to provide real-time observations needed to support operational ocean-forecasting models (see Sections 14.2–14.6). Following international agreements and principles established by the U.N. Committee on Peaceful Uses of Outer Space, the Committee on Earth Observing Satellites[4] (CEOS) was created in 1984 and provides a coordinating framework to ensure U.N. principles are put into practice by all international players in Earth-observing satellites and data.

One aspect of this is concerned with disaster management support. Most agencies are now committed to following the "International Charter on Space and Major Disasters". When they are informed of a disaster for which satellite data can make a useful contribution, both in the immediate crisis and during the ensuing recovery phase, an agency is expected to follow an agreed emergency procedure for acquiring, rapidly processing, and disseminating appropriate satellite data. For example, this may mean switching planned SAR data capture from a science program to monitor a disaster area. There are typically several activations of the charter each month, mainly but not always relevant to land remote sensing.[5]

Recently the international oversight for this type of activity has shifted to the Group on Earth Observations (GEO), a voluntary partnership of governments and international organizations with mandates in Earth observation, who recognize

[4] The CEOS website is at *http://www.ceos.org/*

[5] See *http://www.disasterscharter.org/web/charter/activations* for news of recent activations of the charter.

"that international collaboration is essential for exploiting the growing potential of Earth observations to support decision making in an increasingly complex and environmentally stressed world."[6] This grew out of the 2002 World Summit on Sustainable Development, and its main achievement has been to establish the Global Earth Observing System of Systems (GEOSS). The role of this is to draw together a wide variety of other international activities and enable them to fit together into an effective structure that facilitates the use of Earth observations in their widest sense, including satellites and *in situ* methods, in order to benefit mankind through scientific and operational applications. CEOS is now integrated within the GEO umbrella as the main technical focus for integration of Earth-observing satellites. Issues associated with how ocean remote sensing can be made to work more effectively for the good of human society are discussed further in Chapters 14 and 15.

11.7 REFERENCES

Ablain, M., J. Dorandeu, P.-Y. L. Traon, and A. Sladen (2006), High resolution altimetry reveals new characteristics of the December 2004 Indian Ocean tsunami. *Geophys. Res. Letters*, **33**(L21602), doi: 10.1029/2006GL027533.

Adler, R. F., J. Susskind, G. J. Huffman, D. Bolvin, E. Nelkin, A. Chang, R. Ferraro, A. Gruber, P.-P. Xie *et al.* (2003), The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.*, **4**, 1147–1167.

Allan, T. (2006), The story of GANDER. *Sensors*, **6**, 249–259.

Allen, R., J. Lindesay, and D. Parker (1996), *El Niño Southern Oscillation and Climatic Variability* (405 pp.). CSIRO Publishing, Collingwood, Victoria, Australia.

Anderson, D. L. T., and M. K. Davey (1998), Predicting the El Niño of 1997/98. *Weather*, **53**, 303–309.

Askne, J., and W. Direking (2008), Sea ice monitoring in the Arctic and Baltic Sea using SAR. In: V. Barale and M. Gade (Eds.), *Remote Sensing of the European Seas* (pp. 383-398). Springer-Verlag, Dordrecht, The Netherlands.

Barnston, A. G., M. H. Glantz, and X. He (1999), Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997–98 El Niño episode and the 1998 La Niña onset. *Bull. Am. Meteorol. Soc.*, **80**, 217–243.

Bindoff, N. L., J. Willebrand, V. Artale, C. A. J. Gregory, S. Gulev, K. Hanawa, C. L. Quéré, S. Levitus, Y. Nojiri *et al.* (2007), Observations: Oceanic climate change and sea level. In: S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller (Eds.), *Climate Change 2007: The Physical Science Basis* (Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 385-432). Cambridge University Press, Cambridge, U.K.

Bonjean, F., and G. S. E. Lagerloef (2002), Diagnostic model and analysis of the surface currents in the tropical Pacific Ocean. *J. Phys. Oceanogr.*, **32**(10), 2938–2954.

Carsey, F. (1992), *Microwave Remote Sensing of Sea Ice* (Geophysical Monograph Series, 478 pp.). American Geophysical Union, Washington, D.C.

[6] See the GEO website at *http://earthobservations.org/index.html*

Casey, K. S., and P. Cornillon (1999), A comparison of satellite and in situ-based sea surface temperature climatologies. *J. Climate*, **12**, 1848–1863.

Cavalieri, D. C., P. Gloersen, and W. J. Campbell (1984), Determination of sea ice parameters with the NIMBUS-7 SMMR. *J. Geophys. Res.*, **89**(D4), 5355–5369.

Cavalieri, D. C., K. M. St. Germain, and C. T. Swift (1995), Reduction of weather effects in the calculation of sea ice concentration with the DMSP SSM/I. *J. Glaciology*, **41**(139), 455–464.

Cavalieri, D. C., J. Crawford, M. R. Drinkwater, D. Eppler, L. D. Farmer, R. R. Jentz, and C. C. Wackerman (1991), Aircraft active and passive microwave validation of sea ice concentration from the DMSP SSM/I. *J. Geophys. Res.*, **96**(C12), 21989–22009.

Changnon, S. A., and G. D. Bell (2000), *El Niño, 1997–1998: The Climate Event of the Century* (215 pp.). Oxford University Press, New York.

Charrassin, J.-B., M. Hindell, S. R. Rintoul, F. Roquet, S. Sokolov, M. Biuw, D. Costa, L. Boehme, P. Lovell, R. Coleman *et al.* (2008), Southern Ocean frontal structure and sea-ice formation rates revealed by elephant seals. *Proc Nat. Acad. Sci. U.S.A.*, **105**(33), 11634–11639.

Chavez, F. P., P. G. Strutton, G. E. Friederich, R. Feely, and G. C. Feldman (1999), Biological and chemical response of the equatorial Pacific Ocean to the 1997–98 El Niño. *Science*, **286**, 2126–2131.

Church, J. A., and N. J. White (2006), A 20th century acceleration in global sea-level rise. *Geophys. Res. Letters*, **33**(L01602), doi: 10.1029/2005GL024826.

Delcroix, T., J.-P. Boulanger, F. Masia, and C. Menkes (1994), Geosat-derived sea-level and surface-current anomalies in the equatorial Pacific, during the 1986–89 El Niño and La Niña. *J. Geophys. Res.*, **99**, 25093–25107.

Donlon, C. J., I. S. Robinson, K. S. Casey, J. Vazquez, E. Armstrong, O. Arino, C. L. Gentemann, D. May, P. Le Borgne, J.-F. Piolle *et al.* (2007), The Global Ocean Data Assimilation Experiment (GODAE) High Resolution Sea Surface Temperature Pilot Project (GHRSST-PP). *Bull. Am. Meteorol. Soc.*, **88**(8), 1197–1213, doi: 10.1175/BAMS-88-8-1197.

Fenoglio-Marc, L. (2002), Long-term sea level change in the Mediterranean Sea from multi-satellite altimetry and tide gauges. *Physics and Chemistry of the Earth*, **27**, 1419–1431.

Fetterer, F., K. Knowles, W. Meier, and M. Savoie (2002, updated 2008), *Sea Ice Index*. National Snow and Ice Data Center, Boulder, CO, available at *http://www.nsidc.org/data/seaice_index*

Fischer, A. S., R. A. Weller, D. L. Rudnick, C. C. Eriksen, C. M. Lee, K. H. Brink, C. A. Fox, and R. R. Leben (2002), Mesoscale eddies, coastal upwelling, and the upper-ocean heat budget in the Arabian Sea. *Deep-Sea Res. II*, **49**(12), 2231–2264.

Fu, L.-L., and A. Cazenave (Eds.) (2001), *Satellite Altimetry and Earth Sciences* (463 pp.). Academic Press, San Diego, CA.

Gadgil, S., P. N. Vinayachandran, P. A. Francis, and S. Gadgil (2004), Extremes of the Indian summer monsoon rainfall, ENSO and equatorial Indian Ocean oscillation. *Geophys. Res. Letters*, **31**(L12213), doi: 10.1029/2004GL019733.

Garcia, D., G. Ramillien, A. Lombard, and A. Cazenave (2007), Steric sea-level variations inferred from combined Topex/Poseidon altimetry and GRACE gravimetry. *Pure Appl. Geophys.*, **164**, 721–731.

Gower, J. F. R. (2005), Jason-1 detects the 26 December 2004 tsunami. *EOS, Trans. Amer. Geophys. Union*, **86**(4), 37–38.

Gower, J. F. R. (2007), The 26 December 2004 tsunami measured by satellite altimetry. *Int. J. Remote Sensing*, **28**(13), 2897–2913.

Graham, N. E., and T. P. Barnett (1987), Sea surface temperature, sea surface wind divergence, and convection over tropical oceans. *Science*, **238**, 657–659.

Hayes, S. P., L. J. Mangum, J. Picaut, A. Sumi, and K. Takeuchi (1991), TOGATAO: A moored array for real-time measurements in the tropical Pacific Ocean. *Bull. Am. Meteorol. Soc.*, **72**, 339–347.

Haykin, S., E. O. Lewis, R. K. Raney, and J. R. Rossiter (1994), *Remote Sensing of Sea Ice and Icebergs* (686 pp.). John Wiley & Sons, Chichester, U.K.

Høyer, J. L., and O. B. Andersen (2003), Improved description of sea level in the North Sea. *J. Geophys. Res.*, **108**(C5), 3163, doi: 10.1029/2002JC001601.

Ji, M., and A. Leetma (1997), Impact of data assimilation on ocean initialization and El Nino prediction. *Mon. Weather Rev.*, **125**, 741–753.

Johnson, E. S., F. Bonjean, G. S. E. Lagerloef, J. T. Gunn, and G. T. Mitchum (2007), Validation and error analysis of OSCAR sea surface currents. *J. Atmos. Oceanic Tech.*, **24**, 688–701.

Kyte, E. A., G. D. Quartly, M. A. Srokosz, and M. N. Tsimplis (2006), Interannual variations in precipitation: The effect of the North Atlantic and Southern Oscillations as seen in a satellite precipitation data set and in models. *J. Geophys. Res.*, **111**(D24113).

Lagerloef, G. S. E., G. T. Mitchum, R. Lukas, and P. Niiler (1999), Tropical Pacific near surface currents estimated from altimeter, wind and drifter data. *J. Geophys. Res.*, **104,** 23313–23326.

Lagerloef, G. S. E., R. Lukas, F. Bonjean, J. T. Gunn, G. T. Mitchum, M. Bourassa, and A. J. Busalacchi (2003), El Niño Tropical Pacific Ocean surface current and temperature evolution in 2002 and outlook for early 2003. *Geophys. Res. Letters*, **30**(10), 1514.

Larnicol, G., N. Ayoub, and P. Y. Le Traon (2002), Major changes in Mediterranean Sea level variability from 7 years of TOPEX/Poseidon and ERS-1/2 data. *J. Mar. Syst.*, **33/34**, 63–89.

Le Provost, C. (2001), Ocean tides. In: L.-L. Fu and A. Cazenave (Eds.), *Satellite Altimetry and Earth Sciences* (pp. 267–303). Academic Press, San Diego, CA.

Le Traon, P.-Y., F. Nadal, and N. Ducet (1998), An improved mapping method of multi-satellite altimeter data. *J. Atmos. Oceanic Tech.*, **15**, 522–534.

Legeckis, R. (1986), A satellite time series of sea surface temperatures in the Eastern Equatorial Pacific Ocean, 1982–1986. *J. Geophys. Res.*, **91**(C11), 12879–12886.

Leuliette, E. W., R. S. Nerem, and T. Jakub (2006), An assessment of IPCC 20th century climate simulations using the 15-year sea level record from altimetry. Paper presented at *15 Years of Progress in Radar Altimetry, Venice, Italy, March 13–18*, available at *earth.esa.int/workshops/venice06/participants/1181/paper_1181_leuliette.pdf*

Leuliette, E. W., R. S. Nerem, and G. T. Mitchum (2004), Calibration of TOPEX/Poseidon and Jason altimeter data to construct a continuous record of mean sea level change. *Marine Geodesy*, **27**(1), 79–94.

Liu, W. T., and X. Xie (1999), Spacebased observations of the seasonal changes of South Asian monsoons and oceanic responses. *Geophys. Res. Letters*, **26**(10), 1473–1476.

Lombard, A., A. Cazenave, P.-Y. Le Traon, S. Guinehut, and C. Cabanes (2006), Perspectives on present-day sea level change: A tribute to Christian le Provost. *Ocean Dynamics*, **6**(5/6), 445–451.

Lubin, D., and R. Massom (2006), *Polar Remote Sensing: Atmosphere and Oceans* (xlii + 756 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Lutchke, S. B., N. P. Zelensky, D. D. Rowlands, F. G. Lemoine, and T. A. Williams (2003), The 1-centimeter orbit: Jason-1 precise orbit determination using GPS, SLR, DORIS and altimeter data. *Marine Geodesy*, **26**(3/4), 399–421.

McClain, C. R., J. R. Christian, R. S. Signorini, M. R. Lewis, and I. Asanuma (2002), Satellite ocean-color observations of the tropical Pacific Ocean. *Deep-Sea Res. II*, **49**, 2522–2560.

McPhaden, M. J. (1999), Genesis and evolution of the 1997–98 El Niño. *Science*, **283**, 950–954.

McPhaden, M. J., A. J. Busalacchi, R. Cheney, J.-R. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G. T. Mitchum *et al.* (1998), The Tropical Ocean–Global Atmosphere observing system: A decade of progress. *J. Geophys. Res.*, **103**(C7), 14169–14240.

Ménard, Y., L.-L. Fu, S. Desai, P. Escudier, B. Haines, G. Kunstmann, F. Parisot, J. Perbos, and P. Vincent (2003), The Jason-1 mission. *Marine Geodesy*, **26**(3/4), 131–146.

Mitchum, G. T. (1994), Comparison of TOPEX sea-surface heights and tide-gauge sea levels. *J. Geophys. Res.*, **99**(C12), 24541–24553.

Mitchum, G. T. (1998), Monitoring the stability of satellite altimeters with tide gauges. *J. Atmos. Oceanic Tech.*, **15**(3), 721–730.

Mitchum, G. T. (2000), An improved calibration of satellite altimetric heights using tide gauge sea levels with adjustment for land motion. *Marine Geodesy*, **23**, 145–166.

Murakami, H., J. Ishizaka, and H. Kawamura (2000), ADEOS observations of chlorophyll a concentration, sea surface temperature, and wind stress change in the equatorial Pacific during the 1997 El Niño. *J. Geophys. Res.*, **105**(C8), 19551–19559.

Murtugudde, R. G., R. S. Signorini, J. R. Christian, A. J. Busalacchi, C. R. McClain, and J. Picaut (1999), Ocean color variability of the tropical Indo-Pacific basin observed by SeaWiFS during 1997–98. *J. Geophys. Res.*, **104**, 18351–18366.

Nerem, R. S., and G. T. Mitchum (2001), Sea level change. In: L.-L. Fu and A. Cazenave (Eds.), *Satellite Altimetry and Earth Sciences* (pp. 329–350). Academic Press, San Diego, CA.

Nerem, R. S., J. M. Wahr, and E. W. Leuliette (2003), Measuring the distribution of ocean mass using GRACE. *Space Science Reviews*, **108**(1), 331–344.

Nerem, R. S., A. Cazenave, D. P. Chambers, L.-L. Fu, E. W. Leuliette, and G. T. Mitchum (2007), Comment on ''Estimating future sea level change from past records by Nils-Axel Mörner''. *Global and Planetary Change*, **55**(4), 358–360.

Okal, E., A. Piatanesi, and P. Heinrich (1999), Tsunami detection by satellite altimetry. *J. Geophys. Res.*, **104**(B1), 599–615.

Onstott, R. G., and R. Shuchman (2005), SAR measurements of sea ice. In: C. R. Jackson and J. R. Apel (Eds.), *Synthetic Aperture Radar Marine User's Manual* (pp. 81–115). U.S. Department of Commerce, Silver Spring, MD.

Parthasarathy, B., R. R. Kumar, and D. R. Kothawale (1992), Indian summer monsoon rainfall indices, 1871–1990. *Meteor. Mag.*, **121**, 174–186.

Perovich, D. K., and J. A. Richter-Menge (2009), Loss of sea ice in the Arctic. *Annu. Rev. Mar. Sci.*, **1**, 417–441.

Peterson, B. J., J. McClelland, R. Curry, R. M. Holmes, J. E. Walsh, and K. Aagaard (2006), Trajectory shifts in the Arctic and subarctic freshwater cycle. *Science*, **313**(5790), 1061–1066.

Philander, S. G. H. (1990), *El Niño, La Niña, and the Southern Oscillation* (International Geophysics Series, 293 pp.). Academic Press, San Diego, CA.

Picaut, J., E. Hackert, A. J. Busalacchi, R. Murtugudde, and G. S. E. Lagerloef (2002), Mechanisms of the 1997–1998 El Niño–La Niña, as inferred from space-based observations. *J. Geophys. Res.*, **107**(C5), doi: 10.1029/2001JC000850.

Quartly, G. D., T. H. Guymer, and M. A. Srokosz (1996), The effects of rain on Topex radar altimeter data. *J. Atmos. Oceanic Technol.*, **13**, 1209–1229.

Quartly, G. D., M. A. Srokosz, and T. H. Guymer (1999), Global precipitation statistics from dual-frequency Topex altimetry. *J. Geophys. Res.*, **104**(D24), 31489–31516.

Quartly, G. D., M. A. Srokosz, and T. H. Guymer (2000), Changes in oceanic precipitation during the 1997–98 El Niño. *Geophys. Res. Lett.*, **27**(15), 2293–2296.

Rahmstorf, S. (2006), Thermohaline ocean circulation. In: S. A. Elias (Ed.), *Encyclopedia of Quaternary Sciences* Elsevier, Amsterdam.

Ramesh Kumar, M. R., S. Sankar, K. Fennig, D. S. Pai, and J. Schulz (2005), Air–sea interaction over the Indian Ocean during the contrasting monsoon years 2002 and 2003. *Geophys. Res. Letters*, **32**(L14821), doi: 10.1029/2005GL022587.

Rees, G. (2005), *Remote Sensing of Snow and Ice* (285 pp.). Taylor & Francis/CRC Press, London/Boca Raton, FL.

Robinson, I. S. (2004), *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Ryan, J. P., P. S. Polito, P. G. Strutton, and F. P. Chavez (2002), Unusual large-scale phytoplankton blooms in the equatorial Pacific. *Prog. Oceanogr.*, **55**, 263–285.

Scharroo, R., W. H. F. Smith, and J. L. Lillibridge (2005), Satellite altimetry and the intensification of hurricane Katrina. *EOS, Trans. Amer. Geophys. Union*, **86**(40), 366–367.

Strutton, P. G., J. P. Ryan, and F. P. Chavez (2001), Enhanced chlorophyll associated with tropical instability waves in the equatorial Pacific. *Geophys. Res. Letters*, **28**(10), 2005–2008.

Subrahmanyam, B., and I. S. Robinson (2000), Sea surface height variability in the Indian Ocean from TOPEX/POSEIDON altimetry and model simulations. *Marine Geodesy*, **23**, 167–195.

Subrahmanyam, B., V. Ramesh Babu, V. S. N. Murty, and L. V. G. Rao (1996), Surface circulation off Somalia and western equatorial Indian Ocean during summer monsoon of 1992 from Geosat altimeter data. *Int. J. Remote Sensing*, **17**, 761–770.

Tapley, B. D., and M.-C. Kim (2001), Applications to geodesy. In: L.-L. Fu and A. Cazenave (Eds.), *Satellite Altimetry and Earth Sciences* (pp. 371–406). Academic Press, San Diego, CA.

Tsimplis, M. N., A. G. P. Shaw, A. Pascual, M. Marcos, M. Pasaric, and L. Fenoglio-Marc (2008), Can we reconstruct the 20th century sea level variability in the Mediterranean Sea on the basis of recent altimetric measurements? In: V. Barale and M. Gade (Eds.), *Remote Sensing of the European Seas* (pp. 307–318). Springer-Verlag, Berlin.

Wadhams, P. (2000), *Ice in the Ocean* (351 pp.). Gordon & Breach, London.

Wang, B., and Z. Fan (1999), Choice of South Asian summer monsoon indices. *Bull. Am. Meteorol. Soc.*, **80**(4), 629–638.

Webster, P. J., and S. Yang (1992), Monsoon and ENSO: Selectively interactive systems. *Quart. J. Roy. Meteorol. Soc.*, **118**, 877–926.

Weller, R. A., A. S. Fischer, D. L. Rudnick, C. C. Eriksen, T. D. Dickey, J. Marra, C. Fox, and R. Leben (2002), Moored observations of upper-ocean response to the monsoons in the Arabian Sea during 1994–1995. *Deep-Sea Res. II*, **49**(12), 2195–2230.

Woolf, D. K., and C. P. Gommenginger (2008), Radar altimetry: Introduction and application to air–sea interaction. In: V. Barale and M. Gade (Eds.), *Remote Sensing of the European Seas* (pp. 283–294). Springer-Verlag, Berlin.

Woolf, D. K., A. G. P. Shaw, and M. N. Tsimplis (2003), The influence of the North Atlantic Oscillation on sea-level variability in the North Atlantic region. *J. Atm. Ocean Sci.* (previously *The Global Atmosphere and Ocean System*), **9**(4), 145–167.

Xue, Y., T. M. Smith, and R. Reynolds (2003), Interdecadal changes of 30-year SST normals during 1871–2000. *J. Climate*, **16**, 1601–1612.

Zwally, H. J., J. C. Comiso, C. L. Parkinson, D. J. Cavalieri, and P. Gloersen (2002), Variability of the Antarctic sea ice cover. *J. Geophys. Res.*, **107**(C5), 1029–1047.

12

Internal waves

Co-authored with José da Silva[1]

12.1 INTRODUCTION

Oceanographers generally recognize that the subject of internal waves has greatly benefited from the advent of satellite observations, which is why a full chapter is devoted here to their study, expanding on the introduction already provided in section 10.10 of *MTOFS* (Robinson, 2004). In this chapter we concentrate mainly on imaging sensor observations of short-period internal waves, such as those provided by synthetic aperture radar (SAR). The topics of radar backscatter, Bragg scattering, and SAR image interpretation are discussed at some length in *MTOFS*, and for this reason we concentrate here on applications of remote-sensing observations to internal wave studies. We discuss the various imaging mechanisms for short-period, solitary-like, internal waves (commonly referred to as internal solitary waves, ISWs), because such knowledge helps with interpreting images that reveal the processes of generation, propagation, and dissipation of internal wave energy. The chapter also presents ocean color observations and related model results concerning larger scale internal waves of tidal period, which are important in a multidisciplinary context.

12.1.1 Ocean internal and interfacial waves

Internal waves (IWs) are an important part of small-scale processes in geophysical fluid flows. They occur both in the ocean and the atmosphere through the restoring action of buoyancy forces on fluid parcels displaced from their equilibrium position. In both the atmosphere and the ocean, fluid is density-stratified (i.e., $\rho = \rho(z)$) so that dense fluid underlies lighter fluid. This density gradient supports the propagation of

[1] Dr. José da Silva is associate professor at the Institute of Oceanography in the Faculty of Sciences, University of Lisbon, Portugal.

Figure 12.1. Slick bands associated with internal waves off Cape Cod (Massachusetts, U.S.A.). In the background, Cape Cod and Herring Cove beach can be seen, as well as Race Point lighthouse. Slicks such as these can be seen in satellite images, such as SAR images, revealing the full spatial structure and scales of short-period internal waves (photo taken by the co-author on August 30, 2006, at 15:30 local time).

internal waves, and a simple example is interfacial waves on the steep density gradient at the interface between two layers of a stably stratified fluid. Akin to the air–sea interface, when this interface is disturbed waves radiate away horizontally along the interface, producing subtle roughness patterns at the surface that allow them to be detected by remote-sensing methods and even by the eye (see Figure 12.1). Internal waves, or internal gravity waves, are designated as such because the vertical structure of the waves is oscillatory and most of the vertical displacement occurs within the fluid rather than at the upper boundary. This is in contrast with the case of surface gravity waves (discussed in Chapter 8) where maximum displacement occurs at the surface. The name "gravity waves" derives from the fact that, as for surface waves, the restoring force is due to gravity.

Readers may wonder why—if IW are subsurface phenomena—they deserve a whole chapter in a satellite oceanography book? The reason is that, much to the surprise of some involved in the launch of the first, satellite synthetic aperture radars (SAR), IWs are able to change the sea surface in subtle ways which enable them to create their own signature in SAR images. Not only is this imaging mechanism interesting in its own right but it has unlocked some of the secrets of these waves that would otherwise be hidden, and opened up new oceanographic understanding of their importance. In this chapter we will mainly discuss *oceanic* internal waves, revealed by remote-sensing observations of the sea surface. However, readers should be aware that *atmospheric* internal waves are also observed as manifestations in sea surface roughness patterns, and sometimes it is very difficult, if not impossible,

to distinguish between signatures of oceanic and atmospheric internal waves. But, first, we need to know more about the dynamical character of IWs.

In the interior of a continuously stratified rotating fluid, free IWs are radiated at an angle to the vertical, but are confined to frequencies between f and N, the inertial and the Brunt–Väisälä frequencies, respectively. f is the familiar Coriolis parameter and N, sometimes referred to as the buoyancy frequency, is the natural frequency of oscillation of a fluid parcel displaced vertically from its equilibrium position within a vertical density gradient. Thus:

$$f = 2\Omega \cos \phi, \tag{12.1}$$

where ϕ is latitude; and Ω is the Earth's rotation rate; and

$$N = \sqrt{-\frac{g}{\rho}\frac{\partial \rho}{\partial z}}, \tag{12.2}$$

where ρ is the fluid density; z the vertical co-ordinate pointing upwards; and g is acceleration due to gravity.

These frequencies define the minimum and maximum angles to the vertical with which IWs can propagate. The domain within which the waves can propagate is thus conditioned by their own frequency, σ, the inertial and the Brunt–Väisälä frequencies. For monochromatic waves (with a single frequency) they are evident as "rays" or "beams" of internal wave energy which follow characteristic pathways (see, e.g., Kantha and Clayson, 2000). These rays have a slope c to the horizontal given by:

$$c = \pm \left(\frac{\sigma^2 - f^2}{N^2 - \sigma^2} \right)^{1/2}. \tag{12.3}$$

For this mode of oscillation the frequency of internal waves depends only on the orientation of the wave vector and not on its magnitude, being therefore independent of the wavelength (see, e.g., Pedlosky, 2004). This rather unusual dispersion relation contrasts with interfacial internal waves and surface waves, for which frequency and wavelength have a one-to-one relationship. In addition, energy propagates along the crests and troughs and not perpendicular to them, as in the case of interfacial waves and surface waves. Indeed, the group velocity for three-dimensional internal waves is perpendicular to the wave vector and therefore in the direction of fluid velocity. Analytical demonstrations of these concepts are provided in textbooks such as Gill (1982) and Lighthill (1978).

These rather peculiar relations and geometry are difficult to visualize and somewhat nonintuitive, but good examples are provided by movies of tank experiments (see, e.g., *http://www.phys.ocean.dal.ca/programs/doubdiff/demos/IW1-Low frequency.html*) which were first demonstrated by Mowbray and Rarity (1967). Figure 12.2 shows the result of an experiment in which a small disk is oscillated in a stratified fluid with a constant N (hence the constant slope of the beams) at a constant frequency, σ. In such a case the wave vectors are aligned in a direction such that $\cos\theta = \pm\sigma/N$ (if we neglect the effects of rotation) and there are four such angles (as can be seen from Figure 12.2). It is interesting and important to note

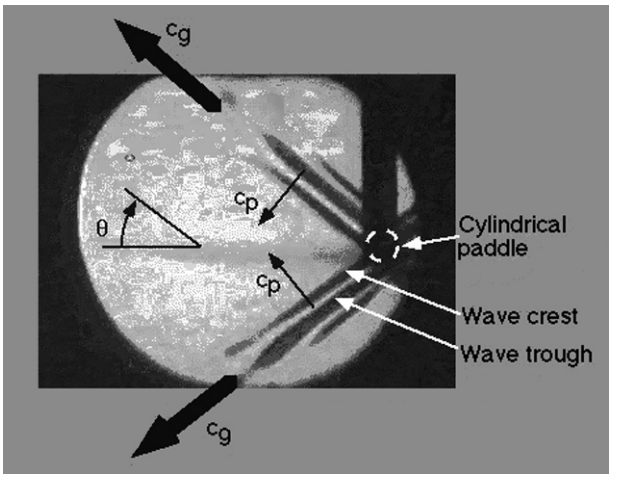


Figure 12.2. A photograph showing the lines of constant phase produced by a small cylindrical paddle oscillating with constant frequency. The angle of constant phase propagation to the horizontal, θ, is determined solely by the fluid's stratification and frequency of oscillation. Phase and group velocity, C_p and C_g, respectively, are indicated. To see a colored movie of this experiment visit *http://www.phys.ocean.dal.ca/programs/doubdiff/demos/IW1-Low frequency.html*

that the disturbance is limited to narrow bands leading away from the oscillating paddle.

In the ocean the "equivalent" of the oscillating paddle is the flow of barotropic tidal currents over bottom topography, forcing the system with semidiurnal or diurnal frequency. New and Pingree (1990) showed that these internal tidal rays (for they correspond to internal waves with tidal period) also correspond to large, internal tidal oscillations of the thermocline, causing amplification of the interfacial internal tide and in some cases generation of nonlinear, short-period, internal interfacial waves. More recently, Gerkema (2001) and Akylas *et al.* (2007) presented model simulations for different stratification regimes showing when nonlinear, short-period internal waves could be "generated by" internal tidal beams, and in Section 12.3 an illustration of such a case will be presented.

12.1.2 The importance of internal waves in physical and biological oceanography

The role of internal waves in vertical mixing of the World Ocean is believed to be an important factor in maintaining ocean structure and circulation (Killworth, 1998; Munk and Wunch, 1998), and also in determining heat transfer between ocean and atmosphere. Understanding and quantifying deep-water mixing processes is essential in explaining how the cold, oceanic waters that sink into the abyss at high latitudes, and flow into low latitudes, rise again into the upper, warm water. Consequently, the study of internal waves is relevant for climatologists. In the upper layers, internal (or interfacial) waves are also responsible for mixing as they propagate into continental shelves where they break or dissipate. This is vital for primary production, because vertical mixing transports nutrients from deeper ocean layers into the upper photic zone where phytoplankton require them for growth. Internal waves can therefore have an important physical impact on marine ecosystems.

The importance of IWs from a biological point of view also stems from their impact on the transport, as well as development, of plankton (Holligan *et al.*, 1985). Nonlinear IWs produce a net transport of in-water particles (phytoplankton, zooplankton, and even small fish), which in the upper surface layer is usually in the same direction as IW propagation. This is likely to affect the exchange of heat, nutrients, and other properties between the shelf and the open ocean (Jeans and Sherwin, 2001). In the lower layer, near the bottom, currents produced by nonlinear internal waves would have the opposite direction to that at the surface, but could be equally important for particle transport. In this case currents may effectively drag sediments (Heathershaw, 1985). Near the surface, typical distances reached by such transport have been modeled by Lamb (1997) and are of the order of several kilometers for a train of internal solitary waves (ISWs—nonlinear, asymmetric, IWs). Some of the early work suggested that IW slicks at the surface are correlated with shoreward transport of pelagic larvae (Shanks, 1983). IWs have the ability to turn scattered distributions of fish and zooplankton into structured distributions, causing the aggregation of organisms in slicks (Pineda, 1999). However, very little research has been done in this field, since traditionally internal waves have been a subject for physical oceanographers, but it is hoped that remote-sensing observations may stimulate further work.

12.2 Internal wave signatures detected with SAR

12.2.1 Introduction

Internal waves are among the most easily recognized of the oceanographic phenomena observed in remote-sensing imagery. The characteristic signatures of alternating bands of light and dark, quasilinear strips have been noted in photographs of the sea surface, in multispectral radiometer images, and in real and synthetic aperture radar images (see Figure 12.3). Once SAR data became widely available, they became the most important remote sensors for IW detection. However, there are different types of radar signatures of short-period, internal wave trains that can be very difficult to interpret. They convey specific information about the characteristics of the internal waveforms that, correctly interpreted, provide unique measurements not only about the IWs but also the interior ocean and the sea surface microlayer. The fact that internal waves (especially internal solitary waves) are among the most coherent and reproducible phenomena in the sea, comparable with the regularity of barotropic astronomical tides, makes them ideal tracers for studying characteristics of the interior ocean such as stratification (thermocline depth) as well as microlayer parameters such as contamination by surface films of organic or hydrocarbon material (see da Silva *et al.*, 1998, 2000; Robinson, 2004). Internal waves follow the tides and seasons, since tidal currents are one ingredient in the recipe for producing most observed IWs, others being stratification and variable bathymetry that perturbs the density structure.

A variety of imaging sensors flown on aircraft and spacecraft have shown that surface manifestations of IWs may be seen in high-resolution images (Apel *et al.*,

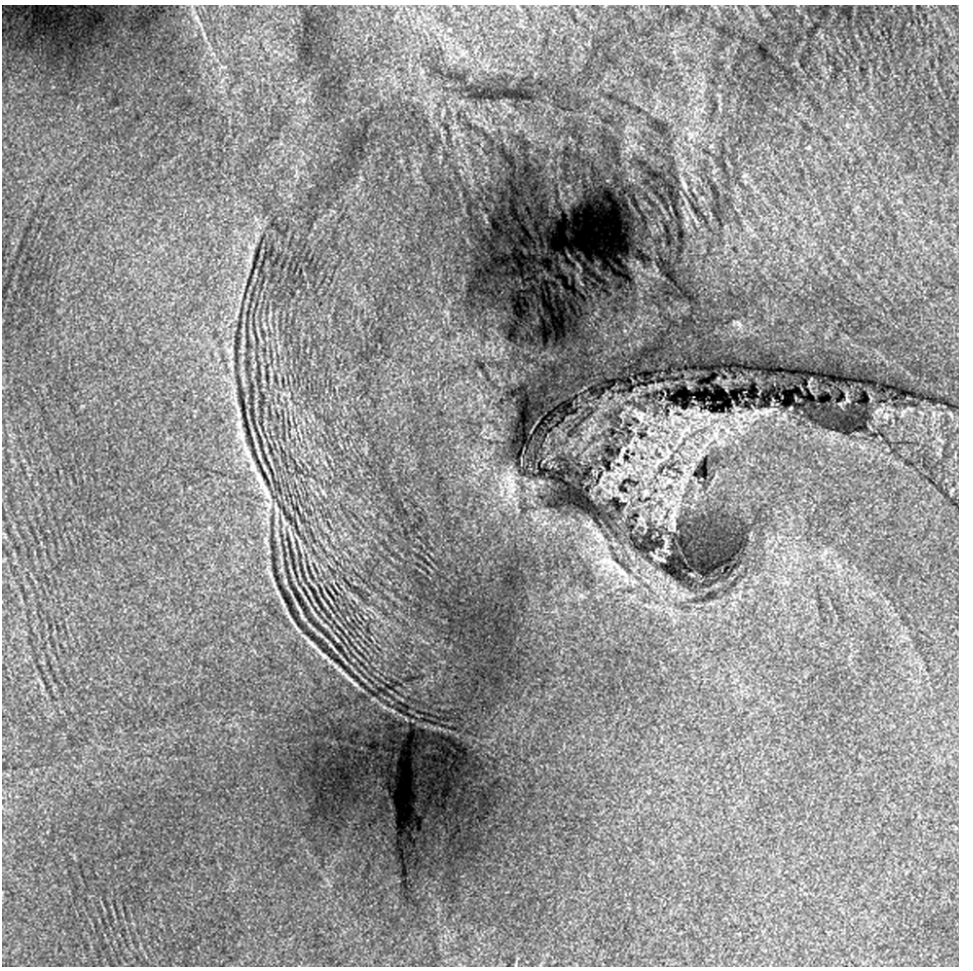


Figure 12.3. ERS-1 SAR image dated August 21, 1994 of the region of Cape Cod (Massachusetts, U.S.A.) showing two trains of internal solitary waves emanating from Race Point Channel. The internal wave crests that are clearly seen in the image have lengths of 10 km to 15 km, and correspond to surface roughness changes such as those shown in the photograph of the same area in Figure 12.1.

1975; Apel and Gonzalez, 1983). Both radar and optical imaging devices have been successful in the observation of these waves, including moderate-resolution optical sensors like the Coastal Zone Color Scanner (CZCS) flown on Nimbus-7 (870 m resolution) which revealed large-scale internal solitary wave signatures in the Andaman Sea in visible wavelength imagery of "sunglint" areas (Apel *et al.*, 1985). Remote sensing has contributed enormously to the study of IWs in the ocean, since such observations can provide details of the two-dimensional spatial structure which cannot easily be obtained *in situ*. These include an overall picture of

the spatial distribution, orientation, propagation direction, and separations of both individual waves and groups or packets of waves.

SAR has several advantages in relation to optical sensors since it is unaffected by cloud cover and, as an active sensor, operates equally well during the day or night. SAR also detects subtle changes in surface roughness more readily than sensors operating in the visible part of the spectrum, which for IW detection are dependent on the Sun's inclination relative to the remote-sensing platform (Melsheimer and Kwoh, 2001; da Silva *et al*., 2003). During the short life of the first SAR in orbit, Seasat imagery revealed signatures that have been interpreted as surface expressions of internal waves (Vesecky and Stewart, 1982) and it became apparent that trains of IWs were far more common than was previously thought. Later, in response to the sustained availability of the European Space Agency ERS satellite series, the number of research projects on internal waves using SAR increased considerably, and numerous papers were published concerning both the explanation of imaging mechanisms (e.g., da Silva *et al*., 1998) and comparison of internal wave model predictions with IW characteristics observed in the images (Brandt *et al*., 1997).

Surface thermal signatures of oceanic internal waves have also been detected *in situ* (by means of a sensor towed at a depth of 15–20 cm), and slick bands were reported to be generally 0.3°C to 0.6°C warmer than ripple bands (Zatsepin *et al*., 1984). Marmorino *et al*. (2004) collected infrared imagery of small-scale internal waves using an airborne infrared camera. Infrared imagery appears to be able to detect internal waves under conditions when the wind is too low to generate surface waves, and hence there are no Bragg scatterers to reveal IWs unambiguously in radar imagery. In such a case, infrared imagery might serve as an alternative or at least an adjunct to radar measurements. Altimeters have also been capable of internal wave detection, but in this case only for the large-amplitude, internal solitary waves characteristic of the Sulu Sea (Kantha and Clayson, 2000). In that region the phase speed of IWs are among the highest in the world and IW/surface wave resonant interactions can lead to amplification of meter-scale surface waves, which may even break and cause strong roughness. This has been known since the 19th century, and there were several reports of sailors who observed "boiling seas" and "tide rips" in otherwise calm conditions. A set of photographs showing the phenomenon has been published (Osborne and Burch, 1980). In such cases, altimeter data may also prove useful.

In the following sections we will introduce some techniques to interpret internal wave signatures, focusing on SAR examples of internal, solitary wave packets, since SAR is the most useful and complete sensor to observe these particular phenomena.

12.2.2 Internal, solitary wave packets observed by SAR

In ocean midlatitudes, in summer, the top 20 m or 30 m can be several degrees warmer than water below, and this gives rise to a thermocline along which interfacial internal waves can propagate. Although the oscillations may typically have amplitudes of 10 m or more, internal waves produce very small surface elevations, on the order of 10 cm or less. Still, the cellular currents that accompany them are of the

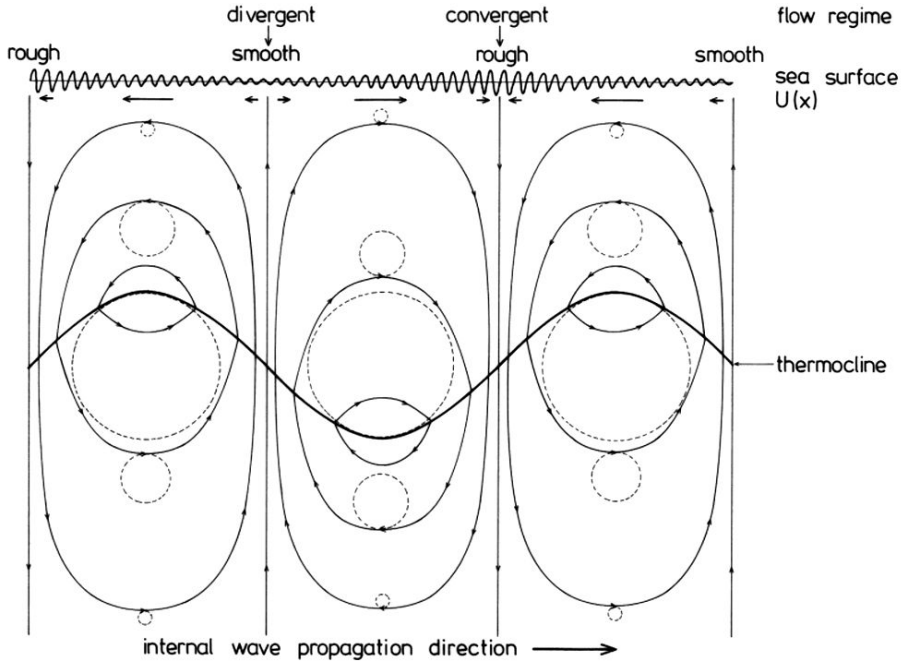


Figure 12.4. Schematic plot of processes associated with the passage of a linear oceanic internal wave. Deformation of the thermocline (heavy solid line), orbital motions of water particles (dashed lines), streamlines of the velocity field (light solid lines), surface current velocity vectors (arrows in the upper part of the image), and variation of the amplitude of Bragg waves (wavy line at the top) (after Alpers, 1985).

same order as wave phase speeds, typically a few tens of centimeters per second to 2.5 m/s. The periodic spatial patterns of surface currents produce convergences and divergences strong enough to modulate short-length, surface gravity waves and capillary waves, resulting in a surface roughness signature characteristic of the underlying internal wave field (as shown in Figure 12.4). The sea surface roughness patterns produced by the internal wave–surface wave interaction are responsible for making them visible to satellite sensors such as SARs and moderate-resolution optical sensors such as MERIS. In the case of SAR, the amplification of Bragg waves in the convergence zones at the surface and attenuation of Bragg wave amplitudes in the divergence zones above the rear slopes of internal waves, are responsible for the image intensity modulations observed as bright and dark bands (as shown in Figure 12.5).

Simplified models assume that the ocean consists of several layers of uniform density with a sharp change in density between each layer. In the simplest case (two-layer model) only two layers are considered corresponding to conditions typical of many parts of the ocean: a top layer warmer and less dense above a sharp thermocline and a lower layer cooler and denser.

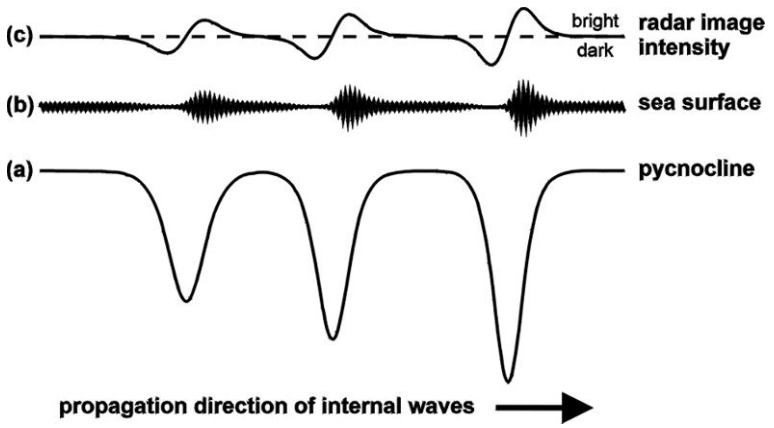


Figure 12.5. Schematic showing an internal solitary wave packet consisting of solitons of depression with decreasing amplitude. (a) Shape of the pycnocline. (b) Sea surface roughness pattern caused by the soliton. (c) SAR image intensity associated with (b) (after Alpers, *http://www.ifm.uni-hamburg.de/ers-sar/*).

Solitons, solitary waves that retain their shape and speed even after collisions with each other, have been found in many branches of physics. The name internal "solitary" waves is used because, in the ocean, these waves are found either as single crests or in isolated packets, and have often been identified with the internal soliton solutions of nonlinear wave theory. They are finite amplitude waves that result from an exact balance between the nonlinear steepening of the waveform and the tendency towards dispersion of the wave in the governing equations, so that their shape and speed remain invariant as they propagate in the ocean. Solitary waves at the interface of a two-layer fluid (upper-layer thickness H_1, lower-layer H_2) are governed by the Korteweg–de Vries (K-dV) equation (Korteweg and de Vries, 1895; Drazin, 1983):

$$\frac{\partial \eta}{\partial t} + c\frac{\partial \eta}{\partial x} + \alpha\eta\frac{\partial \eta}{\partial x} + \gamma\frac{\partial^3 \eta}{\partial x^3} = 0, \tag{12.4}$$

where η is interfacial displacement; and

$$c = \left[g\frac{\Delta\rho}{\rho}H_1(1+r)\right]^{1/2}; \tag{12.5}$$

$$\alpha = -\frac{3c}{2}[(1-r)/H_1]; \tag{12.6}$$

and

$$\gamma = cH_1H_2/6; \qquad \Delta\rho = \rho_2 - \rho_1; \qquad r = H_1/H_2. \tag{12.7}$$

If the upper layer (density ρ_1) is thinner than the lower one (density ρ_2), as is usually the case for midlatitude regions, the internal soliton has a downward displacement of the form

$$\eta(x,t) = -A\,\mathrm{sec}\,h^2[2(x-Ct)/\lambda], \tag{12.8}$$

where A is the amplitude of the waveform, and the wave is called a soliton of depression. Its phase speed, C, can be written as a function of the equivalent linear phase speed, c, as

$$C = c(1 - A\alpha/3c) \tag{12.9}$$

and its "wavelength" is

$$\lambda = 4(-3\gamma/\alpha A)^{1/2}. \tag{12.10}$$

Note that from Equation (12.9) it can be seen that for waves of higher amplitude, A, we will have higher phase speeds, C, since α is negative because $r < 1$ for depression waves. This explains the commonly observed rank ordering of nonlinear internal waves (shown in Figure 12.5), where amplitudes and wavelengths decrease towards the rear of a packet. This is a direct consequence of the fact that bigger waves propagate faster than smaller waves. Given time to develop, a train acquires a hierarchic form since the larger waves will overtake the smaller.

Tidally generated, short-period internal waves are characterized by their time evolution as they propagate across a shelf. Waves observed closer to their generation region (e.g., the shelf break) have been observed not to have such well-developed waveforms and to be more irregular in rank order. Those waves which have had sufficient time to evolve into well-developed, soliton-like waveforms have been observed as organized, rank-ordered wave packets. Ostrovsky and Stepanyants (1989) reviewed internal solitary waves in the ocean, and have considered the extent to which they can be regarded as internal solitons (see also a recent review by Helfrich and Melville, 2006). They concluded that the rank ordering of wave amplitudes and wavelengths are the most commonly observed evidence for the nonlinear nature of internal waves. This fact has been observed not only by detailed *in situ* measurements but also in satellite images (such as Figure 12.3), and in particular by SARs (Apel and Gonzalez, 1983). For a reader with time to spare, it is well worth exploring the richly detailed and diverse imagery that has been assembled in an atlas of internal soliton images and other data (Jackson, 2004).[2]

12.2.3 Identification of internal wave trains and their propagation direction

In general, most internal waves observable in SAR imagery exhibit characteristics that indicate important properties, such as propagation direction and speed, length-scales, polarity, and in some cases it is possible to estimate their amplitudes (when auxiliary data from the interior ocean is available). One of the earliest pioneers of the remote sensing of internal waves, John Apel (1930–2001) described observed, internal wave characteristics as follows:

> "They propagate in separate groups or 'packets,' with each packet being generated by the semidiurnal tidal cycle. The separation between packets can range from about 10 km up to 90 km. Each packet contains a few to a few dozen individual waves and the individual wavelengths can range from 100 m up to

[2] Accessible online at *http://www.internalwaveatlas.com/Atlas2_index.html*

20 km. Their along crest length scales vary from 10 km to more than 100 km. The largest waves (in amplitude, wavelength, and along crest length) are found at the leading edge of each packet, and the waves decrease in all aspects to the trailing edge. Usually, the wave signatures observed in SAR imagery are a series of alternating light/dark linear or curvilinear bands that represent the crests and troughs of the waves'' (Apel, 2004).

A very important capability of SARs is the information they provide about internal wave propagation direction. We have seen that due to the hydrodynamic interaction between surface waves and the variable surface currents induced by internal waves, the amplitude of Bragg waves that backscatter radar energy is increased in convergent flow regions and is decreased in divergent flow regions. As a consequence, the radar signatures of oceanic internal waves consist of alternating bright and dark bands on a uniform background (as shown in Figure 12.3). The polarity of the signature is defined by comparing image intensity modulation of the IW profile relative to the ''background'' intensity of an isolated region containing no internal wave activity, but for which wind speed and direction, and the SAR incidence angle, can be assumed to be the same as for the IW region. In the case of the curved packets of waves in Figure 12.3 the polarity of the IW signature is bright/dark, which is also referred to as positive/negative $(+/-)$ by da Silva *et al.* (1998).

The direction of IWs may be interpreted from their SAR signatures according to three simple rules. First, the brighter band (positive intensity modulation) indicates the propagation direction (i.e., waves propagate perpendicular to the band towards the positive side of the band—assuming that the waves are of depression, which is generally the case in deep waters where $H_1 < H_2$). Second, for a nonlinear, internal wave packet the rank ordering of the waves within the packet indicates the propagation direction: amplitudes and wavelengths decrease towards the rear of a packet. Third, if internal wave signatures consist of curvilinear bands that represent the crests and troughs of the waves, the curvature (e.g., concave or convex) may indicate the direction of propagation. For example, internal waves may be generated at submarine mountains, sills, or straits, and a bathymetry map of the study region might be useful to determine their direction of propagation. One should also compare crest waveforms with an auxiliary bathymetry map and look for shapes of isobaths similar to the wave crests in the vicinity of the satellite observation.

12.2.4 Hydrodynamic and film modulation

Hydrodynamic modulation theory describes the evolution of small-amplitude surface waves in a slowly varying current, and it is derived from the action balance equation (see, e.g., Apel, 1987), which accounts for the conservation of wave energy where the wavenumber and frequency of the wave field vary in space and time (Bretherton and Garrett, 1968). Alpers (1985) assumed that the variable surface current due to the orbital velocity of IWs leads to only small deviations from the equilibrium wave spectrum and solved the action balance equation retaining only

first-order terms. Defining the projection of the radar antenna axis on the horizontal plane as the x-axis, and admitting Bragg scattering as the mechanism for radar cross-section modulation in the IW field, the following equation can be obtained:

$$\frac{\delta\sigma}{\sigma_0} = -(4 + c_g/c_p)\tau\frac{\partial U_x}{\partial x}, \tag{12.11}$$

where $\delta\sigma = \sigma - \sigma_0$ denotes deviation of the normalized radar cross-section, σ, from its mean value σ_0 in a nearby region unaffected by IWs; c_p and c_g are the phase and group speeds of Bragg waves; and $\tau(\boldsymbol{k})$ is the relaxation time (the time duration over which surface waves of wavevector $\boldsymbol{k}$ remain strained until they reach equilibrium with the wave spectrum under the influence of wind forcing and dissipation processes). Thus, if Bragg scattering theory is assumed, the cross-section modulation produced by IWs is quite simply proportional to the product of the surface current gradient in the radar look direction and τ. It is important to note here that this simplification was obtained disregarding the damping effects of surface films.

A consequence of Alpers' results is that IW signatures predicted in this approximation of the hydrodynamic theory are characterized by the positive and negative variations of backscatter from the mean background σ_0 (double-sign signatures), which are exemplified in the C-band SAR image of Figure 12.3 and the X-band image of Figure 12.6a. It also follows that the greater τ is, the stronger the IW signature will be. Since the relaxation time τ is expected to decrease with increased wind speed, SAR signatures of IW are not expected under strong wind conditions. But the series of alternating bright/dark bands is not the only type of signature that short-period internal waves can exhibit. Sometimes the signature consists only of dark bands on a gray background, when wind speed is low or moderate. Obviously, the hydrodynamic theory (also sometimes referred to as kinematic theory) cannot by itself explain the existing variety of imaging radar observations of IW signatures.

Ermakov *et al.* (1992) studied some processes of film slick formation on the sea surface and made *in situ* measurements of internal waves in the presence of surface films. Anyone watching the ocean when winds are low will soon notice that some areas of the ocean surface appear smoother than adjacent areas. These smoothed areas, called surface slicks, are often visible as long bands or patches, sometimes of rather complex form, with dimensions from 10 m to a few kilometers. An observer close to the surface (in a boat, for example) would notice that small waves that make the surface look rough are present outside the slicked areas but are missing or altered within. Figure 12.1 shows a photograph of such a situation. Slicks are believed to be primarily composed of naturally occurring, surface-active organic materials which concentrate in the form of films on the ocean surface. Horizontal convergences due to current field variations at the ocean surface, such as those due to internal waves, can compress surfactant materials and form a surface-elastic film that becomes concentrated enough to attenuate surface waves. Surface films are easily observed by SARs as slicks (dark areas in the image) because they are effective at damping the Bragg waves responsible for radar backscatter. da Silva *et al.* (1998) showed that when films are present they can modulate short-scale surface roughness, so that the radar signature of an internal wave field consists of dark lines or bands only (areas of

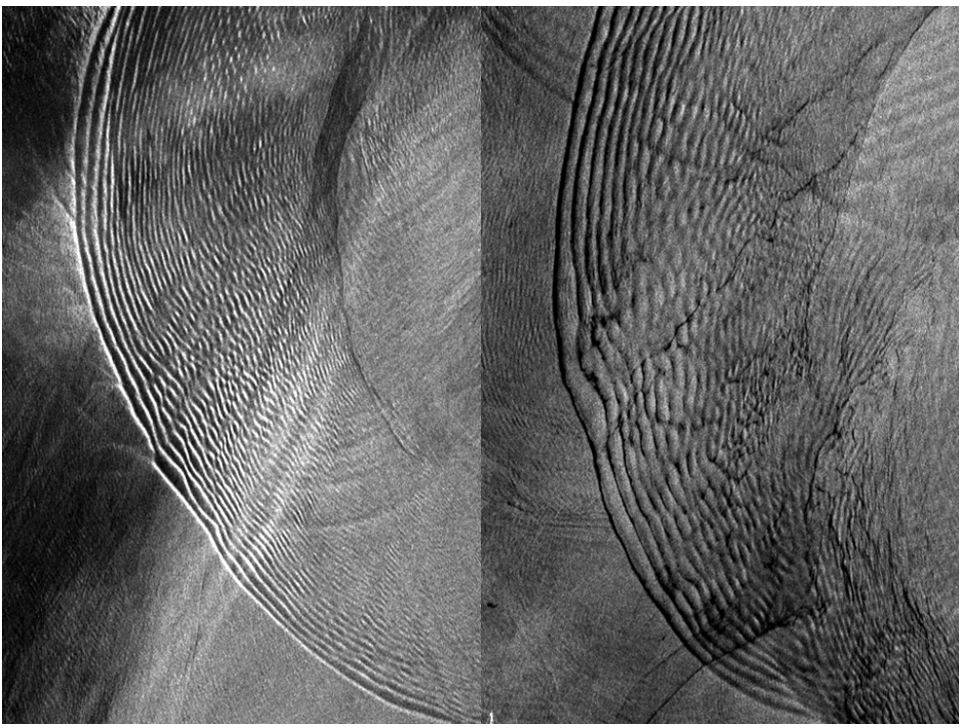


Figure 12.6. (a) On the left a TerraSAR-X image (X-band) dated June 23, 2008 showing a typical example of double-sign signatures (note the bright and dark bands compared with the local gray level). (b) On the right a TerraSAR-X image dated July 4, 2008 of the same region (Cape Cod Bay, U.S.A.) showing internal wave signatures as dark bands on a gray background (single-negative signature). The latter signature is typical in coastal zones in the presence of surface films.

reduced radar backscatter) on a uniform gray background (as shown in Figure 12.6b).

Figure 12.7 shows another IW signature which could be rather puzzling because it has both the classical double-sign signature and also a slick-like negative signature for different crests in the same train. The implication is that both types of imaging mechanism are present here, creating a challenge to the theoretical model. Whereas theories had been advanced to interpret roughness patterns in terms of either straining of short surface waves or their variation due to films, the combined effect of both mechanisms had not been studied until images like this were encountered. In response, a quantitative analysis method was developed (da Silva *et al.*, 2000) to account for the extent to which surface films may transform one type of signature into another. It explains how an increase in background film concentration may trigger transition from double-sign signature modes (bright/dark bands) into single negative signatures (dark bands), such as the example in Figure 12.7 off the coast of

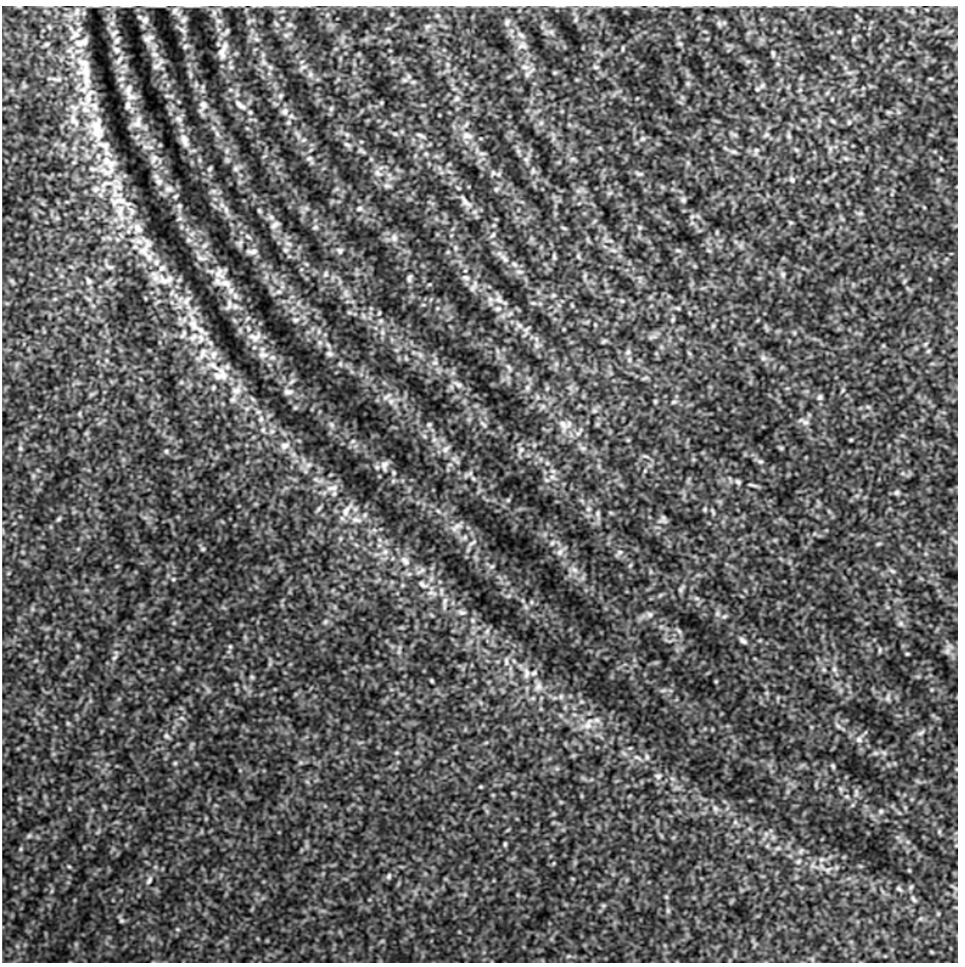


Figure 12.7. Example of SAR image showing signature transition from double- to single-negative sign. This is an ERS-2 SAR image (C-band) acquired over Massachusetts Bay on August 17, 1996 (15:28 UTC), and the sea area covered is 6.4×6.4 km.

Massachusetts. Within a single IW packet, ambient film concentration is likely to increase towards the rear of the packet (see Ermakov *et al.*, 1998) due to the convergence effects of the larger scale, linear interfacial tide.

Figure 12.8 presents some results of the theoretical model (da Silva *et al.*, 2000) in which the center panel shows predicted backscatter when the surfactant is assumed to increase towards the back of the wave train. It shows how the signature for the second and third crests is different in character from the first, exhibiting the signature mode transition. It also shows how such transition is likely to be more prominent for C-band radars such as ERS SAR and Envisat ASAR than for L-band radars such as that on Seasat.

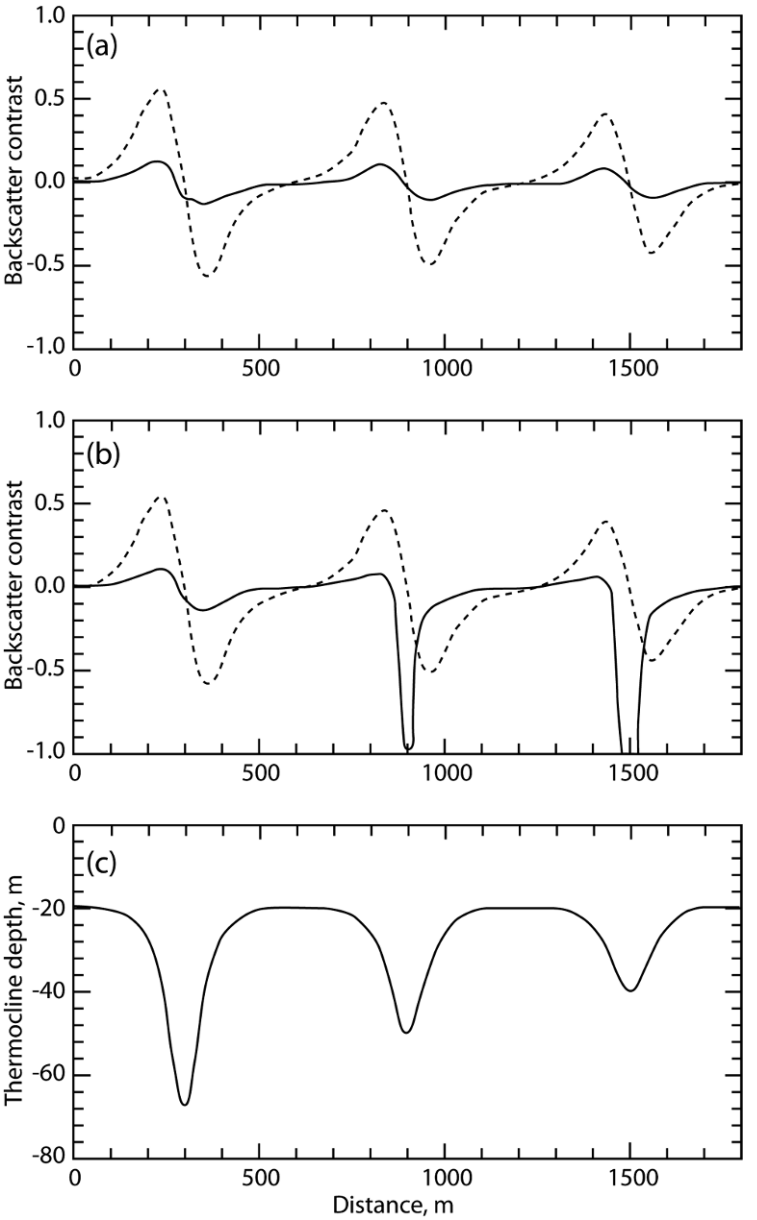


Figure 12.8. Predicted backscatter contrasts across an IW packet with decreasing amplitudes. The contrasts shown are for C-band Bragg wavelengths ($\lambda = 7$ cm: continuous lines) and L-band ($\lambda = 30$ cm: dashed lines): (a) when the unperturbed film concentration is constant across the packet (film pressure $\pi_0 = 0.1\,\mathrm{mN\,m^{-1}}$); (b) when the unperturbed film concentration varies because of the internal tide convergence ($\pi_0 = 0.1\,\mathrm{mN\,m^{-1}}$, first IW; $0.2\,\mathrm{mN\,m^{-1}}$, second IW; $0.3\,\mathrm{mN\,m^{-1}}$, third IW). (c) The assumed IW packet profile used to drive the image contrast model, with characteristic decrease of IW amplitudes toward the rear of the packet.

12.2.5 Internal wave mean propagation speed

Envisat ASAR spatial coverage in wide-swath mode (approximately 400×400 km) allows the simultaneous view of most continental shelf zones, adjacent shelf breaks, and a considerable fraction of open-ocean basins, making them a powerful data source for studying internal waves. Numerous studies of internal waves based on satellite SAR images report the existence of groups of packets of solitary internal waves systematically propagating across the shelf (sometimes with simultaneous offshore- and inshore-propagating packets) with typical interpacket separations of the same order of magnitude as internal tidal waves should have (e.g., Vesecky and Stewart, 1982; Apel and Gonzalez, 1983; da Silva *et al.*, 2007). It is therefore natural to assume that such internal solitary waves are linked to associated, large-period, internal tidal waves. It is believed that solitary waves are generated by the nonlinear steepening of the internal tide, which has been observed by Pingree *et al.* (1983) in the Celtic Sea, having steep and narrow troughs compared with the crests. They explained this through the advective effects of the barotropic tide, which prevented the leading edge of the newly formed on-shelf trough from propagating onto the shelf during strong off-shelf tidal streaming. This resulted in a distorted and steepened trough that was subsequently released to propagate inshore as the tide relaxed.

If packet generation is assumed to be phase-locked with semidiurnal tides, due to tidal current interaction with bottom topography across the slope, it is possible to estimate the average propagation speed of packets from a SAR image by measuring the interpacket distance (see the ASAR wide-swath image in Figure 12.9). average phase speed is simply $c = \Delta x/T$, where Δx is the distance between the first soliton of two consecutive packets (measured in the direction of wave propagation); and T is the tidal period (semidiurnal or diurnal, depending on the study region). Note that the phase speed c will be an averaged value since, for instance, shoaling over the shelf occurs and may alter the internal wave packet phase speed as it progresses towards the coast. Note also that, under this assumption, internal wave propagation speed is found to have a seasonal variability, as might be expected, due to changes in mean stratification throughout the year. Short-term variations in propagation speed are also likely to occur over a few days due to changes in local wind conditions altering the density structure by upwelling.

12.2.6 Inversion of polarity in SAR signatures of internal waves

In Equation (12.6) the sign of α depends on the ratio $r = H_1/H_2$. For $H_1 < H_2$ the upper layer is shallower than the lower one, the internal soliton is a wave of depression, and only an initial waveform that is a depression can generate internal wave solitons (Kantha and Clayson, 2000). A single soliton propagating from deep water onto a shallow shelf can "fission" into a train of rank-ordered solitons (Liu *et al.*, 1998). If $H_1 > H_2$ the upper mixed layer is thicker than the bottom layer, and solitary waves will be waves of elevation, with displacement of the interface upward. When a soliton train in deep water consisting of depression waves propagates into

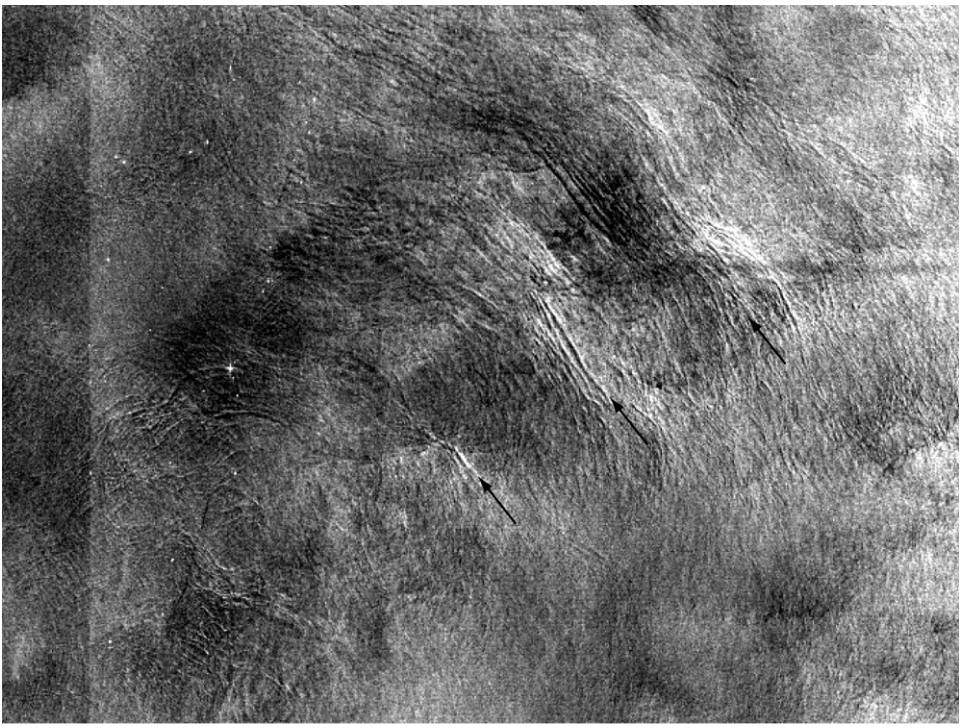


Figure 12.9. Wide-swath ENVISAT ASAR images showing several successive trains generated by tidal flow at the Spanish and French continental shelves. The arrows point to three distinct packet fronts which may be considered to have been generated at the same place but on successive tidal cycles. This image represents an area of sea about 400 km wide.

shallower water, the solitons first disintegrate into dispersive wave trains and then reorganize themselves as a packet of nonlinear elevation waves in shallow water after they pass through a switching point where the upper and lower depths are approximately equal.

We observe this fascinating behavior in an ERS SAR image of the Gulf of Cadiz (see Figure 12.10). This transformation of polarity was first observed in a SAR image by Liu *et al.* (1998) as well as simulated by a numerical model consisting of a K-dV equation of the type discussed in Section 12.2.2. There is another example of transformation of polarity shown in figure 10.47 of *MTOFS*, also in the Andaman Sea, while Figure 12.10 confirms that such phenomena can also occur for moderate latitudes. Figure 12.11 shows a schematic of the depression, solitary wave train evolving into elevation waves. A train of several, rank-ordered elevation solitons can emerge from the disintegration of a single-depression soliton as it passes through the switching depth. There are also observations of similar behavior amongst nonlinear internal waves off Taiwan in the East China Sea and off Hainan in the South China Sea in ERS-1 SAR imagery described by Liu *et al.* (1998).

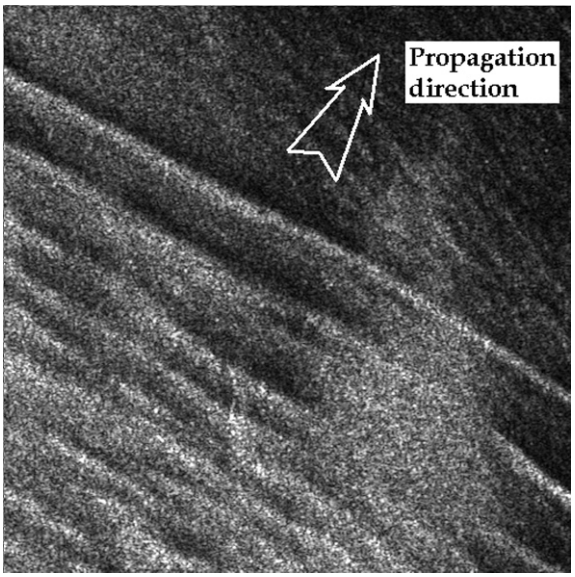


Figure 12.10. ERS-2 SAR image dated July 23, 1998 (11:10 UTC) acquired over the Gulf of Cadiz (Spain). The white arrow indicates the propagation direction of the internal wave train (north is upwards). Note that the leading wave of the wave packet is characterized by negative backscatter variation (dark band) preceding positive variation (bright band) in the propagation direction. Such a pattern is consistent with a SAR signature of a solitary wave of elevation (see also Figure 12.11 and text for details).

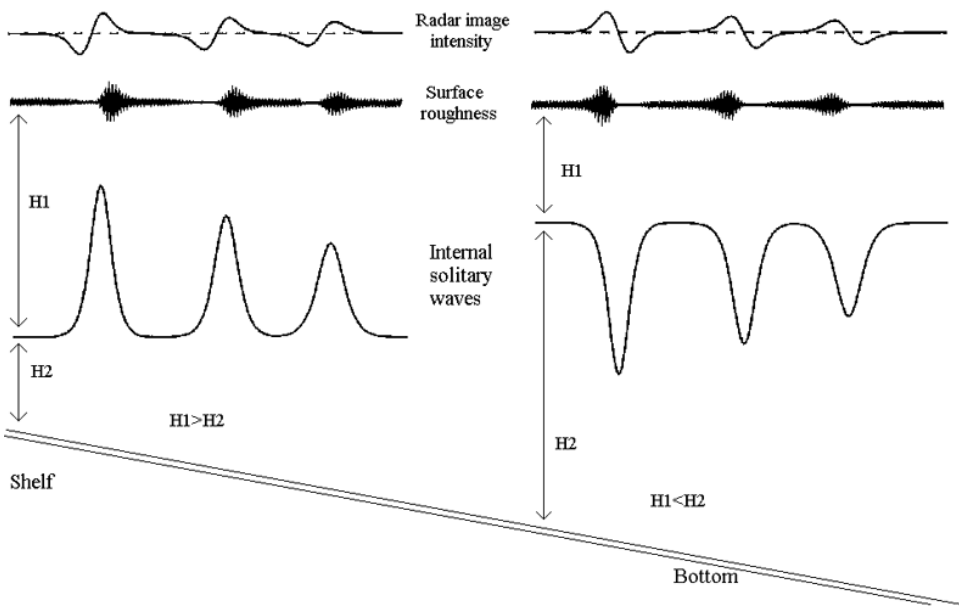


Figure 12.11. Schematic diagram of internal waves, surface waves, and SAR image intensity variation when depression solitary waves move (from right to left) into shallower water. The interface displacements observed can have amplitudes of several tens of meters (adapted from Liu *et al.*, 1998).

12.3 INTERNAL WAVES AND OCEAN COLOR

12.3.1 Observations

In this section we present an example of interfacial, internal tidal waves propagating off the Armorican shelf break in the Bay of Biscay observed in a pair of SeaWiFS and ERS SAR images obtained quasisimultaneously. Large, internal tidal waves have already been extensively studied in the Bay of Biscay, between France and Spain (see Figure 12.12a). These internal waves of semidiurnal tidal period result from the interaction of the surface tide with the steep shelf slope topography, and propagate both onto the shelf and into the deeper ocean. In the upper water column, these internal tides (ITs) are characterized as long-wavelength (30–50 km) depressions and elevations of the thermocline of up to 30 m in amplitude. Pingree *et al.* (1986) observed these waves *in situ* to travel for over 250 km into the deep ocean from the shelf break near 47°30′N, 6–8°W, with typical propagation speeds of about $1\,m\,s^{-1}$ in the summer. These internal tides were also visible in remotely sensed sunglint AVHRR imagery as long-crested features extending for several hundreds of kilometers in a direction parallel with the shelf break (Pingree and New, 1995), and were considered to be propagating directly away from the shelf break.

Figure 12.12b shows an example of a time series of wave motions recorded in the upper-ocean temperature structure, at position B (46°19′N, 7°14′W) over one tidal cycle on July 5, 1988. The mean depth of the thermocline (14°C contour, say) is about 50 m, but there are two pronounced depressions, centered near 10:00 h, and one tidal cycle later, near 22:30 h. These are internal tidal troughs: the thermocline is generally depressed to about 110 m deep in the former and 90 m in the latter. In between (around 15:00–16:00 h), the thermocline rises to about 30 m deep in the internal tidal crest. Superimposed on this long-wavelength tidal motion are internal solitary waves (ISWs) (i.e., the much shorter waves we have discussed in the previous section and that are so well captured in SAR images). These are particularly pronounced in IT troughs, both of which contain at least two (and possibly more) large-amplitude ISWs. ISWs typically have wavelengths between 1 km and 2 km, periods of 20–40 minutes, and result from the action of nonlinear and dispersive forces on the internal tides themselves (New and Pingree, 2000). It is important to note that, at least in this region, ISWs can therefore be considered as marking the positions of internal tidal troughs.

In situ observations by Lennert-Cody and Franks (1999) suggested that short-period internal waves may have an effect on the patchiness distribution of plankton in near-surface layers, and the question then arises as to whether these large ITs or ISWs would have any effect on the distributions of phytoplankton and hence on ocean color satellite imagery. At the shelf break, ITs and ISWs are thought to be responsible for physical mixing of the water column which increases the levels of nutrients in near-surface layers, giving rise to elevated levels of chlorophyll in a generalized band over the shelf break (e.g., Pingree *et al.*, 1986). However, such mixing must be sustained for several days at least in order for the phytoplankton population to respond significantly, so that bands of enhanced chlorophyll asso-

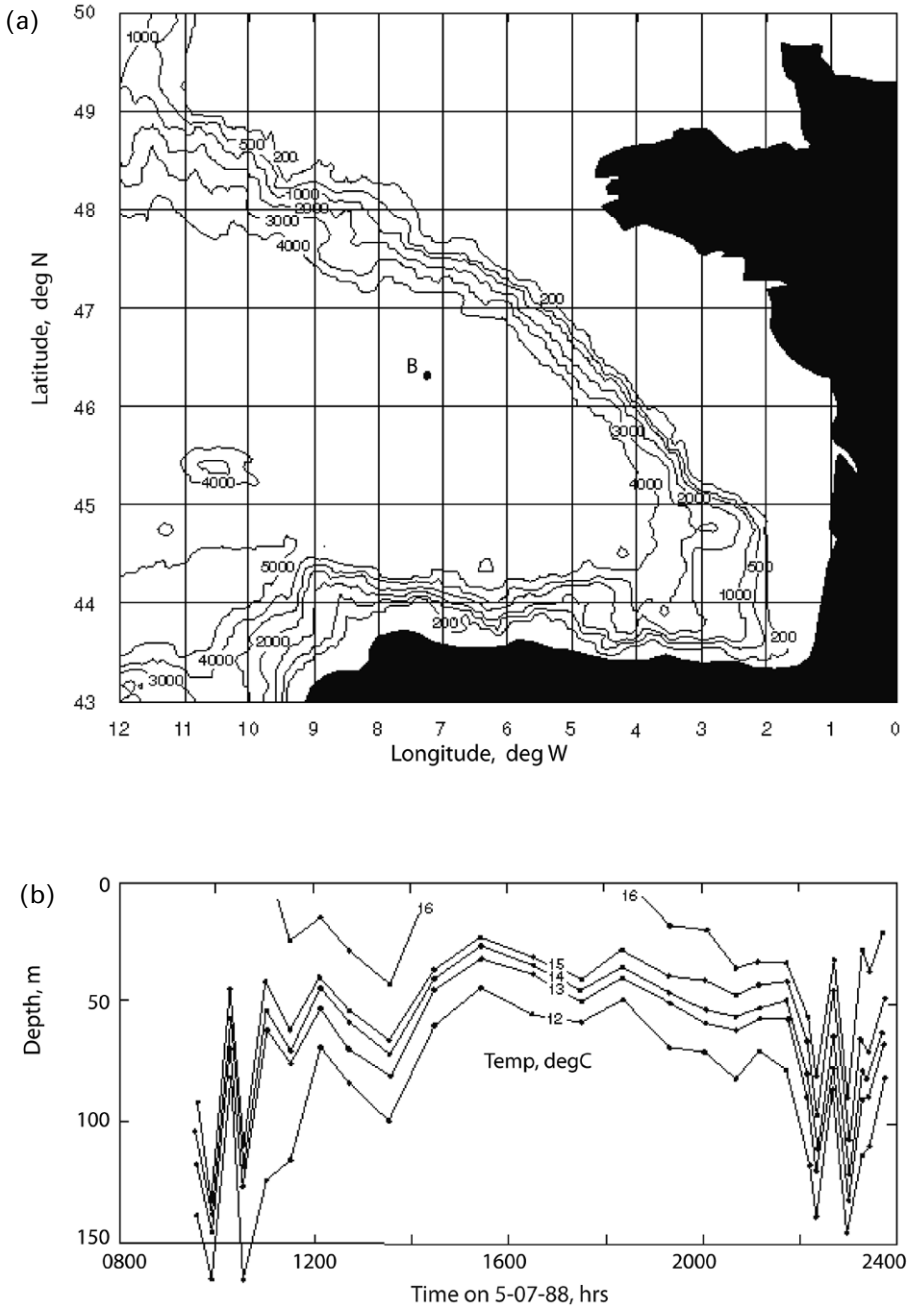


Figure 12.12. (a) Chart of the Bay of Biscay, showing depth contours (m) and the coasts of northern Spain and western France. Position B is 46°19′N, 7°14′W. (b) Time series of the observed thermal structure (°C) from an expendable bathythermograph (XBT) survey at *B* on July 5, 1988 (from da Silva *et al.*, 2002).

ciated with individual, traveling internal tides (with timescales of only a few hours) are unlikely to be caused by such a mechanism.

da Silva *et al.* (2002) investigated bands of enhanced levels of near-surface chlorophyll in the central Bay of Biscay in remotely sensed images from the SeaWiFS ocean color sensor. They showed that these are associated with the crests of internal tidal waves traveling away from the shelf break, which can be "seen" by the satellite sensor because the internal tide is able to lift a subsurface chlorophyll maximum (located near the thermocline) sufficiently close to the ocean surface. An example of this can be seen in Figure 12.13.

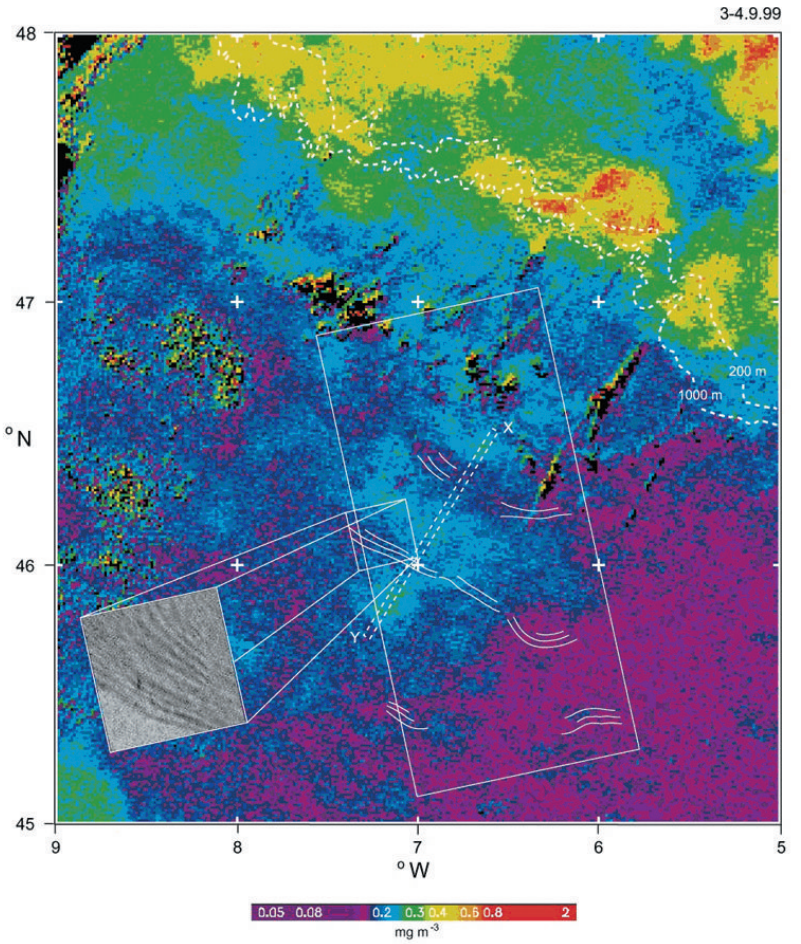


Figure 12.13. Chlorophyll concentration (color) from SeaWiFS on September 4, 1999 and coincident internal waves (white lines) from the ERS-2 SAR on September 3, 1999 (see text for details). The area covered by the SAR is shown by the large white rectangle, and X–Y denotes the transect used for Figure 12.16. Only every second IW (up to a maximum of three per packet) is shown for clarity. Shelf break depth contours at 200 m and 1,000 m are indicated, and an expanded portion of the SAR image itself is overlaid in the lower left corner.

Figure 12.13 shows a combination of SeaWiFS and SAR observations on September 3 and 4, 1999 (5–6 days after spring tides at Brest, France). The SeaWiFS image was acquired at 13:00 h (UTC) on September 4, and the ERS-2 SAR image at 22:36 h on September 3. The SAR image is thus about one tidal cycle (12 h 25 min) plus 1 h 59 min earlier than the SeaWiFS image. We have assumed that the ISW patterns are the same on these two, successive tidal cycles, and that the ISWs move away from the shelf break at $1.03\,\mathrm{m\,s^{-1}}$ (the same speed as the ITs for a typical summer stratification—Pingree *et al.*, 1986). Then the ISWs at 13:00 h on September 4 should be in the same locations as in the SAR image at 22:36 h on September 3, but displaced an additional 7.4 km (the distance they would travel in 1 h 59 min) from the shelf break in their apparent direction of propagation (to the south-southwest). This correction has been made in Figure 12.13, so that ISWs (marked as white lines) can be viewed as though they were coincident with the SeaWiFS image.

We clearly see in Figure 12.13 the band of strongly enhanced chlorophyll levels near the shelf break previously ascribed to the mixing caused by internal waves and tides. Oceanwards from the break, and within the SAR frame shown (white rectangle), we can see that, generally, low chlorophyll strikingly corresponds with ISW packets. On the other hand, two bands of enhanced chlorophyll concentration are sandwiched between packets of ISWs, and so must correspond with internal tidal crests. A third band of enhanced levels of chlorophyll concentration can be seen closest to the shelf break, at about 46°30′N (and 6°20′ to 6°50′W, just south-southwest of position *X*), and may also correspond to another internal tide crest. Elsewhere, chlorophyll levels are considered to be at their background levels.

12.3.2 Remote sensing and depth distribution of ocean chlorophyll

The constraints on ocean color remote sensing to detect chlorophyll at depths below the surface has already been mentioned in Section 7.3 in the context of estimating primary production rates. It is worth looking again at some specific details in order to clarify how ocean color data can be interpreted in relation to the impact of internal waves on phytoplankton populations. Light intensity decreases nonlinearly with depth in the water column. In fact, optical attenuation is exponential, and a parameter can be defined, Z_{90}, to give the depth of penetration of light above which 90% of diffusely reflected irradiance (excluding specular reflectance) originates. This depth, Z_{90}, can also be considered as the depth to which the satellite sensor effectively "sees". Gordon and McCluney (1975) showed that for an homogeneous ocean

$$Z_{90} \approx K^{-1}, \tag{12.12}$$

where K is the diffuse attenuation coefficient for downwelling irradiance. For remote-sensing purposes, the concentration of the water constituent under consideration (e.g., chlorophyll) should be weighted by a factor $g(z)$ when estimating the remotely

sensed concentration. This factor is

$$g(z) = \exp\left\{-2\int_0^z K(z)\,.\,dz\right\}. \tag{12.13}$$

Frequently, as a first-order approximation, $K(z)$ is considered approximately constant with depth so that

$$g(z) = \exp\{-2Kz\}. \tag{12.14}$$

The weighting factor $g(z)$ can be regarded as being derived from irradiance arriving at the surface having been attenuated by $\exp[-K\,.\,z]$ from the surface to the depth z and by the same factor on the return to the surface. Gordon and Clark (1980) proposed an equation to calculate the remotely sensed concentration of chlorophyll, which is:

$$c_{\text{sat}} = \frac{\int_0^{z_{90}} c(z)\,.\,g(z)\,.\,dz}{\int_0^{z_{90}} g(z)\,.\,dz}, \tag{12.15}$$

where $c(z)$ is the concentration of chlorophyll as a function of depth. If, as is generally the case, $c(z)$ is not too complex then c_{sat} can be used as an index of mean chlorophyll concentration in the water column.

Figure 12.14 shows an example of typical chlorophyll profiles plotted as a function of water depth, for a variety of water types, ranging from oligotrophic water to productive coastal water (Cullen and Eppley, 1981). Note that all three chlorophyll profiles present a subsurface maximum at 120 m, 50 m, and 20 m, respectively for geographic sites *A*, *B*, and *C*. This typical deep chlorophyll maximum (DCM) often occurs in the summer when levels of surface nutrients, phytoplankton, and chlorophyll have become depleted following the spring bloom, leaving behind a subsurface maximum near the thermocline.

12.3.3 A model for interpreting ocean color signatures of internal tides

If we consider the existence of a subsurface DCM and the vertical displacements by internal waves that particles on such a layer (specially chlorophyll) would experience, it is reasonable to assume that a plausible mechanism to explain the bands of enhanced chlorophyll observed by ocean color sensors such as SeaWiFS, MODIS, and MERIS in the Bay of Biscay could be the uplifting of a DCM by the passage of internal tidal crests. Chlorophyll in internal tidal crests would rise to such a water depth level as may be seen by the satellite sensor. In order to assess this hypothesis, da Silva *et al.* (2002) quantified the effect with a simple model, as follows.

Let us assume that a DCM exists between depths h_1 and h_2 (with $h_2 > h_1$) in which the chlorophyll concentration is uniform and equal to $(c_b + c_0)$, while elsewhere the concentration is equal to a background value c_b. We then consider that the passage of a sinusoidal wave of amplitude a (with $a < h_1$, and traveling in the x-

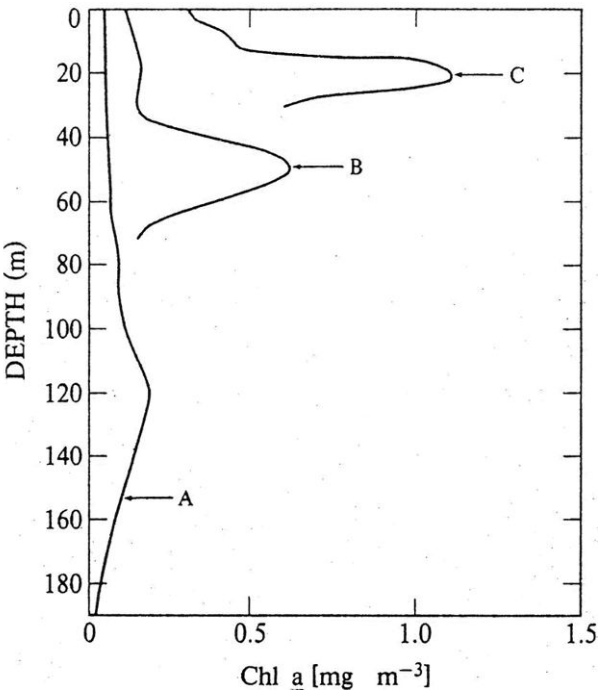


Figure 12.14. Typical chlorophyll profiles plotted as a function of geometrical depth (Cullen and Eppley, 1981): (A) North Pacific Central Gyre, near 28°N, 155°W (from Beers *et al.*, 1975); (B) Southern California Bight Sampling (SCBS) 15, Station 205; (C) SCBS 7, Station 102.

direction) will distort this layer by simply moving it upwards and downwards such that the depth distribution of the chlorophyll concentration is

$$c(z) = \begin{cases} c_b & \text{for } z \le h_1 + a\cos(kx - \omega t) \\ c_b + c_0 & \text{for } h_1 + a\cos(kx - \omega t) \le z \le h_2 + a\cos(kx - \omega t) \\ c_b & \text{for } z \ge h_2 + a\cos(kx - \omega t), \end{cases} \quad (12.16)$$

where k is the wavenumber of the wave; ω its frequency; and z increases with depth. Solving Equation (12.15) for the chlorophyll profile described by (12.16) leads to

$$c_{\text{sat}} = c_b \qquad \text{when} \quad z_{90} < h_1 + a\cos(kx - \omega t), \quad (12.17)$$

but if $h_1 + a\cos(kx - \omega t) < z_{90} < h_2 + a\cos(kx - \omega t)$, then

$$c_{\text{sat}} = c_b + \frac{c_0[\exp\{-2K(h_1 + a\cos(kx - \omega t))\} - \exp\{-2Kz_{90}\}]}{[1 - \exp\{-2Kz_{90}\}]}. \quad (12.18)$$

Thus in order for the satellite to measure enhanced chlorophyll, the top of the DCM will have to be lifted sufficiently upwards (see Figure 12.15) that $z_{90} > h_1 + a\cos(kx - \omega t)$ over part of the tidal cycle, and the largest values measured by the satellite will occur over the wave crest ($\cos(kx - \omega t) = -1$).

If we now substitute values of parameters typical for the Bay of Biscay into this model and take mean thermocline depth as 50 m, internal tidal amplitude a as 20 m,

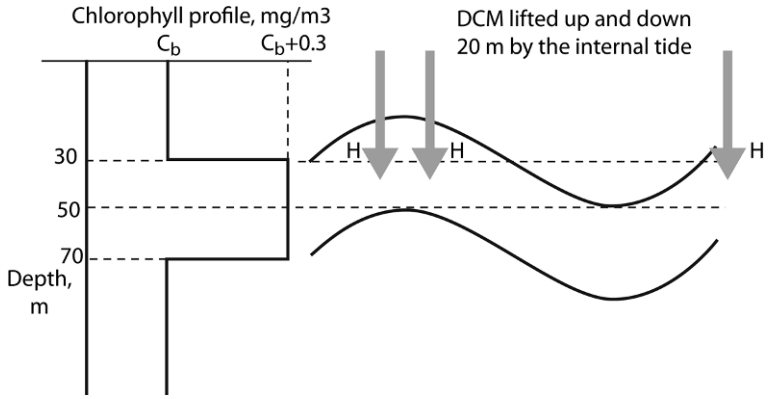


Figure 12.15. Schematic plot of chlorophyll profile and observation of the deep chlorophyll maximum (DCM) by the satellite sensor. The depth to which the sensor effectively "sees" is represented by H, where internal tide (IT) crests are observed as enhanced bands of chlorophyll.

and its wavelength as 40 km, we can reproduce quite well the chlorophyll concentration profile across the SeaWiFS image. Da Silva *et al.* (2002) assumed that the DCM was centered on the thermocline, was 40 m thick, and had chlorophyll levels enhanced by $c_0 = 0.3\,\mathrm{mg\,m^{-3}}$ over background values (as is typically observed on Atlantic Meridional Transect sections near these latitudes—A. Poulton, pers. commun.). Thus $h_1 = 30\,\mathrm{m}$ and $h_2 = 70\,\mathrm{m}$. Finally, we take $K = 0.05\,\mathrm{m^{-1}}$ (typical for the study region, which implies that $z_{90} = 20\,\mathrm{m}$) and the background value of $c_b = 0.26\,\mathrm{mg\,m^{-3}}$ directly from observed trough values in section X–Y (see Figure 12.16). These parameter choices then give the modeled chlorophyll distribution which is compared in Figure 12.16 with observed levels of c_{sat}, along the section X–Y. The modeled distribution (dashed line), chosen to match the positions of crests, is in remarkably good overall agreement with observed levels of c_{sat} (solid line). Although the regions of elevated chlorophyll are somewhat too narrow in the model, the actual increases at tidal crests are nicely captured. Note that, if we had employed a nonlinear tidal wave profile with a relatively broad crest compared with that of a sinusoidal wave, then the regions of elevated chlorophyll would have been correspondingly broader.

12.3.4 Internal waves and primary production

Primary production takes place in the top 50 m to 150 m of the water column (euphotic zone) where there is sufficient light for photosynthesis. The supply of nutrients is mostly from the pumping of nutrient-rich deep water to the euphotic zone through various mechanisms. However, traditionally accepted mechanisms are insufficient to explain observed productivity (McGillicuddy and Robinson, 1997). Consequently there have been intensive searches for new mechanisms to account for observed unexplained production. For long it has been speculated that internal

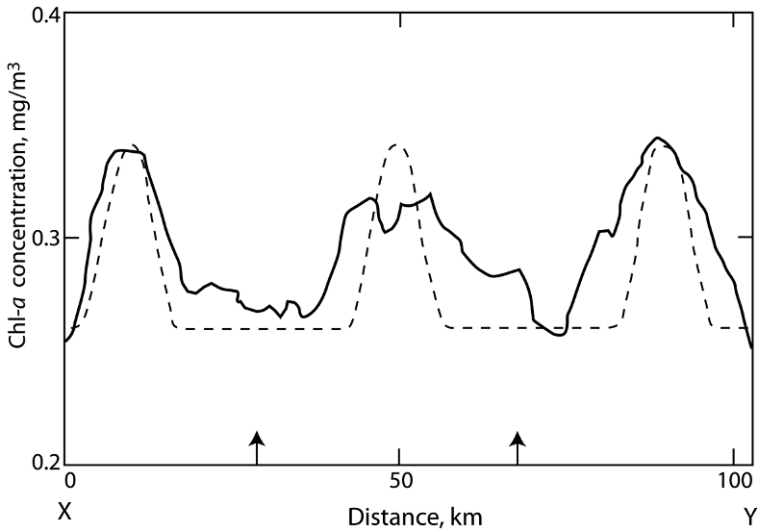


Figure 12.16. Chlorophyll concentration along section X–Y in Figure 12.13 observed by SeaWiFS (solid line) and as modeled by da Silva *et al.* (2002) (dashed line). The arrows show the positions of ISW packets (average position of the first two waves) and so indicate the troughs of the internal tides.

waves have a significant effect in primary production, but due to observational difficulty this process has been poorly quantified. Satellite remote sensing allied to *in situ* measurements has the potential to overcome the abovementioned constraint, making it possible to measure at frequent intervals over large spatial domains.

In the upper pycnocline, internal waves increase new primary production not only by creating shear and turbulence with consequent upward transport of nutrients, but also by increasing average light intensity experienced by phytoplankton there. Because light intensity decreases nonlinearly (exponentially) with depth, a neutrally buoyant or slowly sinking phytoplankton cell undergoing vertical displacements by internal waves is exposed to an average light intensity that is greater than the light intensity at its average depth during a day (in the absence of internal waves). If we assume that photosynthesis is proportional to total daily irradiance (because of dim light conditions near the euphotic zone), it is then clear that the vertical motion of internal waves may significantly increase primary production in eutrophic regions. Surprisingly, so far remote-sensing methods have not been applied to estimate net average enhancement of primary production which an intense internal wave field would experience, such as in this case of the Bay of Biscay.

The relationship between depth-dependent changes in photosynthesis and subsurface irradiance has been recognized for 50 years and its description remains a primary focus of productivity model development. Nearly three decades ago, Shulenberger and Reid (1981) demonstrated that net primary production in the region of the deep chlorophyll maximum often constitutes a substantial fraction of

total, depth-integrated primary production. At about the same time, Kahru (1983) estimated a substantial effect of high-amplitude, long-period internal waves (seiches) on depth-integrated primary production, using a model of eutrophic water in the Baltic Sea. Lande and Yentsch (1988) derived a simple model to estimate the increase in average light intensity on phytoplankton cells that are passively displaced by a random field of internal waves in the upper pycnocline, the lower portion of the euphotic zone. However, further efforts are needed if we are to fully explore the significance of internal waves in primary production, and determine whether they can actually explain the "missing" observed productivity of the global ocean. Remote sensing may play a key role in these efforts.

Imaging sensors, such as MERIS and MODIS with relatively higher spatial resolution (300 m) and wide swath widths, are now capable of resolving short-period signatures of internal waves as well as large-period, internal tidal waves, and may effectively be used to observe the effects of tidally generated internal waves on near-surface phytoplankton distribution. Such sensors may be very useful for determining the spatial scales of plankton distribution compared with the physical features of internal tidal waves and short-period internal waves. They may resolve whether the phytoplankton distribution apparently associated with IWs is simply a consequence of IWs revealing the deep chlorophyll to remote-sensing sensors, or is in fact evidence of enhanced primary production due to internal wave activity.

12.4 IMPACT OF REMOTE SENSING ON OUR KNOWLEDGE OF INTERNAL WAVES

Satellite remote sensing has had a remarkable impact on internal wave research, revealing the ubiquity of internal waves at the global scale (Jackson, 2007), their detailed horizontal structure in the near surface, and contributing to better understanding of their generation mechanisms (New and da Silva, 2002; Nash and Mourn, 2005; da Silva and Helfrich, 2008).

Nonimaging sensors, such as the TOPEX/Poseidon or Jason altimeters, have also made a remarkable contribution to the understanding of internal (tidal) waves in the ocean, providing global field maps of internal tides and their dissipation rates (e.g., Egbert and Ray, 2003). Even today, more than 30 years after the successful Seasat mission, high-resolution satellite images reveal previously unknown hotspots of internal wave activity close to sites where internal waves have been studied for decades, such as in Massachusetts Bay (da Silva and Helfrich, 2008).

Satellite images provide an important data source in remote regions where continuous, long-term *in situ* measurements are difficult to conduct, such as in the Mozambique Channel of the Indian Ocean (da Silva *et al.*, 2009). It is important to continue our efforts in satellite missions, aiming for continuous and more frequent satellite acquisitions, which would enhance our knowledge of internal waves. The use of automated procedures for identifying internal wave signatures on radar images (Simonin *et al.*, 2009) may eventually lead to routine scanning of all SAR data in

order to monitor internal waves worldwide. Future satellite observations are likely to shed light into exciting problems such as mixing by internal waves at the base of the mixed layer and in the thermocline, and will continue to reveal the generation mechanisms of these waves.

12.5 REFERENCES

Akylas, T. R., R. H. J. Grimshaw, S. R. Clarke, and A. Tabaei (2007), Reflecting tidal wave beams and local generation of solitary waves in the ocean thermocline. *J. Fluid Mech.*, **593**, 297–313, doi: 10.1017/S0022112007008786.

Alpers, W. (1985), Theory of radar imaging of internal waves. *Nature*, **314**, 245–247.

Apel, J. R. (1987), *Principles of Ocean Physics* (631 pp.). Academic Press, San Diego, CA.

Apel, J. R. (2004), Oceanic internal waves and solitons. In: C. R. Jackson and J. R. Apel (Eds.), *Synthetic Aperture Radar Marine User's Manual* (pp. 189–206). NOAA/NESDIS, Washington, D.C.

Apel, J. R., and F. I. Gonzalez (1983), Nonlinear features of internal waves off Baja California as observed from SEASAT imaging radar. *J. Geophys. Res.*, **88**(C7), 4459–4466.

Apel, J. R., H. M. Byrne, J. R. Proni, and R. L. Charnell (1975), Observations of oceanic internal and surface waves from the Earth Resources Technology satellite. *J. Geophys. Res.*, **80**(6), 865–881.

Apel, J. R., D. R. Thomson, D. G. Tilley, and P. van Dyke (1985), Hydrodynamics and radar signatures of internal solitons in the Andaman Sea. *Johns Hopkins APL Technical Digest*, **6**(4), 330–337.

Beers, J. R., F. M. H. Reid, and G. L. Stewart (1975), Microplankton of the North-Pacific Central gyre: Population structures and abundance, June 1973. *Int. Revue ges. Hydrobiol.*, **60**, 607–638.

Brandt, P., A. Rubino, W. Alpers, and J. O. Backhaus (1997), Internal waves in the Strait of Messina studied by a numerical model and synthetic aperture radar images from the ERS 1/2 satellites. *J. Phys. Oceanogr.*, **27**(5), 648–663.

Bretherton, F. P., and C. J. R. Garrett (1968), Wave trains in inhomogeneous moving media. *Proc. R. Soc. Lond. A*, **301**, 539.

Cullen, J. J., and R. W. Eppley (1981), Chlorophyll maximum layers of the Southern California Bight and possible mechanisms of their formation and maintenance. *Oceanologica Acta*, **4**, 23–32.

da Silva, J. C. B., S. A. Ermakov, I. S. Robinson, D. R. G. Jeans, and S. V. Kijashko (1998), Role of surface films in ERS SAR signatures of internal waves on the shelf, I: Short-period internal waves. *J. Geophys. Res.*, **103**(C4), 8009–8031.

da Silva, J. C. B., S. A. Ermakov, and I. S. Robinson (2000), Role of surface films in ERS SAR signatures of internal waves on the shelf, III: Mode transitions. *J. Geophys. Res.*, **105**(C10), 24089–24104, doi: 10.1029/2000JC900053.

da Silva, J. C. B., A. L. New, M. A. Srokosz, and T. J. Smith (2002), On the observability of internal tidal waves in remotely-sensed ocean color data. *Geophys. Res. Letters*, **29**(12), 1569, doi: 10.1029/2001GL013888.

da Silva, J. C. B., S. M. Correia, S. A. Ermakov, I. A. Sergievskaya, and I. S. Robinson (2003), Synergy of MERIS ASAR for observing marine film slicks and small scale processes. Paper presented at *Proc. MERIS User Workshop, November, Frascati, Italy*. ESA, Noordwijk, The Netherlands.

da Silva, J. C. B., A. L. New, and A. Azevedo (2007), On the role of SAR for observing "local generation" of internal solitary waves off the Iberian Peninsula. *Can. J. Remote Sensing*, **33**(5), 388–403.

da Silva, J. C. B., and K. R. Helfrich (2008), Synthetic aperture radar observations of resonantly generated internal solitary waves at Race Point Channel (Cape Cod). *J. Geophys. Res.*, **113**(C11016), doi: 10.1029/2008JC005004.

da Silva, J. C. B., A. L. New, and J. M. Magalhaes (2009), Internal solitary waves in the Mozambique Channel: Observations and interpretation. *J. Geophys. Res.*, **114**(C05001), doi: 10.1029/2008JC005125.

Drazin, P. G. (1983), *Solitons* (London Mathematical Society Lecture Note Series 85, viii + 136 pp.). Cambridge University Press, Cambridge, U.K.

Egbert, G. D., and R. D. Ray (2003), Semi-diurnal and diurnal tidal dissipation from TOPEX/Poseidon altimetry. *Geophys. Res. Letters*, **30**(17), 1907, doi: 10.1029/2003GL017676.

Ermakov, S. A., S. G. Salashin, and A. R. Panchenko (1992), Film slicks on the sea surface and some mechanisms of their formation. *Dynamics of Atmos. Oceans*, **16**, 279–304.

Ermakov, S. A., J. C. B. Da Silva, and I. S. Robinson (1998), Role of surface films in ERS SAR signatures of internal waves on the shelf, 2: Internal tidal waves. *J. Geophys. Res.*, **103**(C4), 8033–8043.

Gerkema, T. (2001), Internal and interfacial tides: Beam scattering and local generation of solitary waves. *J. Mar. Res.*, **59**, 227–255.

Gill, A. E. (1982), *Atmosphere–Ocean Dynamics* (International Geophysics Series Vol. 30, 662 pp.). Academic Press, San Diego, CA.

Gordon, H. R., and D. K. Clark (1980), Remote sensing optical properties of a stratified ocean. *Appl. Opt.*, **19**, 3428–3430.

Gordon, H. R., and W. R. McCluney (1975), Estimation of the depth of sunlight penetration in the sea for remote sensing. *Appl. Opt.*, **14**, 413–416.

Heathershaw, A. D. (1985), Observations of internal wave current fluctuations at the shelf-edge and their implications for sediment transport. *Continental Shelf Res.*, **4**, 485–493.

Helfrich, K. R., and W. K. Melville (2006), Long nonlinear internal waves. *Ann. Rev. Fluid Mechanics*, **38**, 395–425.

Holligan, P. M., R. D. Pingree, and G. T. Mardell (1985), Oceanic solitons, nutrient pulses and phytoplankton growth. *Nature*, **314**, 348–350.

Jackson, C. R. (2004), *An Atlas of Internal Solitary-like Waves and Their Properties* (Second Edition, prepared under contract for Office of Naval Research, Code 322PO, Contract N00014-03-C-0176, 560 pp.). Global Ocean Associates, Alexandria, VA.

Jackson, C. (2007), Internal wave detection using the Moderate Resolution Imaging Spectroradiometer (MODIS). *J. Geophys. Res.*, **112**(C11012), doi: 10.1029/2007JC004220.

Jeans, D. R. G., and T. J. Sherwin (2001), The variability of strongly non-linear solitary internal waves observed during an upwelling season on the Portuguese shelf. *Continental Shelf Res.*, **21**, 1855–1878.

Kahru, M. (1983), Phytoplankton patchiness generated by long internal waves: A model. *Mar. Ecol. Prog. Ser.*, **10**, 111–117.

Kantha, L. H., and C. A. Clayson (2000), *Small Scale Processes in Geophysical Fluid Flows* (International Geophysics Series Vol. 67, 888 pp.). Academic Press, San Diego, CA.

Killworth, P. D. (1998), Something stirs in the deep. *Nature*, **398**(24/31), 720–721.

Korteweg, D. J., and G. de Vries (1895), On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Philosphical Magazine*, **39**, 422–443.

Lamb, K. G. (1997), Particle transport by nonbreaking, solitary internal waves. *J. Geophys. Res.*, **102**(C8), 18641–18660.

Lande, R., and C. S. Yentsch (1988), Internal waves, primary production and the compensation depth of marine phytoplankton. *J. Plankton Res.*, **10**(3), 565–571.

Lennert-Cody, C. E., and P. J. S. Franks (1999), Plankton patchiness in high-frequency internal waves. *Mar. Ecol. Prog. Ser.*, **186**, 59–66.

Lighthill, M. J. (1978), *Waves in Fluids*. Cambridge University Press, Cambridge, U.K.

Liu, A. K., Y. S. Chang, M.-K. Hsu, and N. K. Liang (1998), Evolution of nonlinear internal waves in the East and South China Seas. *J. Geophys. Res.*, **103**(C4), 7995–8008.

Marmorino, G. O., G. B. Smith, and G. J. Lindemann (2004), Infrared imagery of ocean internal waves. *Geophys. Res. Letters*, **31**(L11309), doi: 10.1029/2004GL020152.

McGillicuddy, D. J., Jr., and A. R. Robinson (1997), Eddy-induced nutrient supply and new production in the Sargasso Sea. *Deep-Sea Res. I*, **44**, 1427–1450.

Melsheimer, C., and L. K. Kwoh (2001), Sun glitter in SPOT images and the visibility of oceanic phenomena. Paper presented at *Proc. 22nd Asian Conference on Remote Sensing, Singapore, November 5–9*.

Mowbray, D. E., and B. S. H. Rarity (1967), A theoretical and experimental investigation of the phase configuration of internal waves of small amplitude in a density stratified fluid. *J. Fluid Mech.*, **28**, 1–16.

Munk, W., and C. Wunch (1998), Abyssal recipes II. *Deep-Sea Res. I*, **45**, 1976–2009.

Nash, J. D., and J. N. Mourn (2005), River plumes as a source of large-amplitude internal waves in the coastal ocean. *Nature*, **437**(September), 400–403.

New, A. L., and J. C. B. Da Silva (2002), Remote-sensing evidence for the local generation of internal soliton packets in the central Bay of Biscay. *Deep-Sea Res. I*, **49**, 915–934.

New, A. L., and R. D. Pingree (1990), Large-amplitude internal soliton packets in the central Bay of Biscay. *Deep-Sea Res. I*, **37**, 513–524.

New, A. L., and R. D. Pingree (2000), An intercomparison of internal solitary waves in the Bay of Biscay and resulting from Korteweg–de Vries-type theory. *Prog. Oceanogr.*, **45**(1), 1–38.

Osborne, A. R.. and T. L. Burch (1980). Internal solitons in the Andaman Sea. *Science*, **208**(4443), 451–460.

Ostrovsky, L. A., and Y. A. Stepanyants (1989), Do internal solitons exist in the ocean? *Rev. Geophys.*, **27**, 293–310.

Pedlosky, J. (2004), *Ocean Circulation Theory* (453 pp.). Springer-Verlag, New York.

Pineda, J. (1999), Circulation and larvae distribution in internal tidal bore warm fronts. *Limnol. Oceanogr.*, **44**(6), 1400–1414.

Pingree, R. D., and A. L. New (1995), Structure, seasonal development and sunglint spatial coherence of the internal tide on the Celtic and Armorican shelves in the Bay of Biscay. *Deep-Sea Res. I*, **42**, 245–284.

Pingree, R. D., D. K. Griffiths, and G. T. Mardell (1983), The structure of the internal tide at the Celtic Sea shelf break. *J. Mar. Biol. Assoc. U.K.*, **64**, 99–113.

Pingree, R. D., G. T. Mardell, and A. L. New (1986), Propagation of internal tides from the upper slopes of the Bay of Biscay. *Nature*, **321**, 154–158.

Robinson, I. S. (2004), *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Shanks, A. L. (1983), Surface slicks associated with tidally forced internal waves may transport pelagic larvae of benthic invertebrates and fishes shoreward. *Mar. Ecol. Prog. Ser.*, **13**, 311–315.

Shulenberger, E., and J. Reid (1981), The Pacific shallow oxygen maximum, deep chlorophyll maximum and primary productivity, reconsidered. *Deep-Sea Res. I*, **28**, 901–919.

Simonin, D., A. R. Tatnall, and I. S. Robinson (2009), The automated detection and recognition of internal waves. *Int. J. Remote Sensing*, **30**(17), 4581–4598.

Vesecky, J. F., and R. H. Stewart (1982), The observation of ocean surface radar phenomena using imagery from the Seasat Synthetic Aperture Radar: An assessment. *J. Geophys. Res.*, **87**, 3397–3430.

Zatsepin, A. G., A. S. Kazmin, and I. N. Fedorov (1984), Thermal and visible manifestations of large internal waves. *Oceanologia*, **28**, 586–592.

13

Shelf seas, estuaries, and coasts

13.1 INTRODUCTION

This chapter reviews how satellite data contribute to the research and applications of oceanographic knowledge in shelf seas, estuaries, and coastal waters. The reason for devoting a chapter to this subject is that remote sensing in shallow seas and in the near-shore marine environment presents some challenges and opportunities that differ from those encountered in open-ocean satellite oceanography. Where the continental shelf extends more than a few tens of kilometers offshore the dynamics of water movement have a distinct character, which presents a variety of different phenomena to be observed and affects the way that satellite data are interpreted. This is developed in Section 13.2, which shows the importance of medium-resolution imaging sensors for shelf sea oceanography. Until quite recently, the use of altimetry in shelf seas was largely ruled out as impractical because of severe limitations of accuracy. However, recent research activity has pointed the way to promising developments in coastal altimetry, and these are outlined in Section 13.3. As we approach very close to the shore, or where the sea penetrates the coastline in estuaries, standard ocean-imaging sensors lack the required spatial resolution. Having to use alternative sensors has given a different character to the methods of coastal and estuarine remote sensing, which are outlined in Section 13.4, although there is scope here for little more than an introduction to what could become a fairly extensive subject if allied topics such as the remote sensing of lakes, deltas, and wetlands were included. Readers should note that the chapter on ocean biology from space deals in Sections 7.5 and 7.6 with topics that might have been included here—habitats in shallow tropical seas and warm-water coral reefs—while Chapters 8 on waves and 9 on winds contain some coastal applications too.

13.2 OBSERVING SHELF SEAS FROM SPACE

13.2.1 What is distinct about the remote sensing of shelf seas?

Shelf seas are found at the margins of the main oceans of the world, where the sea has inundated low-lying parts of the continental landmass. From a geophysical perspective, the Earth's crust where there are continental landmasses is much thicker than the crust beneath the ocean's abyssal planes. Over geological history the continental plates have moved around the Earth's surface, driven by tectonic processes, to produce the present shape of the World Ocean. At many boundaries between continental plates and oceanic plates there is an abrupt edge to the landmass and within a short distance offshore the seabed slopes steeply down to depths of a few kilometers. In such cases, properties characteristic of the deep ocean are found close to the coast and there is no need to consider a special remote-sensing approach that is different from observing the open ocean although, as shown in Chapters 4 and 5, satellites have an important role to play in observing dynamical phenomena associated with the land boundaries of the ocean, such as coastal upwelling, boundary currents, and their associated fronts.

In other parts of the world, the edges of continental plates have been inundated by rising sea levels and/or by sinking of the continental landmass relative to sea level, to form shallow seas with a depth typically less than 200 m. Sometimes these shelf regions stretch hundreds of kilometers offshore until, at the edge of the continental plate, there is an underwater cliff where the seabed falls away steeply to a depth of 1 km to 2 km or more. The sudden contrast of ocean depth at the continental shelf edge is not immediately apparent to someone at sea level or observing the ocean from above, although there may be subtle evidence to look for in satellite data (see Section 13.2.3). Nonetheless the huge depth difference often results in the processes and properties of the shelf sea being characteristically different from the adjacent deep ocean, in terms of physical and chemical phenomena, biology, and sedimentation. It is these different characteristics of shelf seas that present new opportunities and challenges for the application of remote-sensing techniques and the interpretation of satellite data. This is why it is useful to consider the satellite oceanography of shelf seas separately (Nihoul *et al.*, 1998).

While a narrow continental shelf is found around nearly all coastlines, the particular issues considered in this chapter are applicable mostly in regions where the continental shelf is many tens or hundreds of kilometers wide and may well be semi-enclosed by surrounding landmasses and islands. Figure 13.1 highlights locations around the world where such seas are found. Foremost in terms of oceanographic study and knowledge are regions such as northwest European marginal seas (including the North Sea, Baltic Sea, Irish Sea, English Channel, and Celtic Sea), the Nova Scotian shelf and Newfoundland Grand Banks off northeast America, the Patagonian shelf east of Argentina, the Gulf of Thailand, the Malaysian shelf and Java Sea, the East China Sea, the Yellow Sea and the Bering Sea. However, there are several other coastlines where the continental shelf is over 100 km wide (as shown in Figure 13.1) and where the approach discussed in this

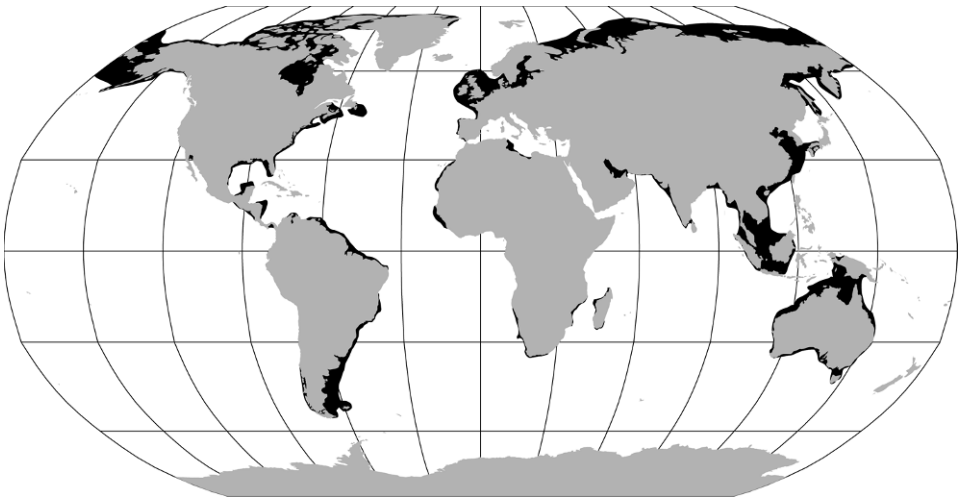


Figure 13.1. World map showing in black the regions where the continental shelf (bathymetric depth <200 m) extends more than about 25 km from the coast. Where these are wide or enclosed the region is likely to display some of the characteristics of shelf seas discussed in this chapter.

chapter is applicable. These include parts of the Brazilian coast, the west African coast between Senegal and Sierra Leone, the Timor Sea and Gulf of Carpentaria off North Australia, the Bass Strait, and the seas along the Siberian and Russian coasts of the Arctic Ocean. This section of the chapter will rely mainly on examples from northwest European shelf seas to illustrate principles that should be applicable in most other shelf seas with large tidal ranges. This is done for convenience, given the ready availability of high-resolution image data and because there is a fairly extensive literature on theoretical and observational studies of the oceanography of this region.

Because the important lengthscales and timescales for shelf seas differ from the open ocean, this has consequences when applying the methods of ocean remote sensing to shallow tidal seas (as Section 13.2.2 will explain). A characteristic of most shelf seas is the strength of tidal currents which provide a source of energy for mixing and stirring the water. This can result in shelf seas being vertically well-mixed, which has the beneficial consequence for remote sensing observations that what satellites are able to measure at the sea surface is characteristic of the whole water column. Where stratification does occur, it is constrained by tidally driven processes to form some interesting dynamical phenomena which lend themselves to study from satellite infrared sensors (as explained in Section 13.2.4).

Another characteristic of shelf seas is the influence of runoff from the adjacent continental landmass, leading to some characteristic regions of freshwater influence (ROFI). River inputs can also significantly increase the concentration of suspended sediments in shelf seas, and these are often highly visible from space, especially where tidal currents resuspend seabed material and bring it to the surface in regions where

there is no stratification. Remote sensing of suspended sediments is discussed in Section 13.2.5. Rivers and land runoff also provide a rich source of nutrients that support algal blooms, with potentially deleterious consequences for water quality. Remote-sensing methods, especially those using ocean color, are in demand for monitoring the health of shelf seas and their ecosystems, although the diversity of sources of colored material optically creates Case 2 conditions, which challenges the capacity to make reliable quantitative measurements. This is discussed in Section 13.2.6.

13.2.2 Variability scales in shelf seas

When using remote-sensing methods to study an aspect of oceanography, it is always important to ensure that the sampling capabilities of satellite data in space and time are optimized to match broadly the spatial and temporal variability characteristics of the ocean phenomenon. This principle is developed and explained in chapter 4 of *MTOFS* (Robinson, 2004). The earlier chapters of the current book have shown the importance of the oceanic mesoscale for setting the sampling requirements for observing many ocean processes. Moving our attention to shelf seas forces us to reconsider scales that need to be detected, which turn out to differ in some important respects from what is encountered in the open ocean. The key factor that causes different behavior is the shallow depth of shelf seas.

Considering the dynamical behavior of water movement in shelf seas, the shallow depth has a profound effect because it causes wave-like perturbations to propagate at a much slower speed. For example, in water 40 m deep, barotropic long waves travel at 20 m/s or 72 km/h, compared with a speed 10 times faster in water of depth 4,000 m. This mismatch of speeds has a strong effect on ocean tides. The much slower energy propagation speed causes the tidal amplitude to increase over shelf seas, resulting in much stronger tidal currents. The slower propagation speed also allows the tidal phase to change rapidly over shorter distances. This, and the creation of shallow-water tidal constituents by the nonlinear effects of high amplitudes makes it much harder to predict tides from altimetry. Until recently this severely limited the use of satellite altimetry in shelf seas, prompting the recent development of new techniques for coastal altimetry that are discussed in Section 13.3.

Baroclinic perturbations in the water column also show differences of propagation characteristics between the deep ocean and shallow shelf seas. Chapter 12 showed how internal waves at the thermocline are produced by the tidal flow across the shelf break. On the other hand, ocean currents and baroclinic eddies tend not to propagate from the deep ocean onto the adjacent shelf. Because a column of water possesses planetary vorticity by virtue of the Earth's rotation, it would be forced to spin up or down if it moved off or on to the shelf in order to conserve its absolute vorticity. In practice when water starts to flow onto the shelf the induced vorticity deflects it. This tends to constrain water movement to follow bathymetric contours, or more precisely contours of constant f/h (where f is the Coriolis parameter; and h is the water depth or the depth of a water layer in the baroclinic case). Consequently this inhibits flow between the shelf sea and the

neighboring ocean, behavior that is expressed theoretically in the Taylor–Proudman theorem (Brink, 1998). Coastal boundary currents in the deep ocean tend to follow the line of the shelf break rather than flowing on to the shelf. All of these factors tend to result in shelf seas developing water mass characteristics that differ from the adjacent ocean. A more complete description of the characteristic fluid dynamics of shelf seas can be found in a number of review articles and textbooks (e.g., Huthnance, 1995; Csanady, 1997; Hill, 1998; Simpson, 1998; Mann and Lazier, 2006).

The semi-isolation from the adjacent deep ocean, the shallowness of seas such that sea bed friction can constrain the whole water column, the efficiency of strong tides and winds to mix the water column vertically, and the inflow of large rivers draining a wide hinterland are all factors which tend to prevent the sea from becoming horizontally homogeneous. Horizontal stirring is much less effective than in the deep ocean for mixing together waters with different temperatures, salinities, and biogeochemical properties. This tends to reduce variability lengthscales in shelf seas, compared with mesoscale behavior in the deep ocean, and results in the formation of a number of distinctive shelf sea phenomena whose characteristics are summarized in Table 13.1. If satellite observations are to serve the science of shelf sea oceanography, or to contribute to operational monitoring of some of the busiest seas in the world, then they must be capable of observing these processes and features.

The horizontal scale of those surface water properties and contents that are observed from satellites (in particular, temperature, chlorophyll, and suspended particulates) is controlled by factors which affect vertical-mixing processes. These factors are the amplitude, U, of tidal currents associated with the dominant tidal species (typically semidiurnal but possibly diurnal or mixed diurnal–semidiurnal in some parts of the world) and the depth, h, of the water column. Because of the relatively slow propagation speed of the tidal phase in shelf seas, the lengthscale

Table 13.1. Length and timescale of shelf sea processes and phenomena observed from satellites.

Process	*Horizontal variability lengthscale*	*Horizontal extent* (km)	*Vertical structure*	*Variability timescale*	*Lifetime*
Tidal mixing fronts	3–7 km	50–200	Stratified/mixed	5–7 days	Seasonal
Regions of freshwater influence	2–10 km	10–100	Stratified/mixed	5–7 days	Seasonal
Near-surface suspended sediments	2–50 km	500	Vertically mixed	1–5 days	Annual
Phytoplankton blooms	1–20 km	20–400	Stratified/mixed	12 h–5 day	Seasonal
Shelf edge fronts	10–50 km	50–250	Baroclinic	5–7 days	Seasonal
Internal waves	500 m–10 km	200	Stratified	1 h	2–5 days

over which U varies because of tidal phase gradients can be as short as 30 km in the vicinity of an anti-node in the co-tidal map. However, where tidal currents flow across bathymetric contours, U tends to vary inversely with h and so tends to vary with the lengthscale of bathymetric depth along tidal streamlines. In regions of sedimentary bed features this can be as short as 1 km to 2 km; close to coastlines and for curved streamlines flowing around headlands and islands it may be 3 km to 10 km. This accounts for the rather short lengthscales of the phenomena shown in Table 13.1, which are individually discussed in subsequent subsections. These are the scales that need to be resolved by imaging sensors observing shelf seas.

It is worth emphasizing that a fundamental element of interpreting satellite images of shelf seas is to know, or to make a judgment about, whether the water column is stratified or not. As Section 13.2.4 explains, the energy density available for vertical mixing is proportional to U^3/h. In the deeper parts of the shelf sea and away from coasts and islands, where U^3/h is sufficiently small, the water column will become stratified, at least during summer months. Here, surface water properties are largely independent of tidal flow over the sea bed and are expected to be fairly uniform with a variability lengthscale greater than 100 km. However, where mixing is vigorous enough, stratification breaks down so that surface properties measured by satellites are representative of the whole water column. Here the surface properties become sensitive to mixing from the sea bed and the properties measured by satellites may vary at the relatively short lengthscales identified in Table 13.1, which are associated mainly with the lengthscale of the bathymetry. On the other hand, if short-lengthscale inhomogeneities are detected in regions of shelf seas known to be stratified, then these are likely to be associated with perturbations of the thermocline. The primary cause is likely to be internal waves. The same is true at the shelf break (discussed in Section 13.2.3).

The timescale that most influences shelf sea dynamics comes from tidal forcing and is the spring–neap modulation caused by the 14-day beat frequency between lunar and solar tide–generating forces. The amplitude of tidal elevation and tidal currents at spring tides may be between 1.5 or 2 times those at neap tides, with 7 days between neap and spring and another 7 days back to neap tide conditions. Thus over a 14-day cycle the energy available for vertical mixing or for resuspending sea bed particulates varies by the spring–neap amplitude ratio raised to a power between 2 and 3 (i.e., by a factor of between 3 and 8). This is the basis for specifying the 5-day to 7-day timescale for processes that are driven mainly by the strength of tidal currents. In practice the cycle of tidal rise and fall has a 12.5 h period and tidal current speed varies at twice that frequency, but it is generally assumed that water column stratification or thermal structure does not change appreciably on such a short timescale of a few hours. However, the sudden onset of very strong winds can rapidly resuspend seabed sediments, reducing the variability timescale for that process, while the shortest timescale for phytoplankton populations may be determined biologically by their natural growth rate and the physiological response to enhanced sunlight or to the availability of nutrients.

Table 13.2 lists the sampling characteristics of sensors used for mapping shelf sea processes, from which it is evident that the use of medium-resolution sensors (infra-

Table 13.2. Sampling capabilities of sensors used to observe shelf seas (other potential sensors shown in italics).

Sensor type	Measured ocean properties	Spatial sampling resolution	Temporal sampling (revisit interval)	Comments
Infrared radiometer	Sea surface temperature	Polar orbit: 1–2 km Geostationary orbit: 3–5 km	~12 h 1 h	Cannot penetrate cloud
Ocean color sensor	Chlorophyll concentration, suspended particulates, colored organic material, optical properties	1–2 km	1–2 days	Cannot penetrate cloud
Radar scatterometer	Wind speed and direction	25 km	Twice daily	
Imaging radar	Surface roughness	25–75 m	~10 days	
Microwave radiometer	*SST* *Wind speed, sea ice*	*25–50 km,* *10–25 km*	*1 day*	*Unreliable within 100 km of land*
Radar altimeter	*Sea surface height, geostrophic currents, wind speed, wave height*	*~7 km along ground track, ~50 km mapped with multimission data*	*7 days mapped with multimission data*	*Presently unreliable within 20–50 km of land*

red and visible waveband imagers) is well matched to the requirements for spatial sampling, despite the finer scales to be resolved. However, in the same way as observing mesoscale phenomena discussed in Chapters 3 to 5, the dependence of these sensors on cloud-free conditions means that they are not able to follow the most rapid changes of some processes except when clear skies are sustained continuously for several days. In many parts of the world this occurs infrequently.

Another consequence of tides for observing shelf seas is the way that the position of certain phenomena can be advected backwards and forwards by tidal currents. For example, if there are strong rectilinear tidal currents with an amplitude of 1 m/s a buoy carried by the tide would have a tidal excursion of about 14 km. Phytoplankton blooms might be expected to be carried by the tidal excursion and when observed in satellite images their position may vary with the phase of the semidiurnal tide. A single polar-orbiting sensor is typically locked to the solar semidiurnal cycle and so would alias this tidal movement. On the other hand those phenomena associated with bathymetric features should remain geographically locked, although strong tidal currents may tend to smear them out. It would therefore be very interesting to use satellite data to detect whether, and to what extent, spatial structures in

temperature or color fields follow tidal excursions. To do this would require either a combination of sensors with different overpass times during the day or else a single geostationary sensor. At present the spatial resolution of geostationary sensors in midlatitudes is about 5 km and this would not reliably detect tidal excursions over most shelf seas where tidal streams are less than 0.5 m/s. The issue of tidal advection remains a potential source of uncertainty, to be considered when interpreting satellite data from regions of strong tidal streams, although little has been reported about it in the literature of shelf sea remote sensing.

13.2.3 Shelf edge phenomena

We now consider examples of the different phenomena of shelf seas that are detected by satellites, starting with those associated with the continental shelf edge itself, or the region above the slope on the ocean side of the edge. Figure 13.2 shows schematically three processes that are capable of producing a surface signature in satellite image data, given suitable conditions. Although these are three quite distinct mechanisms, they have in common that the resulting signatures are associated with, and geographically tied to, the shelf break.

The first is observed when water masses on and off the shelf have different properties for the reasons mentioned in Section 13.2.1. To be detectable, the contrasting property must be one that is observable from satellite sensors. Figure 13.3 shows an example from SST distribution over the northwest European shelf edge in early April 2006. At this stage in the annual thermal cycle of the upper ocean (see

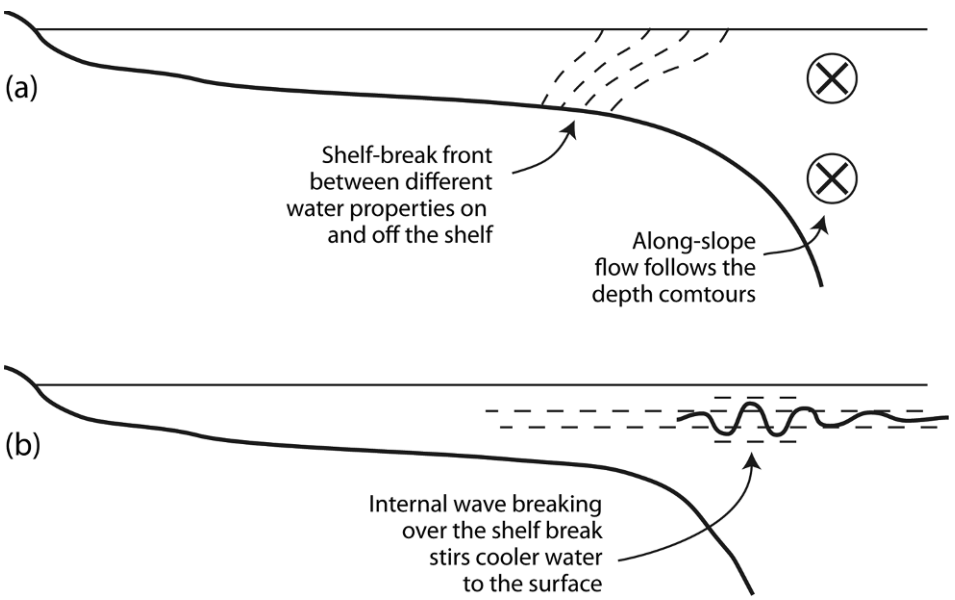


Figure 13.2. Schematic showing different shelf edge processes that create remote-sensing signatures. (a) Shelf break front and along-slope current. (b) Internal wave breaking.

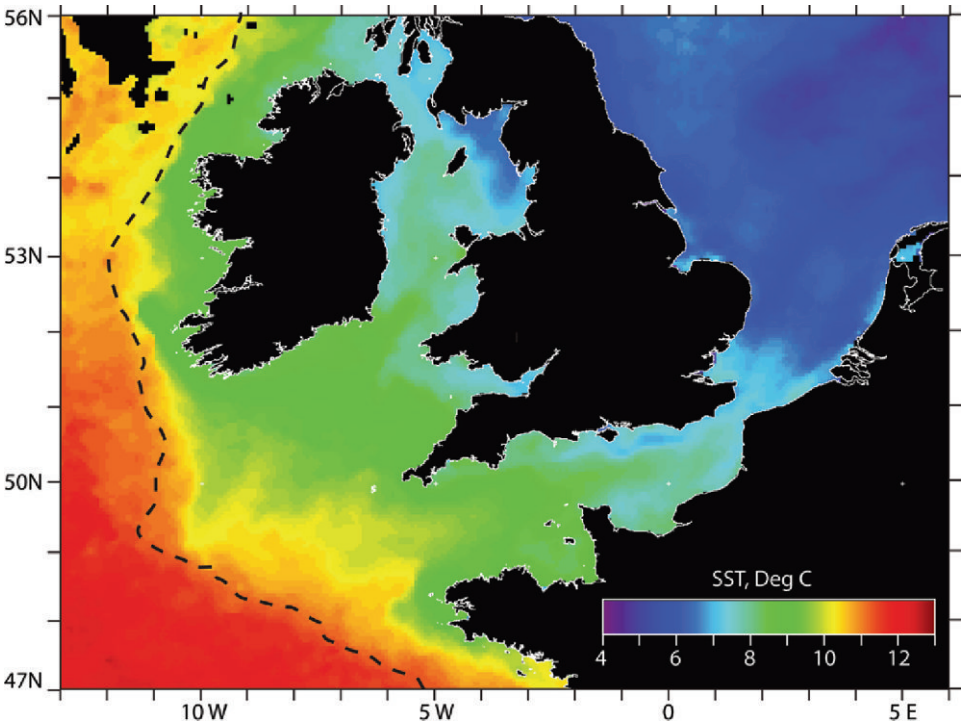


Figure 13.3. SST composite image from AVHRR NOAA-18 over northwest European shelf seas for the week ending April 8, 2006 showing cooler water over the shelf. Black is land or persistently cloudy pixels. The 11°C SST contour approximately follows the shelf edge which is denoted by the dashed line. The figure was created using weekly composite image data produced by NEODAAS (*http://www.neodaas.ac.uk/*).

Figure 6.3) the SST in the North Atlantic Ocean is increasing as the open-ocean seasonal thermocline develops and traps solar heating closer to the surface, whereas over the shelf the enhanced vertical mixing through the water column hinders the development of the seasonal thermocline and so the rise in SST lags behind that in the open ocean. Consequently, for this geographical area at this stage in the annual cycle, although not generally throughout most of the year, SST changes sharply by about 2°C across the shelf edge. In Figure 13.3 the 11°C SST contour follows the characteristic shape of the shelf edge west of Ireland and then south and east towards Brittany.

The second process that produces signatures of the shelf edge in certain satellite images is the presence of along-slope currents flowing on the deep-ocean side of the continental edge along contours of constant f/h (also shown schematically in Figure 13.2a). Some of the strongest examples of such flows have already been encountered in this book (e.g., the Gulf Stream, the Kuroshio Current, and the Agulhas Current presented in Chapter 4). However, weaker and narrower currents are revealed in

other parts of ocean margins wherever the current has a distinctive property that contrasts with the surrounding water. Figure 13.4 shows a good example in the northeast Atlantic where the current flowing northwards along the edge of the northwest European continental slope has been colored by what is probably a coccolithophore bloom with a distinctive broadband reflectance of sunlight. In this enhanced real-color composite of atmospherically corrected, normalized water-leaving radiance, derived from MODIS bands at 667 nm, 551 nm, and 443 nm, the whiter color along the shelf edge near the southwest corner of Ireland contrasts with generally bluer (clearer) water on the ocean side and greener (more chlorophyll-rich) waters on the shelf side, demonstrating how the shelf break acts as a boundary hindering water exchange. The whiter colored water serves as a tracer of the current, showing it to follow the depth contour around the shelf edge. The way in which the bloom seems to be sustained over hundreds of kilometers suggests that the slope current may also be promoting local vertical mixing of nutrients at the shelf edge, although without corroborative *in situ* observations this remains speculation.

A variant of the shelf break current signature is shown in Figure 13.5 where the path of the current along the shelf edge northeast of Scotland, and then turning south with the contour to follow the Norwegian deep, is traced by a trail of warmer SST, contrasting with cooler water both on and off the shelf. Compared with a month earlier (see Figure 13.3), western shelf waters farther south have warmed to become indistinguishable from off-shelf water, a reminder that SST distribution is continuously evolving throughout the year in this dynamic region.

The third shelf edge signature is that caused by breaking internal waves. As mentioned in Chapter 12, internal tidal waves in the deep ocean change their characteristics greatly on encountering the shelf break. As the depth shallows their propagation speed reduces and their amplitude grows to produce the packets of higher frequency internal waves that are observed in SAR images. As part of the same process, increased wave amplitude may cause breaking, promoting mixing which weakens the thermocline, raising cooler and possibly nutrient-rich water to the surface. Although medium-resolution radiometers on satellites do not detect internal tides explicitly, they can reveal zones of cooler water along the shelf break (as in Figure 13.6). This shows monthly averaged SST over the western approaches to the U.K. for June 2004, in which the SST is 0.5° to 1°C cooler over the shelf break than on either side. Although it is difficult to find images from individual overpasses that show the shelf edge thermal signature so consistently, probably because internal wave energy fluctuates with the spring–neap cycle and because of persistent, patchy cloud cover, its appearance in this monthly composite is evidence that when cooling does occur it happens over a geographically restricted area along the shelf break. Shelf break mixing causing cooling appears to be limited to the summer months and is strongest in June.

Confirmation that this is genuinely a mixing phenomenon that delivers a supply of additional nutrients to surface water is to be found in data from ocean color sensors. Figure 13.7 shows monthly averaged chlorophyll concentrations for June 2004 over the Bay of Biscay and the Celtic Sea, accumulated from the SeaWiFS-derived chlorophyll record. At this time of year, although there is enhanced produc-

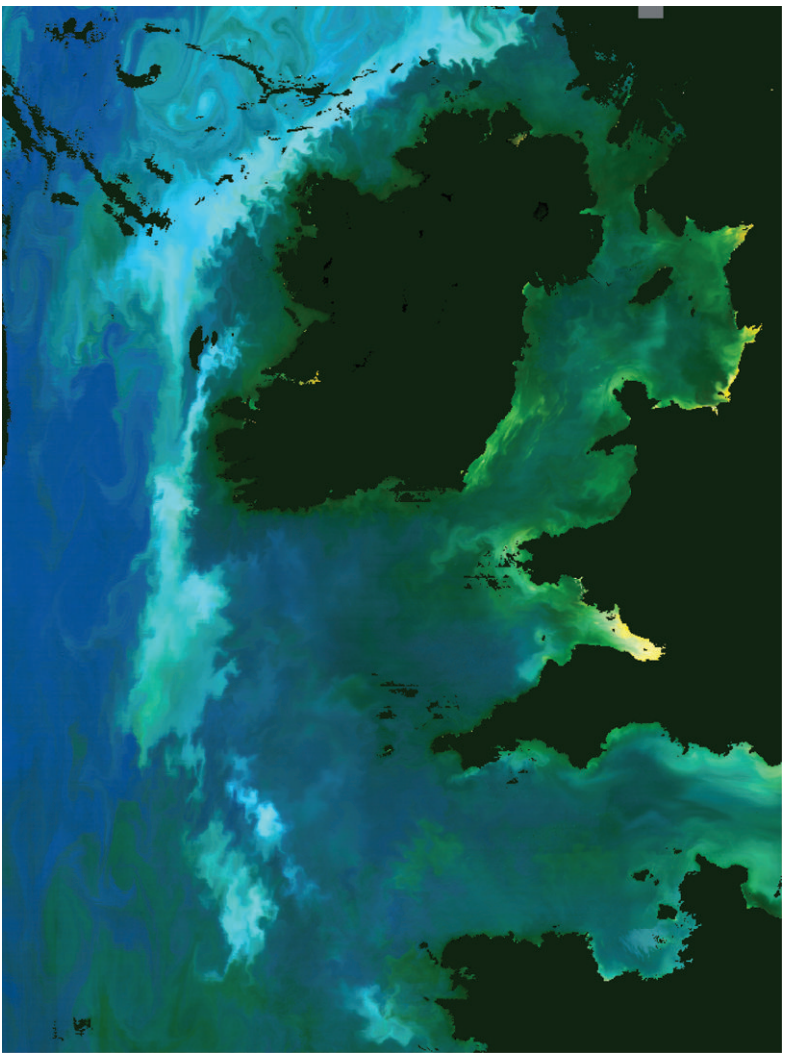


Figure 13.4. Enhanced color composite consisting of normalized water-leaving radiance generated from MODIS (Aqua) data for June 2, 2006 over northwest European coastal waters. The data from wavebands at 667 nm, 551 nm, and 443 nm are overlaid as red, green, and blue images, respectively. The bright white and slightly blue color found all along the line of the shelf break corresponds to high reflectance across the spectrum, typically the result of coccolithophores in the upper mixed layer. Black corresponds to land or detected cloud cover. Where the color is blue corresponds to open-ocean clear water and where it is greener can be associated with the presence of phytoplankton. The bright coastal regions with a yellow tint indicate that there is higher reactance in the red and lower in the blue part of the spectrum, typically the case where there is resuspended sediment producing high turbidity and dissolved organic material absorbing the blue light (image produced by the author using level 2 MODIS data obtained from the NASA Ocean Color website—*http://oceancolor.gsfc.nasa.gov/*).

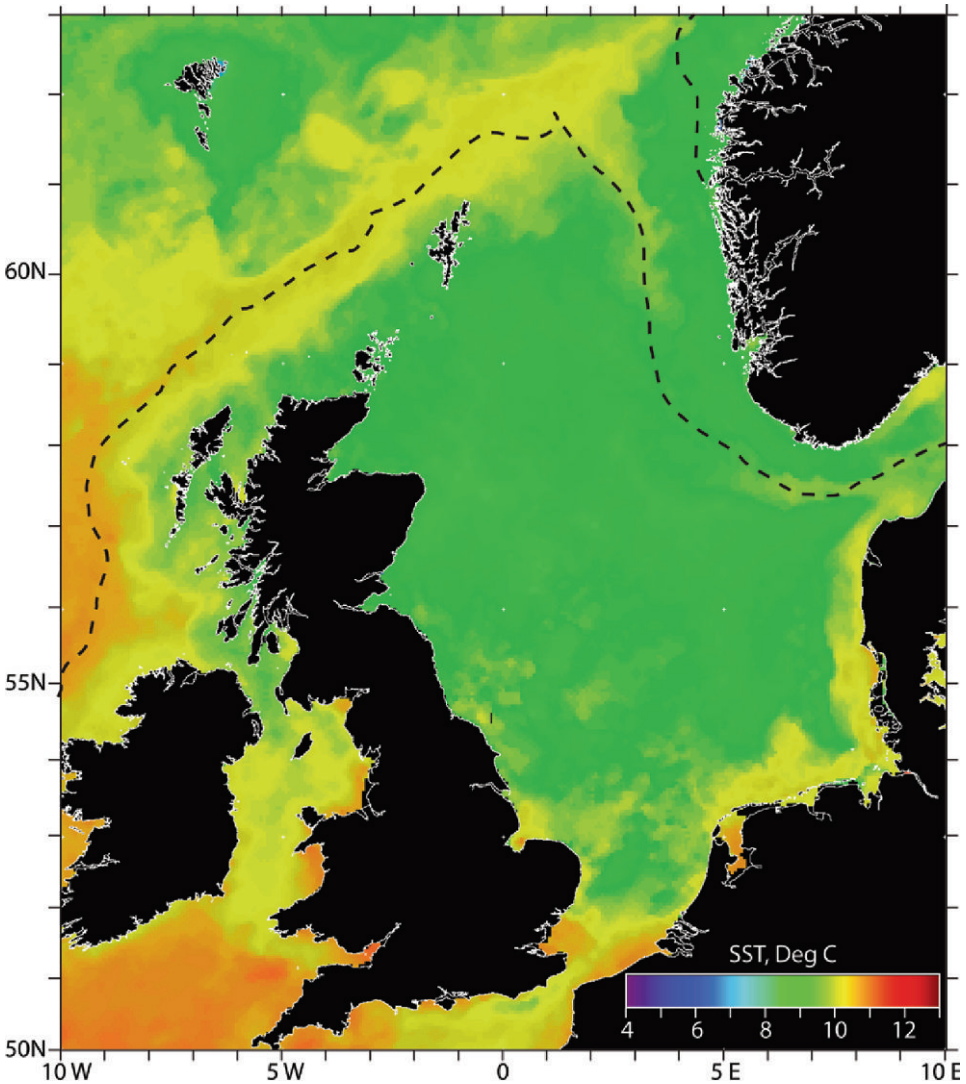


Figure 13.5. SST weekly composite image from AVHRR NOAA-18 over northwest European shelf seas for the week ending May 6, 2006 showing a plume of warmer water transported by the shelf edge current north of Scotand. Black represents land or persistently cloudy pixels. The dashed line shows the position of the shelf edge. The figure was created using weekly composite image data produced by NEODAAS (*http://www.neodaas.ac.uk/*).

tion along the coast and in shallower parts of the shelf region, there is very little in the outer shelf and open ocean, apart from the clear line of raised concentration that precisely follows the line of the continental edge right out to the knee of the shelf edge at about 49°N, 11°W. Chlorophyll concentration along this line is five to ten

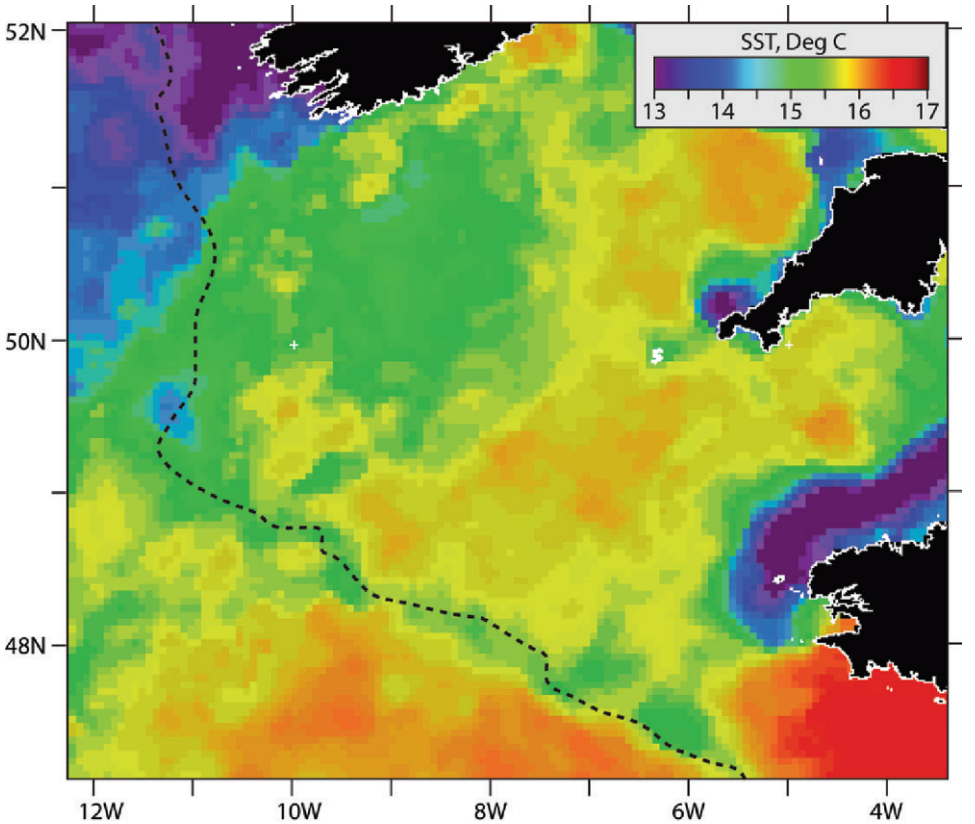


Figure 13.6. SST monthly composite image from AVHRR NOAA-18 over U.K. western approaches for the month of June 2004. It shows a region of cooler SST along much of the line of the shelf edge, marked by the dashed line (figure created using monthly composite image data produced by NEODAAS—*http://www.neodaas.ac.uk/*).

times greater than on either side. While such enhancement of production has been observed by *in situ* measurements, it is the spatial distribution detected from satellite data that provides compelling evidence that this is a phenomenon geographically restricted to the shelf edge.

13.2.4 Thermal signatures of shelf sea dynamical phenomena

There is a lot to be learned from examining SST maps derived from thermal infrared images over the continental shelf, such as those shown in Figure 13.8 which contains four SST weekly composites based on AVHRR, selected to represent different seasons through the annual cycle. While it may seem instinctive to account for the patterns of SST as if they represented the response to horizontal advection and stirring, this is not in general a very helpful approach for interpreting images of

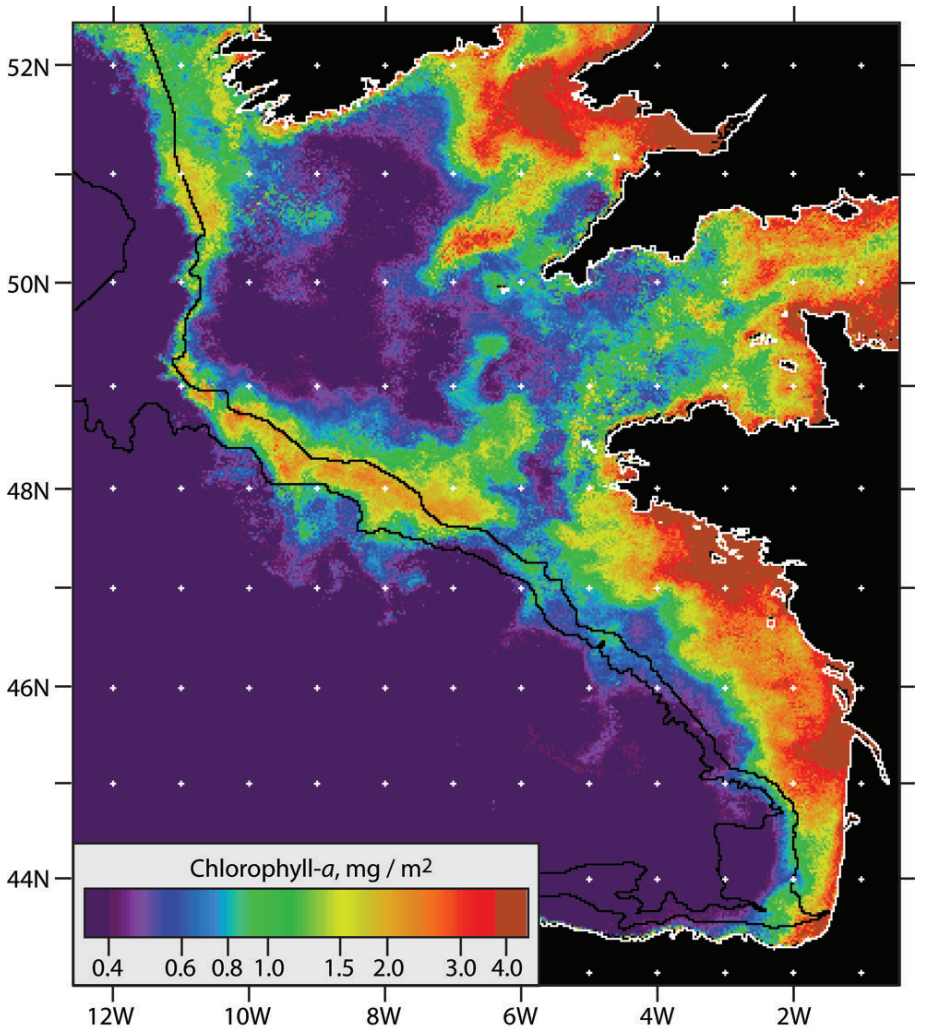


Figure 13.7. Image of monthly averaged chlorophyll-*a* concentration for June 2004 derived from SeaWiFS. The position of the continental slope is delineated by the 200 m and 1,000 m bathymetric contours drawn as black lines. There is a distinct enhancement of chlorophyll concentration over the slope between 11°W and 7°W (figure created using monthly composite image data produced by NEODAAS—*http://www.neodaas.ac.uk/*).

shelf seas. The shallow depth and oscillatory tidal currents tend to promote vertical mixing which destroys stratification where the sea is shallow enough, and also inhibits lateral circulation. This means that horizontal gradients of temperature owe more to variations of water depth and of heat exchange through the surface than to horizontal heat flow by advection or turbulent exchange.

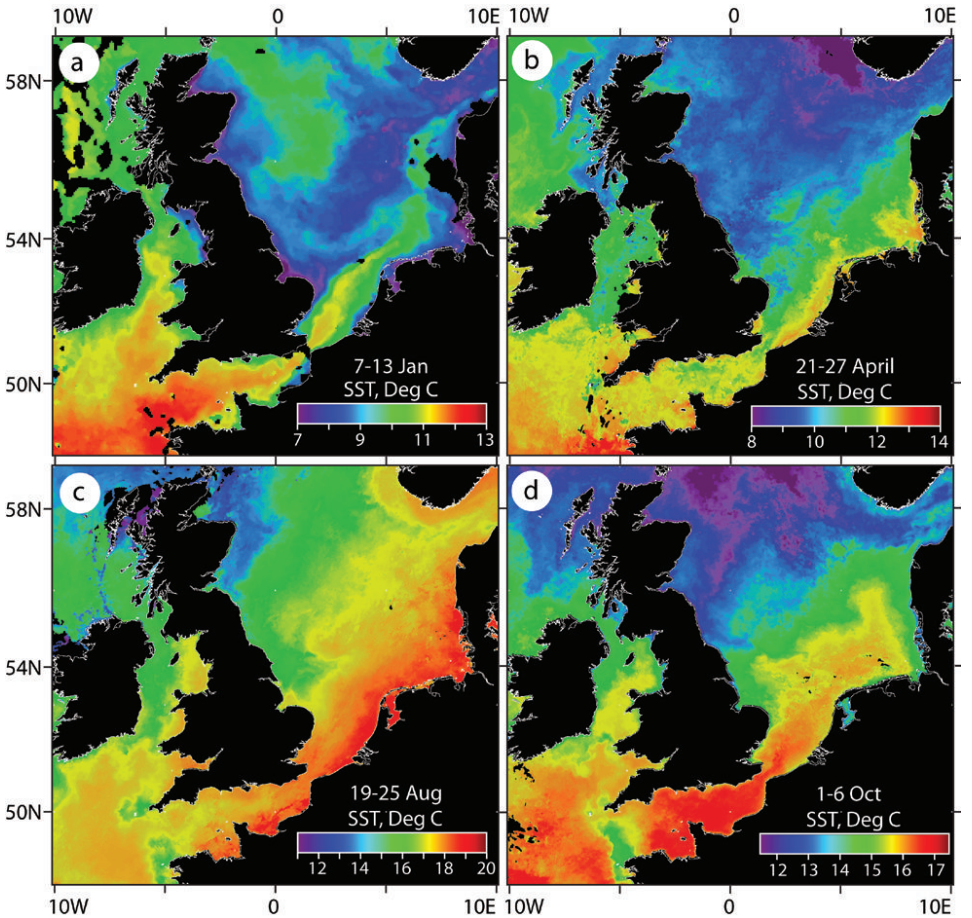


Figure 13.8. Typical, weekly composite sea surface temperature distributions in shelf seas around the U.K. at four times of the year. The figures represent 7-day median values for: (a) January 7–13, 2007; (b) April 21–27, 2007; (c) August 19–25, 2007; (d) September 30–October 6, 2007. Note that the temperature scale is different for each image, selected to emphasize the different spatial temperature structures in each rather than to reveal seasonal variations in absolute temperature (figure created by the author using weekly, composite image data produced by NEODAAS—*http://www.neodaas.ac.uk/*).

For example, in the winter image of the U.K. shelf seas shown in Figure 13.8a, the water along the coastal margins of the English Channel and southern North Sea are cooler because the sea is shallower than farther offshore. At this time of year the sea is losing heat through the surface. If we assume that heat flux per unit surface area (from the sea to the atmosphere at this tine of year) is spatially uniform, the rate of temperature reduction will be greater where the water column is shallower, because of the reduced thermal capacity of the smaller mass of water in the water column. This is the best explanation for cooler coastal margins.

It is tempting at first to interpret the temperature pattern in winter entirely in terms of heat transport by residual flow from southwest to northeast through the region, using the analogy of steady-state temperature distribution of the flow through the pipes of a heating system. This might seem to account for the two warm plumes, one through the English Channel and Straits of Dover, and the other through the Irish Sea, drawing heat from the warm pool of ocean water entering the Celtic Sea in the southwest and carrying it to where it is lost at the cold coastal margins in the Irish and North Seas. However, such an analogy is misleading, as the following discussion of the annual cycle of temperature changes is intended to show.

Seasonal variation of SST

Shelf seas rarely exhibit a steady-state balance between advection and heat loss. Typically the overall temperature is either reducing in fall and winter or increasing in spring and summer. Far from a steady state, temperatures throughout the whole of Figure 13.8a have reduced by about 5°C since the previous October, because heat advection by residual flow is insufficient to supply heat flux to the atmosphere through the surface. Temperature patterns detected by "snapshots" of infrared images during winter depend mainly on local water depth. Where it is shallow the water column cools down more quickly than where it is deep. This is confirmed by the strong similarity between the spatial distribution of temperature in Figure 13.8a and the bathymetry of the region shown in Figure 13.9. It is therefore important to learn to "read the image" more in terms of surface heat loss and thermal capacity of the water column than horizontal heat transfer processes.

In April (Figure 13.8b) when solar heating has started to contribute a net input of heat into the water column, the SST of shallow areas has increased more than that of deeper regions. In the North Sea the shallowest regions, off the Belgian and Dutch coast, in the Thames estuary and over the banks northeast of Norfolk and over Dogger Bank, have already become warmer than deeper regions. The same has happened along the English and Welsh coasts of the Irish Sea, while in most other locations there is little difference in temperature between shallow and deeper regions. Since heating is expected to be higher in the south where the solar elevation angle is greater, there is also a marked north–south gradient of temperature, so that overall the temperature patterns in April are completely different from those in January.

By late summer (Figure 13.8c) there is a tendency for temperature distribution to be the reverse of the winter pattern, especially in the southern North Sea and the eastern English Channel where the water column remains unstratified even in summer. Thus, shallower regions have elevated temperatures compared with deeper zones. However, where a seasonal thermocline develops the dependence of temperature on bathymetric depth is no longer applicable since in this case heat entering the water column through the surface is distributed only through the upper layer. This leads to rather more complex temperature patterns according to whether a region is stratified or not (as discussed below in relation to Figure 13.11). There is one other element of SST distribution in Figure 13.8c which appears to

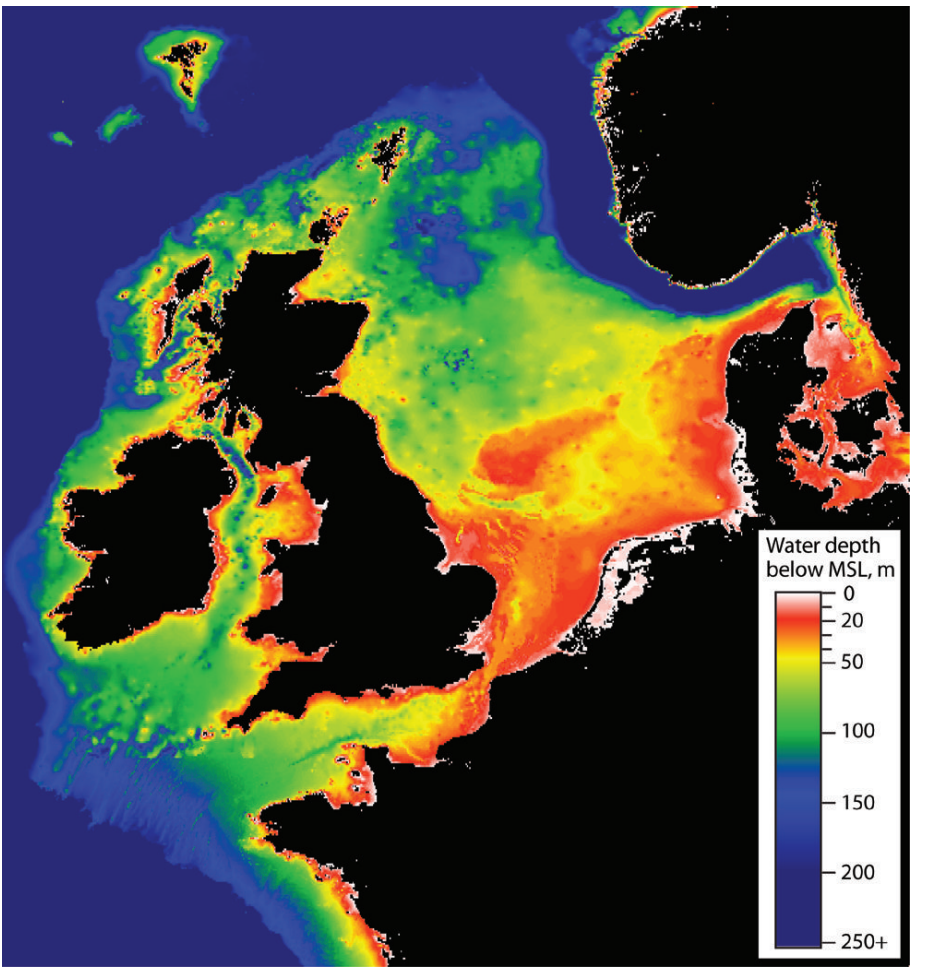


Figure 13.9. Bathymetry of northwest European shelf seas (image created by the author using the U.S. National Geophysical Data Center 2-arcminute digital global topography dataset ETOPO2v2c downloaded from *http://www.ngdc.noaa.gov/mgg/global/etopo2.html*).

contradict the principle of interpreting SST in shelf seas purely in terms of surface heat exchange. This is the cool plume reaching southward down the northeast coast of Scotland. This region is shallower than 100 km offshore and so might be expected to become warmer, instead of cooler, than SST in the deeper part of the North Sea. The most likely explanation is that in this case the tidally driven, counterclockwise residual flow in the North Sea has a significant effect on SST, transporting cooler water down the Scottish coast and overriding local surface heat exchange. Another possibility is that there is some summer stratification in this region, but the prevailing southwest winds are producing offshore Ekman transport which induces upwelling of cooler water adjacent to the coast.

During fall, surface heat exchange is expected to switch from heat input into the water column to a heat loss. In Figure 13.8d, corresponding to early October, this has started to happen, since overall the temperatures are lower than in August. Moreover the shallowest waters close to the coast already have SSTs lower than offshore. It is already fairly clear that as all traces of stratification become broken down and the sea cools further, the dominant winter pattern shown in Figure 13.8a will be re-established.

It needs to be said that—although weekly SST maps chosen to represent different seasons in Figure 13.8 provide a fairly coherent story of the annual temperature cycle—if all 52 weekly composites from 2007 are examined some of them will be found to show anomalous departures from this pattern. This is explainable in terms of the significant impact that weather conditions have on surface heat flux. Thus a period of cold winds and persistent cloud cover can result in net heat losses through the surface in a particular region in summer when positive heat input would normally be expected. Thus local weather can impose its own geographical patterns on SST. This week-to-week variability in SST distribution in shelf seas is consistent with the hypothesis that surface heat exchange and vertical mixing are more important factors in determining SST, at least over the short term, than horizontal advection and diffusion.

Shelf sea tidal-mixing fronts

It was mentioned above that, during summer, seasonal stratification will tend to develop in shelf seas. This leads to the development of a special type of thermal front which is sometimes very clearly defined in satellite infrared thermal images. These shelf sea tidal-mixing fronts deserve special attention here because remote sensing has played an important part in the growth of scientific knowledge about them. They first became the subject of serious scientific analysis in the 1970s (Simpson and Hunter, 1974), at the same time as the first thermal imagery became available from satellites, and this topic provided one of the early examples of how satellite data could make an essential and explicit contribution to advances in oceanographic science (Simpson and Bowers, 1979).

The concept of a shelf sea tidal-mixing front is essentially simple. During the summer months, solar heating provides the buoyancy required to stabilize stratification of the water column into an upper and lower layer. As in the open ocean the development of a seasonal thermocline may be delayed by strong winds that supply energy for turbulent mixing from the surface, which helps to mix heat down through the water column. Once the thermocline is established, further wind stirring tends to mix only the upper layer, steepening the density gradient at the thermocline and further stabilizing it as long as solar heating continues to supply buoyancy flux in the upper layer. However, in the case of shelf seas there is an important additional source of stirring, that provided by the tides.

Tidal currents over the shelf are barotropic and so create a shear layer at the sea bed which acts as a source of energy for turbulent mixing from the bottom up. Tidal currents are regular and the local, mean depth-averaged amplitude, u, of the domi-

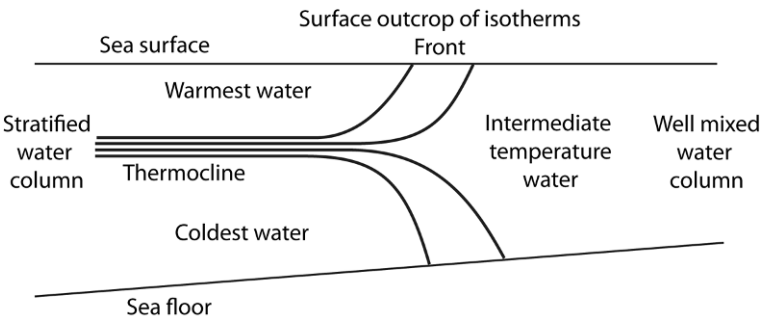


Figure 13.10. Schematic cross-section of isotherms through a shelf sea tidal-mixing front.

nant tidal harmonic current (semidiurnal in U.K. shelf seas) is spatially well defined by regional tidal dynamics. In some places tidal mixing is sufficiently strong to prevent stratification becoming established despite the buoyancy input from surface heating. The magnitude of the tidal current required to prevent a seasonal thermocline from developing is less where the bathymetric depth is smaller. In fact early research into this phenomenon established that the relevant parameter is the ratio h/u^3. Consequently there are places in shelf seas where h/u^3 has a critical value. On one side it is large enough for stratification to be stably sustained, whereas on the other side it is sufficiently small for the water column to remain mixed from top to bottom. In these regions fronts occur that have a cross-section rather like that shown schematically in Figure 13.10.

At the transition between stratified and well-mixed regions of a shelf sea, some of the isotherms in the thermocline outcrop at the surface and the others bend down to the sea bed. Once the stabilizing effect of the steep thermocline is lost, the whole water column rapidly becomes mixed. Thus even if the horizontal gradient of h/u^3 is very gradual, there is still an abrupt transition and the temperature gradient at the surface is quite steep. This is what gives these tidal-mixing fronts a clear remote-sensing signature in satellite-retrieved SST data. The warmer side of such fronts represents a stratified region for which SST characterizes upper-layer temperature. The cooler side corresponds to a well-mixed water column and the measured SST represents the temperature of the whole water column. Although cooler than SST on the stratified side, the well-mixed water column temperature is still greater than the temperature of the bottom layer on the stratified side. The other important thing to note is that the location of such fronts is mostly set by the variables h and u rather than the amount of solar heating or the random variability of wind and weather conditions. Thus, as long as there is enough buoyancy input to stabilize stratification, we expect the front to form in a predictable geographic location. This is confirmed by reference to satellite data such as the examples in Figure 13.11 which contains four SST weekly composite maps of the seas around the British Isles at approximately monthly intervals during the summer of 2007.

The first (Figure 13.11a) is from mid-June by which time tidal fronts have become established. The arrows point to the several different fronts around the Irish Sea for which the warmer side is stratified and the cooler side is well mixed.

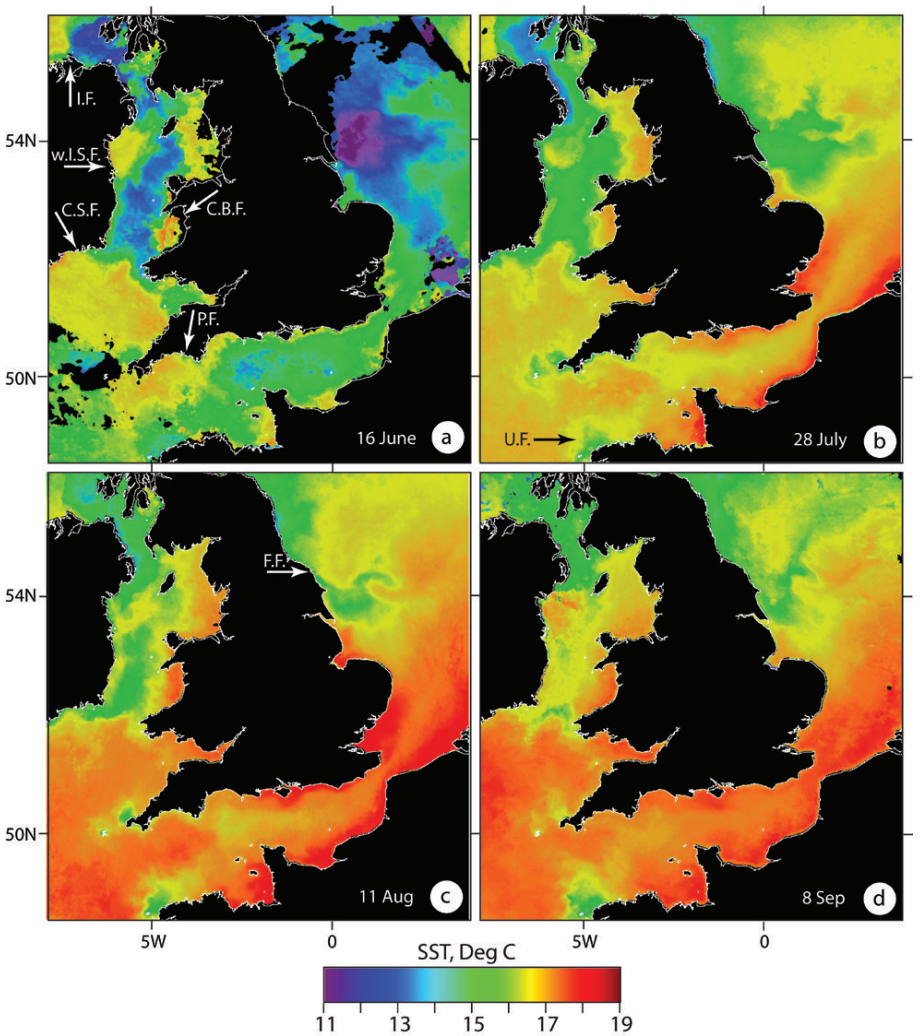


Figure 13.11. Seven-day median composite SST distributions derived from the AVHRR showing the formation of tidal mixing/stratification fronts in U.K. shelf seas during summer of 2007: (a) June 10–16, 2007; (b) July 22–28, 2007; (c) August 5–11, 2007; (d) September 2–8, 2007. The arrows point to particular fronts labeled: I.F. = Islay Front, w.I.S.F. = western Irish Sea Front, C.B.F. = Cardigan Bay Front, C.S.F. = Celtic Sea Front, P.F. = Plymouth Front, U.F. = Ushant Front, F.F. = Flamborough Front.

Note that there are fronts at both entrances to the Irish Sea, the Islay front outside the north entrance and the Celtic Sea front to the south. Although both entrances have fairly deep channels, it is the acceleration of tidal currents which causes the h/u^3 parameter to decrease below the critical level where stratification is no longer sustainable. Inside the Irish Sea there is another interesting frontal region called the

western Irish Sea front southwest of the Isle of Man. This region supports stratification because tidal currents are much reduced here where the semidiurnal co-tidal chart shows there is an anti-node of the tidal standing wave.

In Liverpool Bay, although there appears to be a front, it is not the same as normal tidal-mixing fronts. The stratification that occurs here follows a semidiurnal cycle and is driven by tidal straining of horizontal density gradients created by freshwater inputs from the rivers (Sharples and Simpson, 1995). Stratification can persist for several tidal cycles following a neap tide. This is an example of a region of freshwater influence (ROFI) which has characteristic mixing processes and density-driven residual circulation (Simpson, 1997). ROFIs do not necessarily have signatures in satellite SST images and consideration of the h/u^3 parameter is not applicable. In this case the water is shallow enough to be comparable with the depth of the upper mixed layer in stratified regions and so it heats up to the same extent and gives the appearance of a thermal front. The same may be happening in Cardigan Bay, but here tidal currents are weaker, bypassed by tidal flow up and down the main channel, and the estimates of h/u^3 based on bathymetry and modeled tidal streams show that conditions are suitable for stratification to develop.

Figure 13.12 sketches estimates of the contour where $\log(h/u^3) = 2.0$, based on the original work by Simpson (1981). There has been some debate about which controlling mechanism should be assumed when estimating the critical value of h/u^3 that determines where fronts occur (see, e.g., Simpson and Hunter, 1974; Pingree and Griffiths, 1978; Simpson and Bowers, 1979; Loder and Greenberg, 1986; Bowers and Simpson, 1987; Simpson and Sharples, 1994). An energy-based criterion points to a critical value of $\log(h/u^3)$ above 2.5, whereas a boundary layer–based approach implies that the rotation direction of the tidal stream ellipse may be important and suggests a lower critical value. Theoretical models also suggest that the front should move between spring and neap phases of the tide in a 14-day cycle, since neap tidal streams have a speed that is only about 50% that of spring tides (Sharples and Simpson, 1996). Some shifting has been detected, although by only a few kilometers, much less than predicted. Analysis in other parts of the world also shows some variability (Loder and Greenberg, 1986), but generally the 2.0 contour offers a satisfactory pointer to where fronts should be found (as illustrated by the agreement between Figures 13.11 and 13.12).

The fronts shown in Figure 13.11 generally maintain the same location through the whole summer, while overall SST rises to a maximum in September. There is another front at the western entrance to the English Channel, stretching from Plymouth in a sinuous shape and bending round the Brittany peninsula, in agreement with theoretical prediction, and this is referred to as the Ushant front (see Figure 13.11b). A front also develops in the North Sea at about 54°N, off Flamborough Head. North of here the sea stratifies in summer whereas southwards the North Sea remains well mixed.

As discussed in relation to Figure 13.9, SST is high in the southern North Sea because of the shallow depth. However, there is a characteristic cool region stretching 50 km to 100 km south of the Flamborough front. This might give the erroneous impression of being a cool plume emanating from the coast, but the reason this water

Figure 13.12. Map of modeled contour of the parameter $\log(h/u^3)$, where h is measured in meters; and u in meters per second.

remains cooler than its surroundings is that it continues to mix with the source of colder water in the lower part of the stratified water column to the north. The same is the case in the Irish Sea, where unstratified waters of the main channel remain cooler than coastal waters because they remain in contact with the colder lower layer of the Celtic Sea. This is important because the lower layer is likely to be richer in nutrients. Indeed, shelf sea fronts are zones where there is a supply of nutrient-rich waters available, leading to enhancement of primary production in their vicinity. Figure 13.13 sketches locations within the frontal structure where enhanced production is typically found. Figure 13.14 provides an example where the Celtic Sea front has a distinct signature in the chlorophyll image from SeaWiFS. The ecological impor-

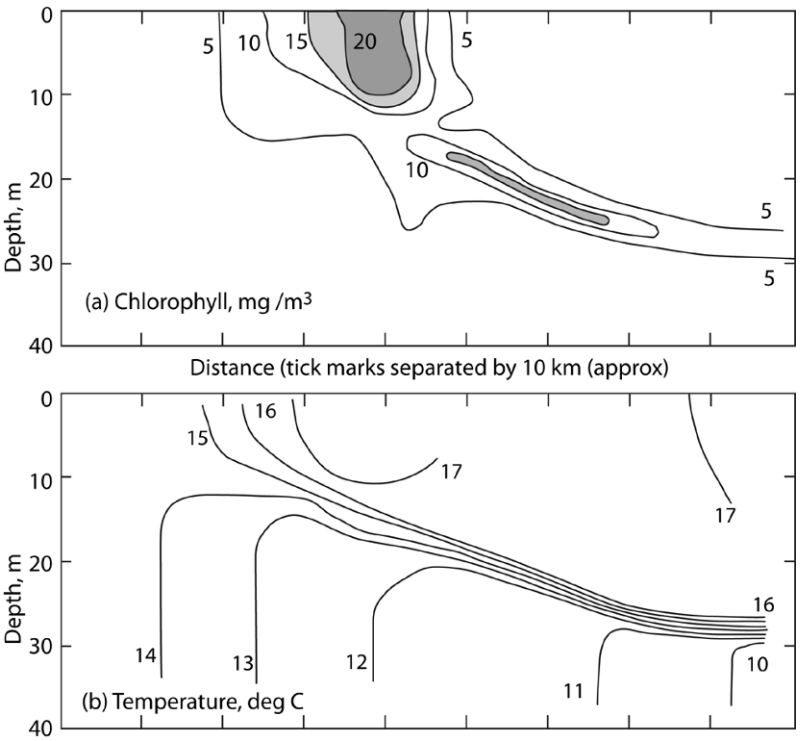


Figure 13.13. Sections showing chlorophyll and temperature distribution through the Ushant front in July 1975 (after Pingree *et al.*, 1975).

tance of shelf sea fronts is a research topic of some importance (Moore *et al.*, 2003), to which ocean color remote sensing has a contribution to make. For example, even before regular ocean color data became available, satellite-derived SST data were used to provide estimates of nitrate flux at the Ushant front (Morin *et al.*, 1993).

Of course satellite data alone are insufficient for such studies, because cloud cover prevents the evolution of fronts from being fully traced, wind stirring can temporarily disturb the thermal structure, and diurnal warming layers may mask the frontal signature in very calm conditions. It must be admitted that not all weekly SST composites from the summer of 2007 present such a clear and unambiguous view of tidal-mixing fronts as the ones selected for Figure 13.11. Nonetheless satellite data continue to offer an important extra perspective to field experiments. Following pioneering work in the Irish Sea, satellite data have contributed to the study of similar fronts in strongly tidal shelf seas in other parts of the world, including both SST detection of fronts and the use of ocean color to observe productivity, and even fisheries, associated with them. Examples of regions observed include the Gulf of Maine (Loder and Greenberg, 1986), the Patagonian Shelf (Glorioso, 1987; Bogazzi *et al.*, 2005; Dogliotti *et al.*, 2009), the shelf seas off China (Tang *et al.*, 1998, 2003; Wang *et al.*, 2001; Hu *et al.*, 2003), and off Japan (Yanagi *et al.*, 1995), while some of the more general literature on the remote sensing of fronts (cited in Chapter 4) includes references to shelf tidal-mixing fronts.

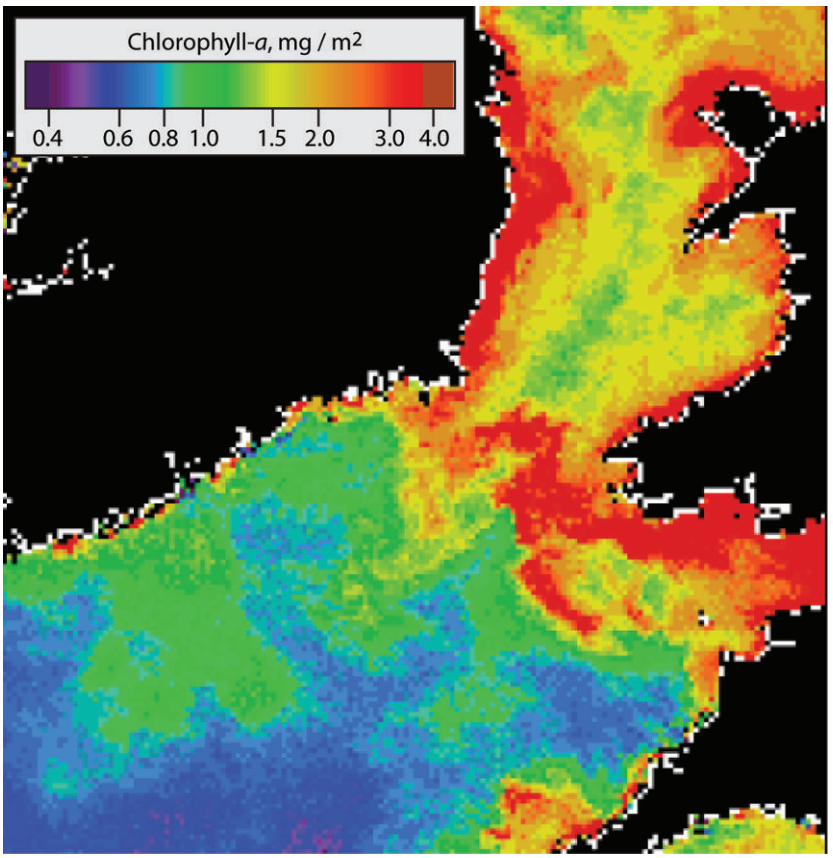


Figure 13.14. SeaWiFS chlorophyll image acquired on July 11, 1999, showing enhanced chlorophyll-*a* concentration along the line of the Celtic Sea tidal front and the western Irish Sea front. Note that the chlorophyll algorithm used is unlikely to be valid in the Bristol Channel (adapted by the author from an image provided by NEODAAS—*http://www.neodaas.ac.uk/*).

13.2.5 Remote sensing of suspended sediments in shelf seas

Riverine inputs and coastal erosion produce considerable quantities of seafloor sediments which tend to spend some time in shelf seas before either becoming compacted into sedimentary layers on the shelf seafloor or eventually migrating over the shelf edge onto the ocean abyssal plain. Strong tidal currents in shelf seas, as well as the energy provided by surface waves in shallower coastal areas, tend to resuspend these sediments. Consequently some areas of shelf seas contain high loads of suspended particulates producing mobile sea bed forms that can create navigational hazards and leading to a highly turbid water column that prevents light penetrating below a few meters, restricting primary production, and impacting strongly on the local shelf sea ecosystem.

Using visible waveband reflectance data

Satellite remote sensing is capable of observing suspended sediments and can provide valuable insights for understanding their distribution and transport paths in shelf seas. The methodology is to use the subsurface reflection of visible light as a measure of the concentration of particulates in the upper part of the water column. This is an important role for dedicated ocean color sensors such as SeaWiFS, MERIS, and MODIS, although in the 1980s and 1990s, when there were no satellite ocean color sensors in operation, the broadband (580–680 nm) visible channel on the AVHRR meteorological sensor was used to produce maps of water turbidity and suspended sediment for shelf seas in northwest European shelf seas (Spitzer *et al.*, 1990; Weeks *et al.*, 1993; Bowers *et al.*, 1998, 2002) as well as in many other parts of the world (e.g., Rucker *et al.*, 1990; Gupta and Krishnan, 1994; Li *et al.*, 1998; Froidefond *et al.*, 1999). There was also at this time considerable work done on the theoretical basis for interpreting single-channel reflectance in the green-to-red part of the spectrum in terms of water turbidity and suspended sediment concentration (e.g., Simpson and Brown, 1987; Prangsma and Roozekrans, 1989; Stumpf and Pennock, 1989; Weeks and Simpson, 1991).

Since the launch of SeaWiFS in 1997, attention has turned to what are spectrally richer ocean color sensors for monitoring and measuring the distribution of suspended sediments. Given a range of narrow-wavelength bands to choose from, which are the most appropriate for measuring suspended sediment? Water-leaving reflectance spectra (see, e.g., figure 6.25 in *MTOFS*—Robinson, 2004) differ considerably for different water types. The most marked effect of increased suspended sediment is to increase reflected light in the green-to-red part of the spectrum at wavelengths greater than 500 nm, extending to 600 nm and beyond in extremely turbid waters. For this reason, a waveband at around 550 nm or 560 nm is normally used as the primary means of representing total suspended sediment (TSS) which includes both organic and mineral fractions of particulate material.

By itself this gives a good qualitative picture of the spatial distribution of sediments in suspension. However, it cannot distinguish between mineral-suspended sediment (MSS), phytoplankton, and organic detritus. Neither can the single-channel, normalized water-leaving reflectance be converted into a quantitative measure of the gravimetric concentration of suspended sediment, except on the basis of calibration compared with *in situ* sampling, and making assumptions about the type of material and particle size distribution being invariant. Furthermore, as with all ocean color remote sensing, satellite measurements can tell us only about particulates in approximately the upper third of the photic zone, which means high-turbidity water near the sea bed is not likely to be detected. Finally, where water turbidity is high, the atmospheric correction procedure is likely to be compromised by misinterpretation of high-water reflectance in near-infrared channels as excess atmospheric scattering, unless special care is taken to adopt special "bright pixel" adjustment techniques (Moore *et al.*, 1999; Lavender *et al.*, 2005; Gao *et al.*, 2007).

Example of a winter turbidity image

Despite these limitations, a great deal can be learned from satellite images such as that shown in Figure 13.15. This level 2 image from MODIS (Aqua) consists of atmospherically corrected, normalized water-leaving radiance from the 551 nm channel, acquired in early February over U.K. shelf seas. At this time of year it can be assumed that there are few phytoplankton present and that the image can tell a story about the patterns of suspended sediments and their movement through the shelf seas. First can be noted the high levels of green light reflected from parts of this scene. It is worth pointing out that in the wider image from which this view was extracted, there is a reduction of radiance at the shelf edge (not shown) from around 1.0 over the shelf to less than 0.2 $mW\,cm^{-2}\,\mu m^{-1}\,sr^{-1}$ over the deep ocean. Thus

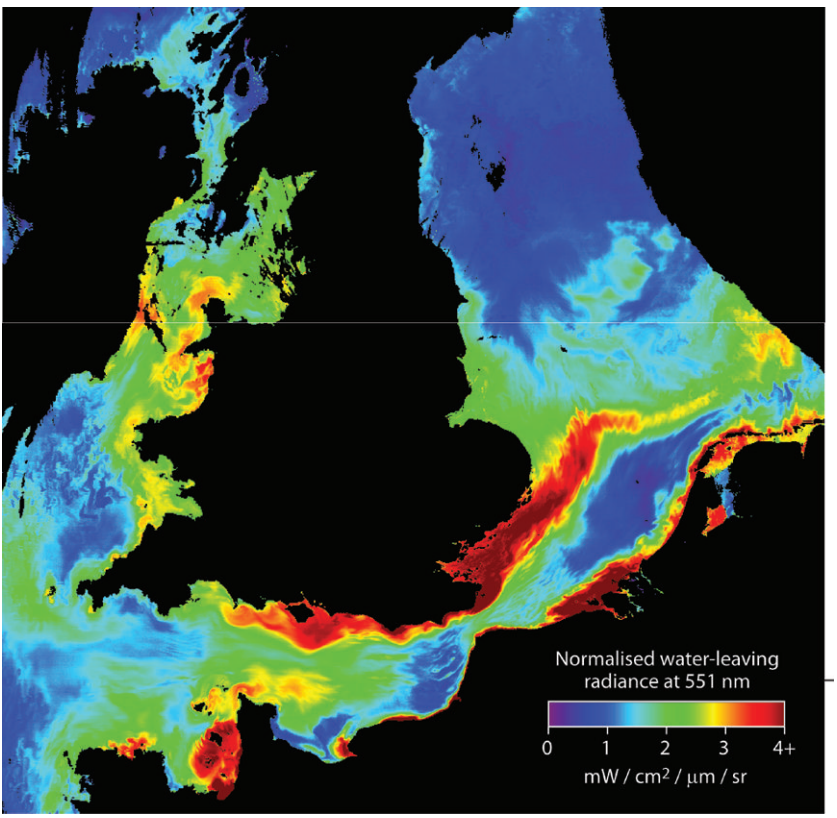


Figure 13.15. Level 2, MODIS (Aqua), normalized water-leaving radiance at 551 nm from a mainly cloud-free overpass of U.K. shelf seas on February 11, 2008. This waveband serves as a qualitative proxy for total suspended sediment in the upper part of the water column. Land and pixels identified as cloud are shown in black (image created using Level 2 data downloaded from the NASA Ocean Color web site—*http://oceancolor.gsfc.nasa.gov/*).

there is an order of magnitude difference between the brightness of pixels on and off the shelf.

The image clearly shows zones of apparently high loads of sediment in the upper water column. These can be loosely interpreted from general knowledge. For example, along the south coast of England there is coastal erosion, especially around the Isle of Wight, while the rivers Thames and Rhine are also strong sources feeding particulates into the North Sea. The Wadden Sea enclosed by the Frisian Islands off the northwest coast of the Netherlands and the Gulf of St. Malo off northwest France are regions where mobile sediments accumulate. However, a detailed study of a time series of images would yield further insights, showing how patterns change with tidal amplitudes and with wind. If the type and size distribution of sea bed sediments are known, this can also aid interpretation of the image.

Such images show plumes of high sediment load apparently transporting and dispersing suspended material away from source regions. A dominant signature in Figure 13.15 is the plume from the Thames that heads northeast along the Essex coast and then tracks out into the center of the North Sea. Its route corresponds to shallow regions with sandbanks and changing bedforms. On this image, suspended particulates from the Rhine appear to be dispersed by tidal oscillation in both directions along the coasts of Belgium and the Netherlands.

Another striking feature of this image is the strong spatial heterogeneity it reveals, with very sharp gradients that amount to turbidity fronts, separating the plumes of high sediment load from closely adjacent clearer water. These turbidity fronts are aligned parallel to the dominant tidal stream directions, although they bulge out off the prominent headlands along the south coast of England where there are known to be residual circulation eddies produced by tidal rectification of vorticity in oscillatory flow. Thus we can consider this image to serve as a tracer, not only of suspended sediments, but also of some of the fluid-dynamical processes taking place. In particular the sharpness of turbidity fronts implies that they are maintained by a density front, and may represent shear lines in the tidal current field. Thus such images have the capacity to provide a signature for ROFIs (mentioned in Section 13.2.4) that generally do not show up in thermal images. It is interesting to compare Figure 13.15 with Figure 13.16 which is the MODIS SST field measured at the same time (demonstrating one of the benefits of measuring color and temperature on the same sensor or platform). This shows a rather different set of patterns, consistent with interpretation of the winter thermal behavior made in Section 13.2.4 and shown in Figure 13.8a. However, careful comparison shows that most turbidity frontal lines do align with isotherms in the temperature field, but in the latter they give little impression of being the distinct boundaries that appear in the visible waveband image.

Another feature of the 551 nm water-leaving radiance field is streakiness at smaller lengthscales of a few kilometers which can be seen on closer inspection of Figure 13.15. Figure 13.17 attempts to enhance this in a gray-tone image of the same data. The first impression is that the streaks line up with tidal streams, but east of the Dover Strait there is also evidence that in fact they are related to more prominent bed forms. The same is found over the Norfolk banks off the coast of East Anglia.

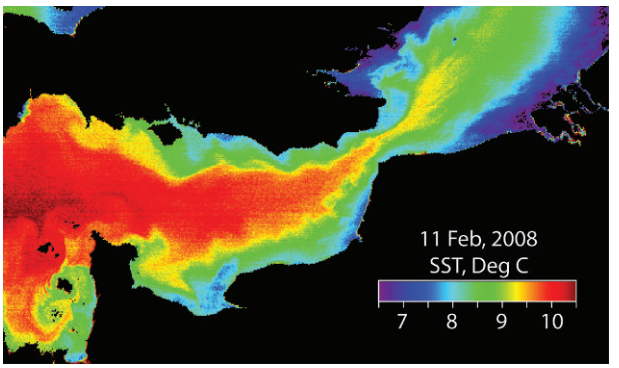


Figure 13.16. MODIS SST image from the same overpass as Figure 13.15 on February 11, 2008.

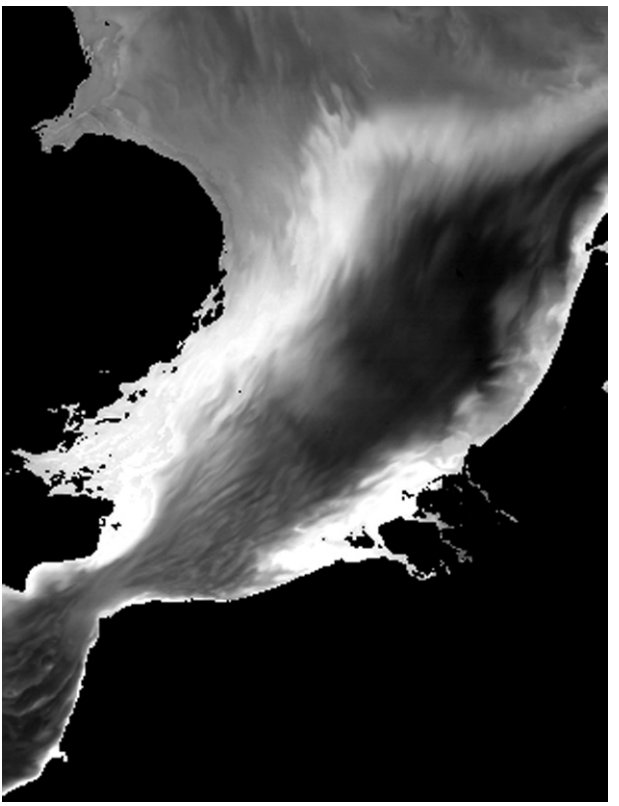


Figure 13.17. Part of Figure 13.15 enlarged as a gray-tone image to reveal the fine-resolution streaks aligned with the current or the bathymetry.

Thus we may be seeing the differentiation of suspended sediment concentration, either simply according to bathymetric depth, or more subtly in relation to differentiation of flood and ebb channels that occurs in tidal flows interacting with complex sandbanks and sand waves. A more definitive analysis and interpretation than this superficial review would require the study of a sequence of images acquired at different phases of the tide.

Summer turbidity distribution

The analysis of summer, water-leaving radiance images as indicators of suspended particulates introduces the complication that light scattering may come from phytoplankton (i.e., organic rather than mineral particulates). Figure 13.18 is such an image. To overcome the limitations caused by extensive cloud cover, a level 3, monthly composite, MODIS dataset at a resolution of 4 km has been used. Although it lacks the crisp detail of single-overpass 1 km images, and despite the persistent cloud patch over the Thames estuary, it shows a broadly similar pattern of high turbidity along the English south coast and stretching northeast from East Anglia.

What does differ from the winter image is the strong contrast between the high reflection of light in coastal areas and the much lower reflection over the deeper shelf. In fact, the date of this example was chosen to coincide with the thermal image of Figure 13.11c. When compared with thermal images, it is striking how closely the boundary between low reflectance (shaded blue on Figure 13.18) and higher reflectance (shaded green) follows the line of tidal-mixing fronts in thermal data. The Islay, Celtic Sea, Flamborough, Plymouth, and Ushant fronts are all clearly delineated at the locations shown on Figure 13.11a. The explanation is very simple: In stratified water any particulates resuspended by tidal flow over the sea floor cannot penetrate above the thermocline. Thus the upper layer remains clear of suspended material and so reflects less light. The western Irish Sea front is also detectable, although not so clearly.

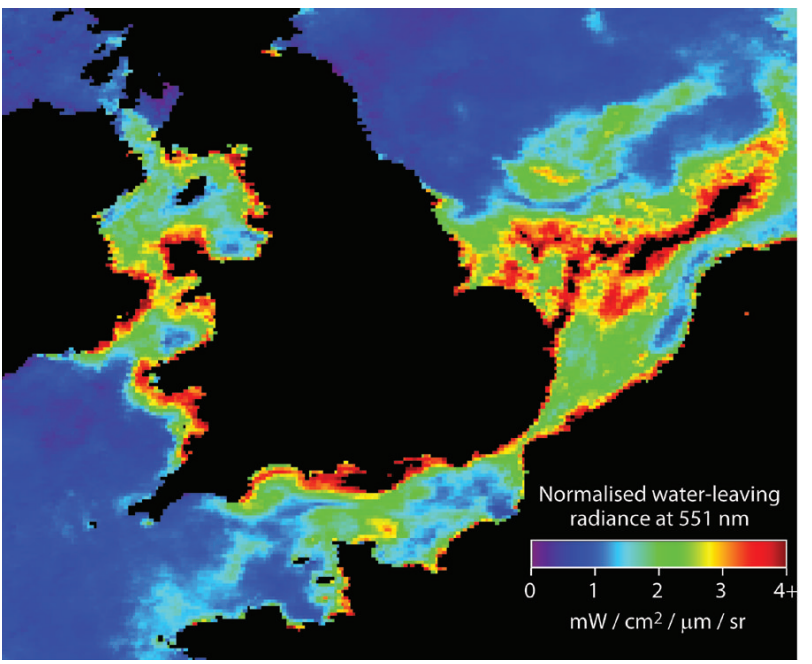


Figure 13.18. Monthly composite of normalized water-leaving radiance at 551 nm from MODIS (Aqua) data for August 2007.

What was also evident in the wider area around the U.K., from which Figure 13.18 was selected, is that there are patches of high reflectance found elsewhere in stratified shelf waters, and over the deep ocean, which correspond to phytoplankton blooms (see also Figure 13.4). If the blooms are coccolithophores the 551 nm reflectance signature is very strong and comparable with coastal patches of high mineral-suspended sediment content. This should remind us that, while it may be reasonable to interpret single-band winter images qualitatively as suspended sediments, to do so in the rest of the year risks confusion from the conflicting scattering and absorbing effects of organic material.

Retrieval of quantitative estimates of suspended sediment

Even in a sea containing no organic content, a quantitative estimate of suspended sediment load requires knowledge of the optical scattering properties of sediment, which can vary with its mineralogy and its size distribution. In order to use satellite ocean color data to retrieve quantitative estimates of suspended sediment concentration requires the more complex approach applicable to Case 2 optical conditions, in which the full spectral band set is used. Anticipating further discussion of these issues in the next section, Figure 13.19 shows an example of TSM distribution in U.K. shelf seas retrieved from the MERIS sensor, and a corresponding map of z_{90}, the depth of light penetration (from above) from which 90% of water-leaving radiance originates (see Section 12.3.2). These products are generated using a retrieval algorithm based

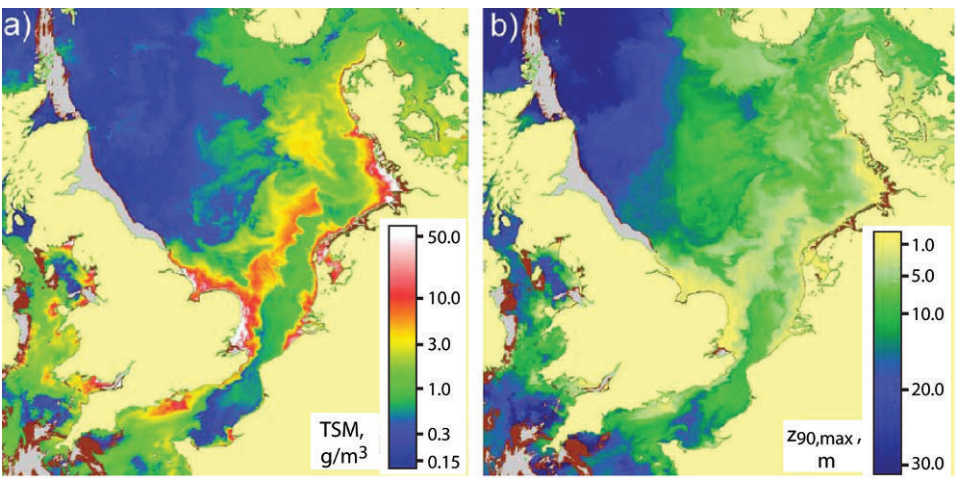


Figure 13.19. (a) Total suspended matter (TSM) derived from a MERIS scene over the North Sea, acquired on March 27, 2007 using the Case-2 R-BEAM processor. (b) The corresponding signal depth, $z_{90,\max}$. Patches of brown or gray correspond to where the algorithm does not produce a result (e.g., because of cloud), as published in IOCCG (2008) (data products and images produced by Roland Doerffer, using MERIS data provided by the European Space Agency, and reproduced with his permission).

on an artificial neural net (ANN) approach (Doerffer and Schiller, 2007). The ANN is trained with a large simulated dataset of bidirectional radiance reflectances derived from a bio-optical model that uses inherent optical properties based on empirical measurements of absorption and scattering coefficients. These came from a variety of coastal waters, including those of northwest European shelf seas, with concentrations of SPM up to 50 mg/L.

Using synthetic aperture radar

Finally, when considering the role of remote-sensing methods in the study of sediment transport in strongly tidal shelf seas, it is important not to forget the potential for using synthetic aperture radar. Tidal flow over bathymetric features such as sandbanks can sometimes be revealed in SAR images in shallow seas (typically less than 20–30 m deep) because tidal currents forced over and around such banks produce regions of convergence and divergence in the surface current field. This modulates short wind waves on the surface that determine the strength of radar backscattering, which is represented as brightness on the SAR image. Figure 13.20 provides a striking example of the imaging of complex underwater topography,

Figure 13.20. Example of a SAR image showing detailed shallow-water bathymetry. This is a full ERS SAR image (100×100 km) at about 33°N, 121°E, over the Chinese coast 200 km north of Shanghai, on July 8, 1995 (ESA image obtained online from Alpers *et al.*, 1999).

showing the waters off the Chinese coast near Shanghai. This imaging mechanism is explained further in section 10.11.2 of *MTOFS* (Robinson, 2004). The scientific basis of a practical bathymetric measuring system has been developed in detail (Vogelzang, 1997; Vogelzang *et al.*, 1997) and its commercial application has been demonstrated for monitoring the complex and ever-changing channels in the Wadden Sea along the north coast of the Netherlands and Germany.

By estimating the two-dimensional bathymetry of complex bed forms it is possible, in principle, to detect their movement from a sequence of SAR images acquired over weeks, months, or years. Thus, whereas ocean color methods tell us about the finer sediments that are suspended in the upper water column, SAR may be able to tell us about the result of sediment transport as it affects the position of mobile bed features. However, application of SAR data for monitoring bed forms must be approached with caution. Because of the complexity of procedures for inverting SAR scenes into bathymetric maps, requiring precise knowledge of such things as tidal currents and wind conditions, there remain large uncertainties in the results except under the most favorable circumstances such as found in the Wadden Sea. Nonetheless, the method has wider applicability as a qualitative monitoring tool for regions of navigational importance that are prone to shifting sandbanks. SAR images can be surveyed routinely and comparison made between the images of bed forms acquired under similar conditions of wind and tide. If these show significant differences, a ship-based bathymetric survey can be initiated to make precise measurements from which to determine whether the changes are serious enough to justify the issue of new navigation charts and relocation of buoys marking the safest navigation channel.

13.2.6 Monitoring ecosystems and water quality

Shelf seas receive flows from rivers which also bring large quantities of dissolved organic material from land drainage and sewage, as well as nutrients which support enhanced primary production that can lead to eutrophication. The monitoring of water quality is therefore essential for most shelf seas which over the centuries have attracted large human populations as well as commerce and industry to adjacent coastal countries. The capacity of ocean color sensors to measure the distribution of chlorophyll concentration in near-surface water holds out the promise that we can use this methodology to monitor phytoplankton in shelf seas. Since phytoplankton are the primary producers on which shelf sea ecosystems are based, it is reasonable to expect that satellite ocean color measurements should provide the basis for operational monitoring, and possibly forecasting, of shelf sea ecosystems and water quality.

Unfortunately, despite the importance of this application, the pullthrough of ocean color remote sensing into improvements to shelf sea ecosystem models has been slow to emerge in comparison with the benefits of other satellite oceanography techniques. A review of the reasons for this (Robinson *et al.*, 2008) points to the complexity of optical processes and their coupling to biological, chemical, and physical processes in shallow-sea ecosystems. Our capacity to retrieve quantitative

measures of useful ocean properties from ocean color in Case 2 waters has not yet achieved the degree of accuracy and reliability that are needed for operational applications. There is still some way to go with research to improve our knowledge of the inherent optical properties of shelf seas. The important research issues are discussed below under the subheading "The Case 2 challenge".

The consequence of the present immaturity of ocean color methodology in Case 2 waters is that maps of chlorophyll concentration derived from ocean color sensors in shelf seas are unreliable in comparison with the confidence we can place in oceanographic data retrieved from other remote-sensing methods, such as maps of SST or wind speed and direction. The careless use of shelf sea regional data abstracted from global chlorophyll datasets based on Case 1 retrieval algorithms can give a very erroneous picture (e.g., suggesting high chlorophyll concentrations in winter that are actually caused by high concentrations of mineral-suspended sediments, or by land-derived CDOM). Even explicitly targeted retrieval algorithms adapted for Case 2 waters may have uncertainties in excess of 100%. Nonetheless it would be equally foolish to ignore the potential information content of ocean color datasets over shelf seas. Images such as Figures 13.4 have a lot to tell us qualitatively. The medium-term goal for developing routine operational use of ocean color satellite data is somehow to couple satellite data with ecosystem and optical models (as discussed in Section 14.3). However, given the present state of the art it would not work to simply assimilate satellite-derived chlorophyll data into ecosystem models in the way that SST can be assimilated into water circulation models; the large uncertainties in the data would seriously weaken their ability to influence the model. Until that can be remedied by improved quantitative retrieval methods, how can such data assist those monitoring water quality in shelf seas? This is addressed below under the subheading "Monitoring algal blooms".

The Case 2 challenge

As mentioned in Chapter 2 and elaborated in section 6.3.2 of *MTOFS* (Robinson, 2004), during the early development of ocean color remote-sensing methods it was found helpful to distinguish seawater into two optical categories (Morel and Prieur, 1977). Case 1 refers to waters whose optical properties are dominated by phytoplankton and their associated degradation products only. Case 2 applies to all other situations; that is, those where particulate material independent of the local phytoplankton population, and/or land-derived yellow substance influence ocean color instead of, or in addition to, phytoplankton. While this classification has simplified the retrieval of ocean variables in Case 1 conditions, allowing progress to be made in global and regional applications of ocean color remote sensing (as discussed in Chapter 7), it has not solved the problem of how to make progress in situations where water conditions are, or may be, Case 2. Shelf seas, by virtue of their shallowness allowing the resuspension of bottom sediments and because they are strongly influenced by river inflows, must always be suspected as being Case 2 unless it can be objectively confirmed that they are genuinely Case 1.

The development of reliable algorithms for retrieving oceanographic variables from ocean color under Case 2 conditions has not progressed as rapidly as for Case 1. Various approaches have been attempted (IOCCG, 2000), most of them requiring a well-defined spectral characterization of the inherent optical properties (IOPs) of seawater and knowledge of how these can be related quantitatively to the concentration of color-affecting water constituents through the quantification of specific inherent optical properties (SIOPs). The simplest approach to Case 2 algorithms uses knowledge of local IOPs to select appropriate wavebands for an empirical band ratio algorithm with limited geographical applicability (Darecki *et al.*, 2003). Another broad class of methods seeks to retrieve estimates of IOPs (spectrally dependent absorption and scattering coefficients) from analysis of subsurface reflectance derived pixel by pixel from the ocean color sensor (IOCCG, 2006). Then from a knowledge of regionally appropriate SIOPs, water constituents are deduced from the derived IOPs.

The other main approach is to use forward optical modeling incorporating SIOPs appropriate to the area of interest in order to simulate a large dataset of artificial reflectance spectra defined at sensor wavebands, each corresponding to different combinations of concentrations of phytoplankton, CDOM, and suspended sediments, and different viewing directions and Sun angles. In this approach the Case 2 algorithm must match the observed reflectance spectrum at each pixel to the artificial dataset, in order to find conditions most likely to account for the observed ocean color at the given viewing geometry. Artificial neural network (ANN) methods have been demonstrated to do this quite successfully, leading to Case 2 algorithms which are quite easy to apply (Doerffer and Schiller, 2007). Under well-controlled conditions where there are extensive *in situ* measurements of optical properties of a particular shelf sea, such algorithms have shown promise. However, until there has been extensive validation by *in situ* observations covering a wide variety of shelf sea conditions it would be premature to assume that ANN-based algorithms are more universally applicable beyond the particular seas which furnished the relatively limited set of optical measurements that defined the SIOPs used to create the training dataset. Even when there are comprehensive local observations of optical properties, ''closure'' experiments still show large discrepancies between measured water-leaving radiances and predictions of optical models (Bulgarelli *et al.*, 2003; Chang *et al.*, 2003).

The same problem afflicts the previously mentioned class of Case 2 retrieval methods which still rely on there being an extensive set of SIOPs available for a wide variety of seas. Unfortunately this is not the case. The literature on SIOPs is quite limited and not always reliable in terms of the accuracy of measurement of either optical properties or water-constituent concentrations or both (Berthon *et al.*, 2008). The relatively small pool of available *in situ* optical measurements, resulting in inadequate specification of SIOPs, was identified by a European Science Foundation working group (Robinson *et al.*, 2008) as a primary reason reliable Case 2 retrieval methods have not yet been successfully demonstrated. The complexity of interactions between different optically effective agents in Case 2 waters can make it more important than under Case 1 conditions to specify the SIOPs of the particular type of phytoplankton present, the particular compounds making up CDOM, or the miner-

alogy and size distribution of mineral-suspended solids. Atmospheric correction of ocean color data to estimate water-leaving radiance is also more difficult over shelf seas, not only because of accommodating Case 2 seawater optical conditions but also because of the diversity and spatial heterogeneity of atmospheric aerosols associated with industrial activity in adjacent coastal regions.

The working group identified other issues where further scientific knowledge is needed to support progress in the application of ocean color data to shelf sea ecosystems including

- The capacity to determine phytoplankton groups (or species in some cases) from ocean color, which will only become feasible after the optical properties of these groups are well documented.
- The optical properties of atmospheric aerosols (spectral dependence of scattering and single-scattering albedo) and their vertical structure.
- The modeling of reflectance due to white caps and foam.
- Specifying the IOPs related to gas bubbles in near-surface water.
- The surface accumulation of phytoplankton species and their accompanying products (when ocean color remote sensing tends to resemble land vegetation remote sensing).
- Interpretation of the natural phytoplankton fluorescence signal in terms of chlorophyll concentration or of phytoplankton physiological status, particularly in the presence of high sediment loads.
- The inability of present inversion algorithms to tackle the problem of ambiguities, when very different sets of water constituents can produce almost identical spectral reflectance signatures.

Not all of these are specific to Case 2 waters, but they are all factors that can be important for particular shelf seas

Monitoring algae blooms

Despite the somewhat negative tone of the previous subsection, it is important not to lose sight of the real opportunities that ocean color image data offer for applications to the monitoring and management of shelf sea ecosystems. Although we need to be cautious about using quantitative retrievals of ecosystem variables from ocean color data until they can be fully justified from scientific understanding and validated empirically, qualitative analysis and interpretation of ocean color images can still reveal a lot.

An important issue for those charged with the ecological management of shelf seas is the detection of harmful algal blooms (HABs), classified as such when the algae involved produce toxins that kill fauna or endanger human health. More generally any phytoplankton bloom may be considered as a nuisance when it becomes dense enough to result in anoxia, shading or smothering, and loss of fauna and flora. HABs can have a substantially deleterious economic impact on a region. In the context of mariculture they may cause mass mortality of caged fish and

the shutting down of shellfish farms contaminated by toxins. They affect wild shellfish fisheries, endanger human health, force the closure of recreational beaches, and result in loss of tourism. Changes from expected occurrences of HABs can occur when exotic algae are unwittingly introduced into a region by shipping, or when the pattern of discharge of pollutants or agricultural runoff enhances the nutrient load, and may be a response to climate change. Since it is the sudden appearance of an unexpected bloom that tends to cause the greatest harm, the early detection of HABs and the monitoring of their growth and dispersion has become an important task for marine environment agencies. The more warning that is given of an approaching bloom, the greater the possibility of taking action to mitigate their impact (e.g., by moving fish cages or harvesting shellfish early).

For these reasons there has been strong interest in using satellite ocean color data to monitor HABs (see, e.g., Stumpf *et al.*, 2003; Pitcher and Weeks, 2006). Harmful algal blooms are typically associated with variability lengthscales from hundreds of meters to tens of kilometers, and timescales from days to weeks. Time series of satellite data products with a spatial resolution of 30 m to 1 km at intervals of one to a few days are highly desirable (see chapter 5 in IOCCG, 2000). Although the smallest of these lengthscales cannot be resolved at such regular intervals from satellites, sensors such as MERIS, in its high-resolution (300 m) imaging mode with overpasses every other day, can make a serious contribution to monitoring algal blooms and their evolution (chapter 9 of IOCCG, 2008). This is, of course, subject to cloud-clear conditions, which may rule out the use of satellites as the primary means of early warning of HABs in some shelf seas. Figure 13.21 shows an example of chlorophyll distribution derived from a full-resolution MERIS image dataset, and depicting a harmful algal bloom encroaching the Scottish Orkney Islands.

A sensor such as MERIS, which maintains its fine spectral resolution even in full spatial resolution mode, holds out the promise of being able to distinguish one algal species from another by its spectral signature (Cullen *et al.*, 1997). However, this is rarely possible unless the algal bloom has a distinctive color that contrasts with harmless blooms typically encountered in a region (e.g., a "red tide" in which the red pigmentation is evidence of the harmful nature of the bloom—Kahru and Mitchell, 1998). In most cases ocean color sensors are not able to discriminate between toxic and nontoxic species, although with additional local knowledge it is sometimes possible to identify high-risk blooms from their intensity, their unusual time of occurrence, or spatial structure. In most cases, like the example shown in Figure 13.21, once a bloom has been detected in satellite image data it is necessary to use conventional oceanographic methods to obtain *in situ* samples from which the species of the bloom can be determined. If a harmful bloom is identified, satellite data can continue to be used qualitatively to monitor the movement of what is now classified as a HAB, and thus provide advice on mitigation of potential dangers. Case studies of actual operational scenarios (Durand *et al.*, 2002) provide evidence of how even the qualitative use of ocean color data can make a major contribution to monitoring water quality.

Other types of blooms which produce a thick scum of material at the sea surface

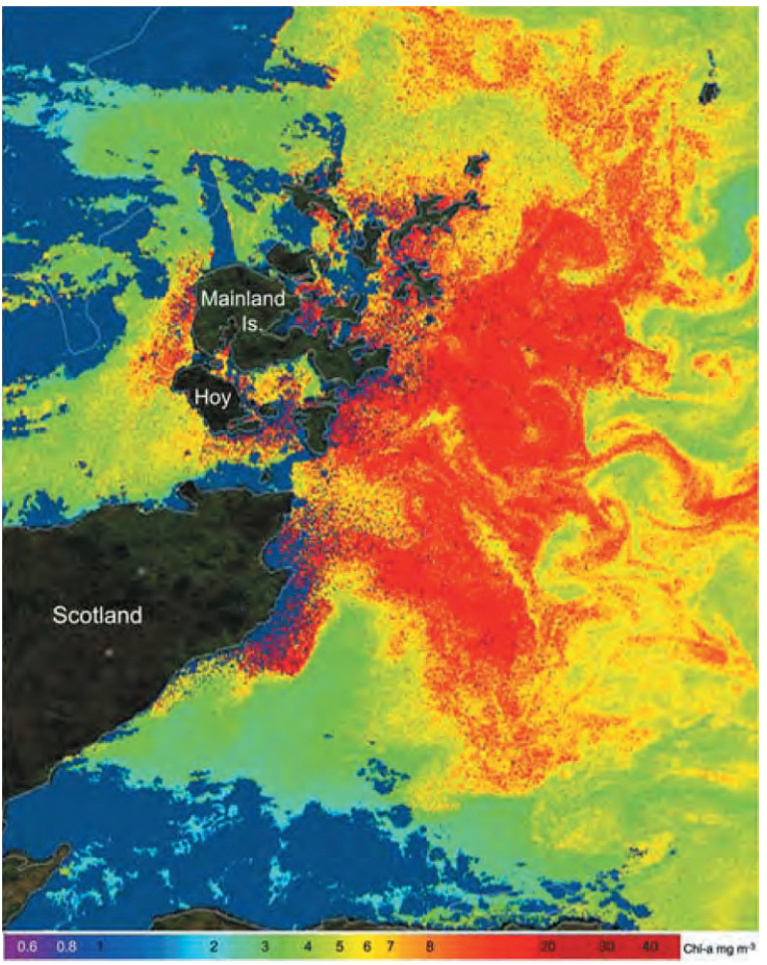


Figure 13.21. A bloom of harmful *Karenia mikimotoi* to the east of the Orkney Isles, Scotland captured by MERIS in full-resolution mode (300 m) on July 31, 2006 (image provided by Steve Groom, Plymouth Marine Laboratory, U.K.; MERIS data provided by ESA).

cause a great nuisance, especially for leisure users of beaches and sheltered seas. Whether or not they are actually toxic, they make swimming or sailing very unpleasant. Figure 13.22 is an example of such a bloom which occurs some years in summer over parts of the Baltic Sea. The high surface reflectance, with patterns of enhanced brightness along surface convergence lines, can be seen from broadband visible sensors as well as dedicated ocean color imagers, enabling satellites to be used to accumulate statistics about the days per year when nuisance blooms occur, an important dataset for research into the causes of such blooms. The coasts and archipelagos of the Swedish east coast are very popular for summer holidaymakers: In recent years the duty oceanographer at the Swedish Meteorological and Hydrological Institute, responding to public enquiries about which regions are contaminated and which are clear, has made use of near real–time ocean color images, as well as forecasts of wind direction, to monitor wind-driven movement of surface

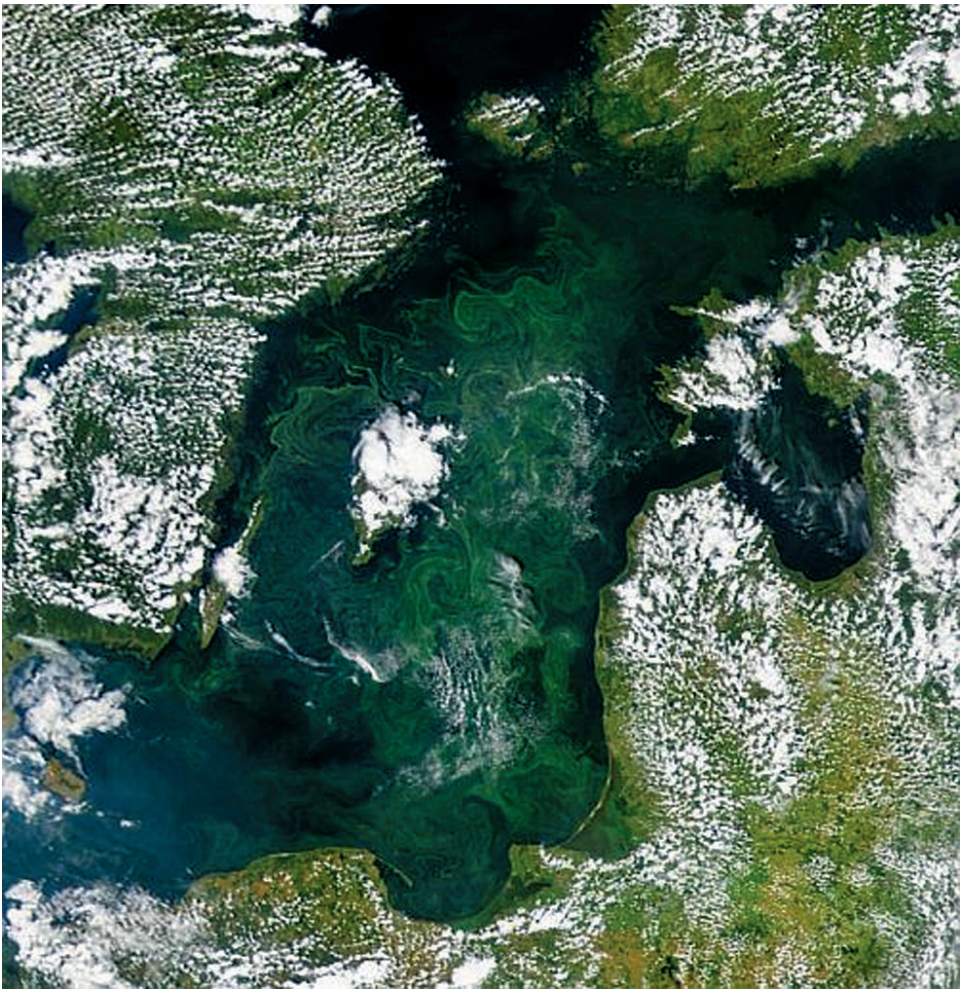


Figure 13.22. A near real–color composite from SeaWiFS over the Baltic Sea on July 24, 2003, showing a surface manifestation of a bloom of the cyanobacteria, *Nodularia spumigena* (image obtained from the NASA Ocean Color website—*http://oceancolor.gsfc.nasa.gov/*).

material. This is one example among many from around the world of how the ready availability of such images is being put to use for the public good, using a largely qualitative approach.

More penetrating scientific applications, such as the monitoring of chlorophyll concentration levels throughout the year in order to assess the potential for eutrophication, requires accurate quantitative retrievals of ocean variables from ocean color satellite data. Such studies will benefit from integration of ocean color information within biogeochemical ecosystem models embedded within physical circulation models, to produce a continually renewed model forecast or nowcast of ocean con-

ditions, regularly updated with new observational data. This new, evolving, and important application of satellite data over shelf seas, which probably holds the key to maximixing information retrieval from satellite data in future, will be discussed further in the section on operational oceanography in Chapter 14. Nonetheless, while we wait for such technological developments, we should never let the implementation of machine-based ocean-forecasting systems obscure the richness of direct insight into the ocean which comes from simple color images such as Figures 13.22 and 13.4.

13.3 COASTAL ALTIMETRY

13.3.1 Challenges and opportunities for altimetry in coastal seas

It is widely accepted that satellite altimetry is a mature observational technique that makes a significant contribution to a variety of research areas and operational applications in oceanography, as illustrated by examples in many of the earlier chapters of this book. However, established methodologies for using altimeters to measure sea surface height (SSH) and significant wave height, as outlined in chapter 11 of *MTOFS* (Robinson, 2004), have explicitly flagged the unreliability of altimeter observations in coastal seas and within about 200 km of the coast. Generally such data are excluded from standard, gridded, multi-sensor data products.

There are several reasons that such caution is justified, relating to the number of different corrections that must be applied to retrieve the distance between the sensor and the sea from the pulse return time, and then to estimate the height of the sea surface relative to the geoid, or to a reference surface. The accuracy of these corrections is compromised by proximity to the coast or by the different behavior of the ocean in shelf seas compared with the open ocean. Sometimes altimeter records cannot be interpreted at all close to coasts using assumptions, such as geostrophy, that are applicable in the open ocean. Moreover, close to the coast standard tracker algorithms that characterize the shape of the altimeter pulse echo are less able to deliver useful results.

Amongst the factors which introduce problems to standard altimeter processing methods are contamination of the signals from both the altimeter and the accompanying microwave radiometer, used for atmospheric correction, when the footprint contains some land. Another important issue is the uncertainty of tidal corrections in shelf seas where tidal amplitudes are typically much larger than over the open ocean, and much harder to predict confidently because of nonlinear tidal dynamics and the interactions between tides and wind-driven currents in shallow waters. For similar reasons it is more difficult to remove atmosphere effects (wind and pressure) from the sea surface height anomaly (SSHA) over shelf seas. Additionally the behavior of sea surface roughness (the shape of the surface profile at scales of mm and cm which affect the reflection and scattering of electromagnetic energy at comparable radar wavelengths) in response to wind waves and swell tends to differ close to the coast depending on factors such as wave fetch, wave age, and the response of incoming

swell to shallow depths. This is likely to introduce uncertainty into the retrieval of significant wave height and wind speed products themselves, and to add further uncertainty to the SSHA through sea state bias correction.

In the early years of ocean altimetry, when satellite oceanographers were still learning about how altimeters perform and perfecting sensor and systems design, it made no sense to worry about the problems caused to the altimeter record by coastal and shelf seas. Given that the driving goals of ocean altimetry were to measure large-scale ocean circulation and mesoscale variability of currents far from land, to determine tides of the open ocean where they were largely unknown, and to detect trends and variability in mean sea level well away from land, there was little motivation to invest scientific effort into exploring coastal issues. Given the empirical evidence that altimeter geophysical data records close to the coast were very noisy, and since the reasons for this were broadly understood, it was logical simply to flag such data as bad records and ignore them. This allowed attention to be devoted to critical tasks like improving orbit monitoring accuracy to a few centimeters and learning how to interpret the SSHA relative to a long-term mean sea surface when there was no independent measure of the geoid.

Fifteen years after the launch of the TOPEX/Poseidon mission, the situation is different. Mainstream dynamical oceanography now makes full use of altimetry as a measurement tool alongside floats and drifters and is linked closely to numerical models for research and monitoring of open-ocean circulation. Aware of this success, shelf seas and coastal oceanographers start to ask why altimetry cannot offer them similar successes for studying shelf sea dynamics. After all, in broad terms the scientific challenge to scientists to monitor and forecast ocean-dynamic behavior is most applicable in shelf seas where the benefits of such endeavors to human society are most immediate. Hence in the last few years a number of ocean altimetry specialists have started to explore options for making better use of data already collected over coastal waters for many years, but discarded or ignored (e.g., Vignudelli *et al.*, 2005). They have also started to consider designing new altimeters or better processing systems to overcome problems which have hitherto attached large error bars to coastal altimetry records.

Starting in February 2008 (Smith *et al.*, 2008) a series of international workshops on coastal altimetry have attracted growing interest by bringing together coastal oceanographers, who are potential users of altimeter records within 200 km of the coast, and altimeter specialists with ideas about how to improve the accuracy and usability of such data. Although at the time of writing there are few published papers this is likely to change and a book is in preparation (Vignudelli *et al.*, 2010) which summarizes activities reported at the second workshop in November 2008. In the rest of this section we shall look first at the promise and potential benefits to coastal and shelf oceanography if coastal altimetry can be made to work, and then outline the different technical approaches being developed to allow this to happen.

13.3.2 Potential applications of coastal and shelf altimetry

Perhaps the fields of research and operational applications from which demand has

been strongest for improved quality of altimeter records within 200 km of the coast are those for regions where there is not a wide continental shelf. Here, studies of ocean dynamics which routinely use altimeter data are frustrated by the inability to extend their monitoring close to the coast. One clear example, already touched on in Section 5.1.2, is the detection of coastal upwelling and its distinction from equatorward eastern boundary currents that occur in the same locations. As illustrated in Figure 5.6 there is a strong sea surface height signature associated with the front which bounds the ocean side of the upwelling region. Its magnitude is of order 25 cm over a distance of 20 km to 40 km, determined by the local Rossby radius. This should be detectable even in noisy altimeter records. Recent studies of U.S. west coast upwelling have already shown how plotting the SSHA along individual altimeter tracks crossing the coastline can pinpoint the upwelling front quite effectively.

Fisheries scientists, who already use altimetry to detect mesoscale patterns of ocean currents in relation to other water properties and fish behavior, want to be able to extend their studies closer to the coast than 200 km in regions where the continental shelf is narrow and open ocean–dynamical features extend close to the coast. This type of research already makes use of coastally based, longwave radar (grazing incidence) installations to monitor current fields, and also makes use of current vector measurements derived by maximum cross-correlation methods from satellite SST imagery, but this is weather-dependent. The attraction of using coastal altimetry is that it would be consistent with current measurements already made by altimetry in the adjacent open ocean beyond 200 km from the coast. In this and the previous example of upwelling studies, uncertainties of tides and dynamic decoupling of the shelf from the adjacent ocean are unlikely to be important factors affecting the accuracy and interpretation of the altimeter SSHA record. The main uncertainty issues seem to be that altimetric height measurement is corrupted by the effects of the continental landmass on standard atmospheric corrections applicable over the open ocean, and direct corruption of land in the field of view of the altimeter and its accompanying radiometer.

Another group of oceanographers looking for improved coastal altimetry data consists of shelf sea dynamics modelers whose work increasingly underpins the monitoring and management of shelf seas. As well as seeking to describe circulation and water movements they also have to represent interactions between tides, winds, and atmospheric pressure in order to predict coastal sea levels and possible storm surges. For this they rely on coastal and offshore tide gauges, but they would benefit greatly from being able to use altimeter records in near-real time. In this case, the invalidity of assumptions about geostrophy made in open-ocean applications of altimetry is not an issue, since what is needed is not estimates of geostrophic currents but direct observations of actual surface height and slope along altimeter tracks. These are needed in order to validate, or adjust through assimilation, their model predictions of sea surface height.

What is common to all these dynamical oceanographers is, on the one hand, an underlying need for measurements of SSH variability at fine spatial resolution promised in the uncertain future by wide-swath altimetry (as outlined in section 11.5.5 of *MTOFS*—Robinson, 2004) but, on the other hand, an immediate requirement to

improve the accuracy of data from altimeters already operating, and recording echoes, over the coastal ocean and shelf seas around the world. Given the wide spacing between altimeter swaths, they have no illusions that they can obtain dense coverage with existing altimeters, but they would like to exploit more fully the along-track resolution of a few kilometers which altimeters achieve. They envisage using records from individual altimeter tracks in conjunction with dynamical models. They are therefore much more interested in immediate access to individual tracks as soon as they are acquired than in waiting several days for gridded datasets integrated over an altimeter cycle, which have little if any validity in shelf seas. There would still be a need for reprocessing the record within near-real time, say within 6 to 24 hours of acquisition, allowing improvement to orbit corrections and giving time to pull together the additional regional information about atmospheric and sea state conditions that would enable them to apply their own, locally informed, improved corrections to the dataset, matched to their specific operational needs.

What is emerging is specification of an operational coastal altimetry product. This would deliver along-track records from all altimeters within 200 km of the coast, covering the domain illustrated in Figure 13.23. Records should be processed or (in the case of historical data covering previous altimeter missions) reprocessed to the highest precision possible and with records sampled at the highest frequency available (as discussed in the next subsection). The dataset should also contain auxiliary data for correction purposes that are based on the best available regional models. There needs to be a common format between data providers that will make it easy to deal with data from all available altimeters, since this will increase the spatiotemporal density of coverage.

It should also be pointed out that there is a similar demand for specially processed altimeter products for significant wave height (SWH) and wind speed. SWH data delivered readily in locally truncated coastal segments is attractive to

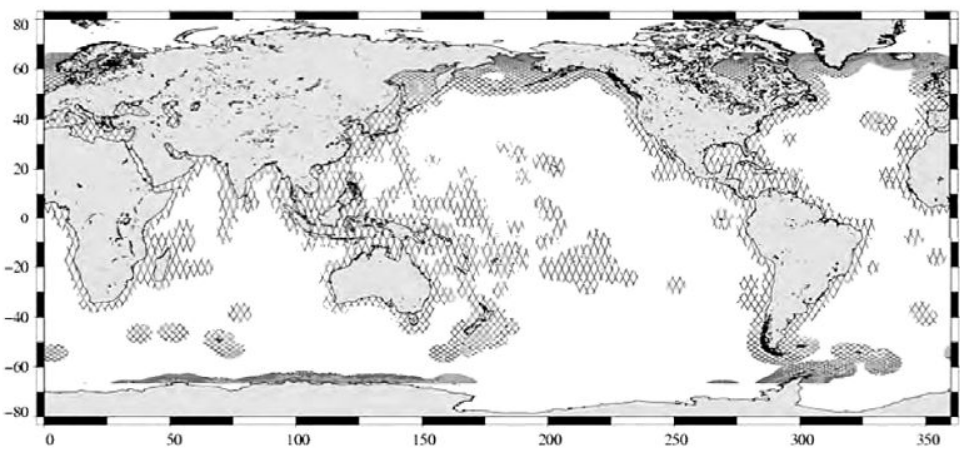


Figure 13.23. Proposed geographical domain in which a specialized coastal altimetry data product is required.

wave modelers and forecasters. Similarly wind speed estimates can be compared with other sources of wind speed, from scatterometry, radiometry, and NWP, in research towards improved high-resolution observation and forecasting of coastal winds.

13.3.3 Practical approaches to improving altimeter accuracy in shelf seas

Of the different classes of problems for coastal altimetry that were identified in Section 13.3.1, those related to the uncertain effect on SSHA of tides, atmospheric pressure, wind stress, and small-scale ageostrophic processes in shelf seas become much less of a problem in the context of shelf sea applications outlined in the previous subsection. They would certainly present difficulties if the objective were to remove the effect of tides and the inverse barometer in order to determine sea level topography and associated geostrophic currents independently of these factors. However, when determining the dynamics of shelf seas it is important to explicitly include interactions between tides and other dynamic processes and, therefore, sea surface topography uncorrected for such factors is the most useful measurement. For those open-ocean coasts without extensive shelves, corrections for tides and barometric pressure remain appropriate, and there it will be important to use the best regional or local tidal models and the best meteorological measurements so that corrections can be as accurate as possible.

Standard corrections to adjust SSH for effects of water vapor in the troposphere need to be refined to cope with greater spatial variability in coastal areas and contamination of microwave radiometry by land emissions. This requires co-ordinated international effort while recognizing that different areas may require different treatment. The same is true for ionospheric correction. The issue here is that the use of dual frequencies by most altimeters to measure this effect can cause problems close to the coast where different frequencies have different footprints so that they encounter land reflections at different times. Ways are being developed to cope with this. Meanwhile the uncertainty of sea state bias remains an unresolved issue. It is recognized that surface slope distribution and asymmetry of waves probably differ in coastal seas from conditions in the open ocean on which empirical constants in sea state bias correction are based, but there is no consensus on how to make improvements. The broad approach to improving these corrections near the coast, where standard methods are less reliable, is to encourage the development of locally tuned corrections and to seek to provide appropriate regional ancillary data alongside the coastal altimetry product, allowing users to apply corrections themselves.

The aspect of altimeter processing that requires the most fundamental attention to deliver meaningful altimeter data within 10 km to 30 km of the coastline is that of echo waveform analysis. There is consensus that retracking of the original data is essential, with the aim that specific coastal retrackers should give better accuracy and precision than generic deep-ocean retrackers. Different approaches are being tried out. One approach uses physically based parametric algorithms which build on open-ocean trackers and try to allow for the effect of some land surface reflections within the field of view. Another is to develop empirical algorithms. An alternative is to classify the shape of the waveforms into those that are recognizably ocean

reflections and can be analyzed conventionally and those that cannot. There is a need for improved digital terrain models and coastline models so that trackers can be matched to particular coastlines. This also calls for a higher frequency of tracking that, where possible, analyzes echoes averaged over smaller spatial intervals along-track.

Work to improve retracking algorithms for coastal situations is also linked to new altimeter design concepts such as Delay-Doppler used in the SIRAL class of altimeters developed by ESA. This raises the possibility that future altimeters may be designed specifically to improve capacity to resolve coastal effects and maximize the accuracy of height retrieval from sea surfaces adjacent to land.

13.4 COASTAL AND ESTUARINE REMOTE SENSING

13.4.1 Important edges of the ocean

This section considers the ways in which satellite data are applied to research and operational tasks in estuaries and in the coastal zone. The distinction between this and what has already been presented about shelf sea remote sensing is one of scale; the focus here is on the processes that occur in estuaries, on beaches, and in coastal embayments whose geographical extent is less than about 10 km to 20 km and may be much smaller. These are the regions close to the coast which are not shown in detail in the typical shelf sea images displayed earlier in this chapter. Indeed many of the estuaries and embayments are not adequately resolved at all in those data.

Nonetheless, despite their small geographical extent, these regions have great importance by virtue of being marginal parts of the ocean adjacent to land. For some people this is the only part of the sea that they encounter. It is certainly the zone of the sea that has most impact on their lives and for which scientific understanding is of relevance; warning of high tides and the danger of coastal flooding, explaining the causes of coastal erosion and sedimentation, monitoring pollution of coastal and estuarine waters, alerting coastal navigation about hazardous currents, and informing the tourist about water temperatures for bathing and the best beaches for surfing, among many other applications. There is a growing expectation within the populations of many maritime nations that reliable scientific information related to such issues should be readily available for public access. Consequently there is a growing demand for just the kind of regular, spatially detailed monitoring of the coastal marine environment, with wide regional coverage, that satellite remote sensing, in principle, can supply.

Of all the subareas in which satellite oceanography can be applied, we might expect this to be the most active. Satellite data have certainly attracted serious attention from coastal and estuarine scientists since the first Landsat images became available in the 1970s. However, in practice the improving sensors and maturing methodology of satellite oceanography have not so far led to the same opening of new scientific opportunities in coastal oceanography as they have for the open ocean. With a few exceptions, there are in fact serious limitations on the ability

of satellite techniques to provide quantitative measurements of coasts and estuaries that are needed operationally or scientifically. For that reason this section on coastal oceanography from space is brief, although if this were a book about remote sensing generally and not constrained to satellite techniques the story would be rather different. The rest of this section first addresses the fundamental reasons for this situation and then briefly reviews some applications where satellite data have found an established, if limited and typically qualitative, role.

13.4.2 A mismatch of scales?

The problem lies in the relatively short lengthscales and timescales of processes encountered in coastal oceanography, and the difficulty for satellite remote sensing methods of sampling at high enough frequencies. This issue is discussed at some length in section 4.5 of *MTOFS* and illustrated here in Figure 13.24. As soon as users demand resolution of lengthscales shorter than 1 km at timescales of less than 1 day, most of the techniques of satellite oceanography applied to other applications in this book are found wanting. Those sensors that can sample daily, such as wide-swath, medium-resolution, visible and infrared radiometers, do not have the spatial resolution needed to explore estuaries and the coastal zone. Scatterometry, altimetry, and microwave radiometry cannot approach resolutions less than 10 km and their

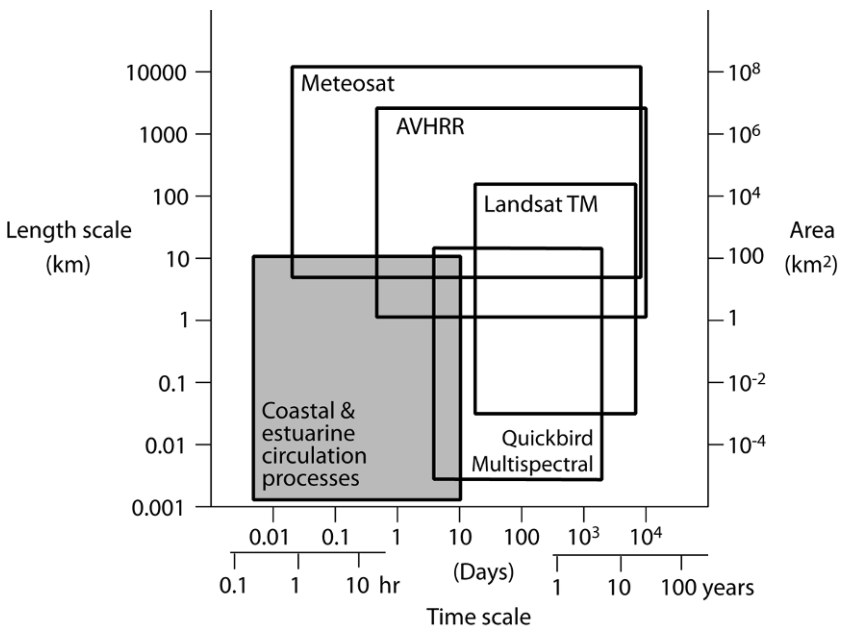


Figure 13.24. Typical length and timescales of coastal and estuarine processes (shaded square) compared with the spatial and temporal sampling capabilities of typical classes of satellite image data.

performance is compromised close to the coast. Even the hoped-for improved performance of coastal altimetry (discussed in Section 13.3) is not able to image the coastal zone, but simply to extend studies of shelf-dynamic processes to closer inshore. Microwave radiometers encounter noise through their antenna side lobes from contrasting microwave emissions from the land surface, and their coastal data would therefore be unreliable even if their resolution could be improved.

Sensors more appropriate for the task of measuring small-scale features in coastal seas and estuaries are the high-resolution visible imagers such as Landsat TM and SPOT HRV (see section 6.4.5 in *MTOFS*), and the SARs (chapter 10 of *MTOFS*). While these can approach spatial resolution of about 20 m, none of them has a revisit interval shorter than about 10 days. Until there are constellations of such sensors in orbit to ensure at least daily overpasses, they are unable to sample often enough to see some of the ephemeral features that come and go in shallow seas. To monitor important dynamical and physical processes in estuaries it would be necessary to explicitly resolve the semidiurnal tide which is the dominant timescale of estuarine variability in many parts of the world. One possible configuration being contemplated is to place a fine-resolution sensor in geostationary orbit, allowing it to "stare" at a limited geographical area and thus sample a few times per hour. It presents a sensitivity challenge for an imager at an altitude of 36,000 km to resolve down to a few tens of meters, and so far no such sensor has been developed.

Moreover, for visible waveband sensors to be useful for measuring water content they would require fine spectral resolution comparable with MERIS and MODIS. The finest resolution sensors, resolving down to 1 m to 5 m, have been panchromatic, until very recently (see Table 7.2). They have a role for mapping beaches and shorelines: with the Sun at a propitious angle they may also reveal the patterns of swell refraction on beaches, but without spectral discrimination their useful information content is limited.

Therefore we should not expect to transfer successful applications of remote sensing in mesoscale and global oceanography to short lengthscales and timescales of coastal oceanography. In particular it must be accepted that the best time sampling achievable from satellites will not resolve most dynamic processes in coastal water, and we should therefore not expect to apply satellite data in that way. For example, it would be inappropriate to use satellite data to study how the distribution of suspended sediment concentration in an estuary varies over a 12-hour tidal cycle. To monitor that type of process would require the use of airborne remote sensing. Instead, we should make the most of infrequent snapshots from fine-resolution imaging sensors, which may be separated by many days. The applications considered in Section 13.4.3 are mainly of this type.

Accepting this limitation, it would be wrong to underestimate the potential usefulness of a few high-resolution images per year in certain types of application where satellites are effective at detecting changes that contain operationally or scientifically important information. Thus, for example, the seasonal advance and retreat of coastal vegetation, or changes in wetlands or exposed banks, may be characterized by a few satellite images per year, unless there is also high-frequency variability (e.g., within the 12.4-hour tidal cycle) which would alias the interpretation of occasional

snapshot views. In contrast to this, it has proved to be possible to retrieve tidal information (e.g., the way suspended sediment distribution varies through a tidal cycle) by careful use of the available, historic Landsat record, if it can be assumed that any long-term changes are very much smaller than the massive semidiurnal tidal changes occurring in macrotidal regions, such as the Bristol Channel and Severn Estuary in the U.K. In such cases it is possible to select examples of images from different tidal phases in order to construct a sequence of the tidal cycle, even though actual images were acquired in a completely different sequence in real time.

Another application context for using a sequence of infrequently acquired images is when a singular event takes place, such as a damaging storm, an earthquake, a tsunami, or another cause of coastal flooding. Then the changes caused by the event to a variety of environmental parameters can be monitored by comparison of images acquired before and after the event. However, when disasters occur, it is not very helpful if rescue and relief agencies have to wait days or weeks for the next images to be available. There is often a more pressing need to know about changes to coastlines and water levels, damage to the coastal infrastructure, extent of flooding, etc. as soon as possible. Satellites offer a powerful means of quickly mapping changes to an area which has previously been monitored by the same sensor, but a daily sampling capability is essential if disaster recovery operations are to be able to rely on them.

In response to this very situation, a constellation of small satellites has been designed and built by Surrey Satellite Technology Ltd. (SSTL) in the U.K. and at least six are in operation (DMC, 2009). The Disaster Monitoring Constellation (DMC) was designed as a proof-of-concept constellation, capable of multispectral imaging of any part of the world every day. It is unique in that each satellite is independently owned and controlled by a separate nation, but the set of satellites has been evenly distributed around a single Sun-synchronous orbit to provide daily imaging capability using all of them together to sample the same location. Each small satellite has an imaging capability of 600×600 km with a resolution of 32 m or better. Although not designed for continuous monitoring around the whole orbit, the 600 km swath allows a 6-day revisit from an individual satellite, which is normally used by the owner nation for their own national mapping or environmental monitoring purposes. However, when a disaster occurs, acquisition plans can be changed so that all the satellites acquire data over the disaster zone, under the terms of the International Charter for Space in Major Disasters.[1] Daily resolution enables immediate mapping of the situation, and how it changes from day to day. Some disasters for which this constellation has been used were coastal events, including the Indian Ocean Tsunami of December 2004, and the flooding of New Orleans by Hurricane Katrina in August 2005.

In principle daily sampling may become possible using constellations of very high–resolution mapping sensors like Quickbird or WorldView (see Table 7.2) which resolve down to 1 m or less in several spectral bands. Although their swaths of about 10 km to 20 km allow them to cover only a small fraction of the Earth's surface each

[1] See *http://www.disasterscharter.org/web/charter/home*

day, their along-track and cross-track pointing ability can select a particular region to view from orbits that do not pass exactly overhead. This would allow sampling of a study region nearly every day or two, and with two or more similar satellites daily sampling can be guaranteed. However, to do so would commit a significant fraction of the sensor's schedule and prevent it from observing other scenes. For most coastal science research programs this makes it prohibitively expensive at today's commercial prices, but in principle daily sampling at 1 m resolution over a 10 km to 20 km scene will soon be technically achievable and the scientific potential for such data is being explored by agencies such as the NOAA Coastal Services Center in the U.S.A. (NOAA-CSC, 2009). As with all visible waveband remote sensing, sampling frequencies are dependent on cloud-free conditions.

13.4.3 Coastal remote-sensing applications using satellite data

For the reasons outlined in the previous subsection, there has been little reported scientific work that has focused on the unique use of satellites for studying coasts and estuaries. There is a much wider literature on coastal and estuarine remote sensing involving airborne and other remote-sensing methodologies, and a number of books on the subject (Green *et al.*, 2000; Miller *et al.*, 2005; Yang, 2009). This subsection offers a brief overview of some of the ways that satellite data can contribute to the science of estuarine and coastal processes. The following examples are not intended to be exhaustive, and the references cited are only representative pointers to the literature.

Waves and beach management

Surface waves are one of the most important factors which control the coastal environment. In addition to tides, it is swell waves reaching a beach which help to shape small-scale dynamical structures such as longshore drift and rip currents that influence the deposition, transport, or erosion of beach material. The various ways in which waves are measured from satellites is discussed in Chapter 8. The direction, wavelength, and amplitude of incident swell are important parameters for a beach engineer. Typically these are obtained from regional wave models, validated by wave buoys although those may be a long distance from the beach of interest. Synthetic aperture radar (SAR) images can provide clear patterns of surface swell showing the refraction of waves as they approach the shore. Although quantitative retrieval of wave amplitude and spectra from SAR still contains some uncertainty, and SAR data are not available frequently enough to record the full wave history of a beach, a combination of SAR data and wave models has the potential to provide the beach engineer with a richer set of information.

Research studies have shown how complementary use of SAR image data and WAM third-generation wave models can lead to a more complete picture of the wave field (Ocampo-Torres *et al.*, 1997; Ocampo-Torres, 2001). For example, analysis of archived SAR images can reveal the transition of the wave field from far offshore, where the wave model is most reliable, to the near-shore region where it interacts

with the shallowing sea bed as waves propagate towards the beach. The archived SAR record should also be able to show how such wave transitions vary according to the initial direction(s) of incoming wave energy. These studies were performed along the Pacific coast of Mexico, where another benefit of using SAR data was that they detected swell that had traveled thousands of kilometers across the Pacific and which could be missed if only a regional scale wave model were used. There is scope for similar studies to be extended to many other regions, with the potential for wider operational use of SAR to complement wave models, now that long-term provision of SARs for operational ocean tasks is expected to be secure (see Section 14.5.3).

Coastline and beach management also requires careful monitoring of how beach material is distributed Although beach surveys to ascertain movements of material require careful *in situ* surveying, high-resolution aerial photography, which may in future come from satellites, can provide a wider overall picture of what is happening to the beach. For example, aerial maps can show changes in the shoreline, the distribution of beach cusps and how they interact with incoming swell, patterns of suspended sediments just off the shoreline, or major changes following storm events. It also allow the behavior of a beach to be monitored during and following beach-engineering work such as material replenishment or the introduction or removal of hard engineering such as groynes and breakwaters.

Coastal flooding and shoreline protection

Another important role for coastal remote sensing is in monitoring the height of land adjacent to the coast and especially the height relative to the sea surface of any banks used for flood defence. From such information flood risk assessment maps can be compiled, to be used for a variety of tasks such as emergency planning, identifying which roads, railways, and other services are at risk of flooding, influencing decisions on new building and development plans, and providing baseline information for insurance purposes.

Mostly such work uses conventional, land remote-sensing techniques for mapping and wetland classification using high-resolution visible sensors. Coastal elevation mapping makes use of airborne lidars, which can measure the height of the primary sea defence crest line and monitor changes resulting from erosion. In future there is the possibility that laser altimeters may be flown on satellites so that such observations could be made more widely. There is also the interesting example of a study that used SAR interferometry to monitor the rate of subsidence in New Orleans during the period 2002–2005, prior to Hurricane Katrina (Dixon *et al.*, 2006). It shows that there was a mean subsidence rate of 6.4 mm/yr and a maximum of 33 mm/yr at some critical locations.

Coastal ecosystems

High-resolution satellite data have been used quite extensively for monitoring coastal and estuarine ecosystems, especially in the tropics. This is discussed in more detail in Section 7.5. This partly includes the mapping of tropical coral reef

systems, but the wider use of satellite oceanography methods to identify the risk of coral bleaching is also discussed in Section 7.6.

13.5 REFERENCES

Alpers, W., L. Mitnik, L. Hock, and K. S. Chen (1999), *The Tropical and Subtropical Ocean Viewed by ERS SAR*. ESA ESRIN, available at *http://www.ifm.uni-hamburg.de/ers-sar/* (last accessed April 25, 2008).

Berthon, J.-F., F. Mélin, and G. Zibordi (2008), Ocean colour remote sensing of the optically complex European seas. In: V. Barale and M. Gade (Eds.), *Remote Sensing of the European Seas* (pp. 35–52), Springer-Verlag, Dordrecht, Netherlands.

Bogazzi, E., A. Baldoni, A. S. Rivas, P. Martos, R. Reta, J. M. Orensanz, M. Lasta, P. Dell'Arciprete, and F. Werner (2005), Spatial correspondence between areas of concentration of Patagonian scallop (*Zygochlamys patagonica*) and frontal systems in the southwestern Atlantic. *Fish. Oceanogr.*, **14**(5), 359–376.

Bowers, D. G., and J. H. Simpson (1987), The mean position of tidal fronts in European-shelf seas. *Continental Shelf Res.*, **7**, 35–44.

Bowers, D. G., S. Boudjelas, and G. E. L. Harker (1998), The distribution of fine suspended sediments in the surface waters of the Irish Sea and its relationship to tidal stirring. *Int. J. Remote Sensing*, **19**, 2789–2805.

Bowers, D. G., S. Gaffney, N. White, and P. Bowyer (2002), Turbidity in the southern Irish Sea. *Continental Shelf Res.*, **22**(15), 2115–2126.

Brink, K. H. (1998), Deep-sea forcing and exchange processes. In: K. H. Brink and A. R. Robinson (Eds.), *The Sea, Vol. 10, The Global Coastal Ocean: Processes and Methods*. John Wiley & Sons, New York.

Bulgarelli, B., G. Zibordi, and J.-F. Berthon (2003), Measured and modeled radiometric quantities in coastal waters: Toward a closure. *Appl. Opt.*, **42**(27), 5365–5381.

Chang, G. C., T. D. Dickey, C. D. Mobley, E. Boss, and S. Pegau (2003), Toward closure of upwelling radiance in coastal waters. *Appl. Opt.*, **42**, 1574–1582.

Csanady, G. T. (1997), On the theories that underlie our understanding of continental shelf circulation. *J. Oceanogr.*, **53**, 207–229.

Cullen, J. J., A. M. Ciotti, R. F. Davis, and M. R. Lewis (1997), Optical detection and assessment of algal blooms. *Limnol. Oceanogr.*, **42**, 1223–1239.

Darecki, M., A. R. Weeks, S. Sagan, P. Kowalczuk, and S. Kaczmarek (2003), Optical characteristics of two contrasting Case 2 waters and their influence on remote sensing algorithms. *Continental Shelf Res.*, **23**(3/4), 237–250.

Dixon, T. H., F. Amelung, A. Ferretti, F. Novali, F. Rocca, R. Dokka, G. Sella, S.-W. Kim, S. Wdowinski, and D. Whitman (2006), Space geodesy: Subsidence and flooding in New Orleans. *Nature*, **441**(June 1), 587–588.

DMC (2009), *Disaster Monitoring Constellation International Imaging*, available at *http://www.dmcii.com/index.html* (last accessed July 31, 2009).

Doerffer, R., and H. Schiller (2007), The MERIS case 2 water algorithm. *Int. J. Remote Sensing*, **28**(3/4), 517–535.

Dogliotti, A. I., I. R. Schloss, G. O. Almandoz, and D. A. Gagliardini (2009), Evaluation of SeaWiFS and MODIS chlorophyll-a products in the Argentinean Patagonian Continental Shelf (38°S–55°S), *Int. J. Remote Sensing*, **30**(1), 251–273.

Durand, D., L. H. Pettersson, O. M. Johannessen, E. Svendsen, H. Søiland, and M. Skogen (2002), Satellite observation and model prediction of toxic algae bloom. *Operational Oceanography: Implementation at the European and Regional Scales* (Elsevier Oceanography Series Vol. 66, pp. 505–515). Elsevier.

Froidefond, J.-M., P. Castaing, and R. Prud'homme (1999), Monitoring suspended particulate matter fluxes and patterns with the AVHRR/NOAA-11 satellite: Application to the Bay of Biscay. *Deep-Sea Res. II*, **46**, 2029–2055.

Gao, B.-C., M. J. Montes, R.-R. Li, H. M. Dierssen, and C. O. Davis (2007), Modification to the atmospheric correction of SeaWiFS ocean colour images over turbid waters. *IEEE Trans. Geosc. Remote Sensing*, **45**(6), 1835–1843.

Glorioso, P. D. (1987), Temperature distribution related to shelf-sea fronts on the Patagonian Shelf. *Continental Shelf Res.*, **7**(1), 27–34.

Green, E. P., P. J. Mumby, A. J. Edwards, and C. D. Clark (Eds.) (2000), *Remote Sensing Handbook for Tropical Coastal Management* (Coastal Management Sourcebooks, x + 316 pp.). UNESCO, Paris.

Gupta, A., and P. Krishnan (1994), Spatial distribution of sediment discharge to the coastal waters of South and Southeast Asia. *Variability in Stream Erosion and Sediment Transport* (IAHS Publication No. 224, pp. 457-463). International Association of Hydrological Sciences, Christchurch, New Zealand [*Proceedings of the Canberra Symposium, December 1994*].

Hill, A. E. (1998), Buoyancy effects in coastal and shelf seas. In: K. H. Brink and A. R. Robinson (Eds.), *The Sea, Vol. 10, The Global Coastal Ocean: Processes and Methods*. John Wiley & Sons, New York.

Hu, J. Y., H. Kawamura, and D. L. Tang (2003), Tidal front around the Hainan Island, northwest of the South China Sea. *J. Geophys. Res.*, **108**(C11), 3342, doi: 10.1029/2003JC001883.

Huthnance, J. M. (1995), Circulation, exchange and water masses at the ocean margin: The role of physical processes at the shelf edge. *Prog. Oceanogr.*, **35**, 353–431.

IOCCG (2000), *Remote Sensing of Ocean Colour in Coastal, and Other Optically Complex Waters* (edited by S. Sathyendranath, Reports of the International Ocean Colour Co-ordinating Group No. 3, 140 pp.). IOCCG, Dartmouth, Canada.

IOCCG (2006), *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms and Applications* (edited by Z.-P. Lee, Reports of the International Ocean Colour Coordinating Group No. 5). IOCCG, Dartmouth, Canada.

IOCCG (2008), *Why Ocean Colour? The Societal Benefits of Ocean-Colour Technology* (edited by T. Platt, N. Hoepffner, V. Stuart, and C. Brown, Reports and Monographs of the International Ocean Colour Coordinating Group No. 7, 141 pp.). IOCCG, Dartmouth, Canada.

Kahru, M., and B. G. Mitchell (1998), Spectral reflectance and absorption of a massive red tide off southern California. *J. Geophys. Res.*, **103**, 21601–21609.

Lavender, S. J., M. H. Pinkerton, G. F. Moore, J. Aiken, and D. Blondeau-Patissier (2005), Modification to the atmospheric correction of SeaWiFS ocean colour images over turbid waters. *Continental Shelf Res.*, **25**(4), 539–555.

Li, Y., W. Huang, and M. Fang (1998), An algorithm for the retrieval of suspended sediment in coastal waters of China from AVHRR data. *Continental Shelf Res.*, **18**(5), 487–500.

Loder, J. W., and D. A. Greenberg (1986), Predicted positions of tidal fronts in the Gulf of Maine region. *Continental Shelf Res.*, **6**(3), 397–414.

Mann, K. H., and J. R. N. Lazier (2006), *Dynamics of Marine Ecosystems: Biological–Physical Interactions in the Oceans* (Third Edition, 496 pp.). Blackwell Publishing, Oxford, U.K.

Miller, R. L., C. E. Del Castillo, and B. A. McKee (Eds.) (2005), *Remote Sensing of Coastal Aquatic Environments* (345 pp.). Springer-Verlag, Dordrecht, Netherlands.

Moore, C. M., D. Suggett, P. M. Holligan, J. Sharples, E. R. Abraham, M. I. Lucas, T. P. Rippeth, N. R. Fisher, J. H. Simpson, and D. J. Hydes (2003), Physical controls on phytoplankton physiology and production at a shelf sea front: A fast repetition-rate fluorometer based field study. *Mar. Ecol. Prog. Ser.*, **259**, 29–45.

Moore, G. F., J. Aiken, and S. J. Lavender (1999), The atmospheric correction of water colour and the quantitative retrieval of suspended particulate matter in Case II waters: Application to MERIS. *Int. J. Remote Sensing*, **20**(9), 1713–1733.

Morel, A., and L. Prieur (1977), Analysis of variations in ocean colour. *Limnol. Oceanogr.*, **22**, 709–722.

Morin, P., M. Wafar, and P. Le Corre (1993), Estimation of nitrate flux in a tidal front from satellite-derived temperature data. *J. Geophys. Res.*, **98**(C3), 4689–4695.

Nihoul, J. C. J., P. T. Strub, and P. E. La Violette (1998), Remote sensing. In: K. H. Brink and A. R. Robinson (Eds.), *The Sea, Vol. 10, The Global Coastal Ocean: Processes and Methods*. John Wiley & Sons, New York.

NOAA-CSC (2009), *Remote Sensing for Coastal Management*. National Oceanic and Atmospheric Administration, Coastal Services Center, available at *http://www.csc.noaa.gov/crs/rs_apps/* (last accessed August 3, 2009).

Ocampo-Torres, F. J. (2001), On the homogeneity of the wave field in coastal regions as determined from ERS-2 and RADARSAT synthetic aperture radar images of the ocean surface. *Scientia Marina*, **65**(Suppl. 1), 215–228.

Ocampo-Torres, F. J., A. Martínez Diaz de León, and I. S. Robinson (1997), Synergy of ERS radar information and modelled directional wave spectrum to estimate coastal region wave characteristics in the Gulf of Tehuantepec, Mexico. Paper presented at *Proc. of the Use and Applications of ERS in Latin America, Viña del Mar, Chile, November 25–29, 1996* (ESA SP-405, pp. 219–224). ESA, Noordwijk, The Netherlands.

Pingree, R. D., and D. K. Griffiths (1978), Tidal fronts on the shelf seas around the British Isles. *J. Geophys. Res.*, **83**, 4615–4622.

Pingree, R. D., P. R. Pugh, P. M. Holligan, and G. R. Forster (1975), Summer phytoplankton blooms and red tides along tidal fronts in the approaches to the English Channel. *Nature*, **258**, 672–677.

Pitcher, G. C., and S. J. Weeks (2006), The variability and potential for prediction of harmful algal blooms in the southern Benguela ecosystem. In: V. Shannon, G. Hempel, C. Moloney, J. D. Woods, P. Malanotte-Rizzoli (Eds.), *Benguela: Predicting a Large Marine Ecosystem* (pp. 125–146). Elsevier.

Prangsma, G. J., and J. N. Roozekrans (1989), Using NOAA AVHRR imagery in assessing water quality parameters. *Int. J. Remote Sensing*, **10**, 811–818.

Robinson, I. S. (2004), *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Robinson, I. S., D. Antoine, M. Darecki, P. Gorringe, L. Pettersson, K. Ruddick, R. Santoleri, H. Siegel, P. Vincent, M. R. Wernand *et al.* (2008), *Remote Sensing of Shelf Sea Ecosystems: State of the Art and Perspectives* (edited by N. Connolly, Marine Board Position Papers No. 12., 60 pp.). European Science Foundation Marine Board, Ostend, Belgium.

Rucker, J. B., R. P. Stumpf, and W. W. Schroeder (1990), Temporal variability of remotely sensed suspended sediment and sea surface temperature patterns in Mobile Bay, Alabama. *Estuaries*, **13**(2), 155–160.

Sharples, J., and J. H. Simpson (1995), Semi-diurnal and longer period stability cycles in the Liverpool Bay region of freshwater influence. *Continental Shelf Res.*, **15**(2/3), 295–313.

Sharples, J., and J. H. Simpson (1996), The influence of the springs neaps cycle on the position of shelf sea fronts. *Coast. Estuar. Stud.*, **53**, 71–82.

Simpson, J. H. (1981) Sea surface fronts and temperatures. In A. P. Cracknell (Ed.), *Remote Sensing in Meteorology, Oceanography and Hydrology* (pp. 295–311). Ellis Horwood, Chichester, U.K.

Simpson, J. H. (1997), Physical processes in the ROFI regime. *J. Mar. Syst.*, **12**, 3–15.

Simpson, J. H. (1998) Tidal processes in shelf seas. In: K. H. Brink and A. R. Robinson (Eds.), *The Sea, Vol. 10, The Global Coastal Ocean: Processes and Methods*. John Wiley & Sons, New York.

Simpson, J. H., and D. G. Bowers (1979), Shelf sea fronts' adjustments revealed by satellite IR imagery. *Nature*, **280**, 648–651.

Simpson, J. H., and J. Brown (1987), The interpretation of visible band imagery of turbid shallow seas in terms of the distribution of suspended particles. *Continental Shelf Res.*, **7**, 1307–1313.

Simpson, J. H., and J. Hunter (1974), Fronts in the Irish Sea. *Nature*, **250**, 404–406.

Simpson, J. H., and J. Sharples (1994), Does the Earth's rotation influence the location of the shelf sea fronts? *J. Geophys. Res.*, **99**(C2), 3315–3319.

Smith, W. H. F., P. T. Strub, and L. Miller (2008), First Coastal Altimetry Workshop. *EOS, Trans. Amer. Geophys. Union*, **89**(40).

Spitzer, D., R. Laane, and J. N. Roozekrans (1990), Pollution monitoring of the North Sea using NOAA/AVHRR imagery. *Int. J. Remote Sensing*, **11**, 967–977.

Stumpf, R. P., and J. R. Pennock (1989), Calibration of a general optical equation for remote sensing of suspended sediments in a moderately turbid estuary. *J. Geophys. Res.*, **94**, 14363–14371.

Stumpf, R. P., M. E. Culver, P. A. Tester, M. Tomlinson, and G. J. Kirkpatrick (2003), Monitoring *Karenia brevis* blooms in the Gulf of Mexico using satellite ocean color imagery and other data. *Harmful Algae*, **2**(2), 147–160.

Tang, D. L., I. H. Ni, F. E. Muller-Karger, and Z. J. Liu (1998), Analysis of annual and spatial patterns of CZCS-derived pigment concentrations on the continental shelf of China. *Continental Shelf Res.*, **18**, 1493–1515.

Tang, D. L., H. Kawamura, M. A. Lee, and T. V. Dien (2003), Seasonal and spatial distribution of chlorophyll-a and water conditions in the Gulf of Tonkin, South China Sea. *Remote Sens. Environ.*, **85**, 475–483.

Vignudelli, S., P. Cipollini, L. Roblou, F. Lyard, G. P. Gasparini, G. Manzella, and M. Astraldi (2005), Improved satellite altimetry in coastal systems: Case study of the Corsica Channel (Mediterranean Sea). *Geophys. Res. Letters*, **32**(L07608), doi: 10.1029/2005GL022602.

Vignudelli, S., A. Kostianoy, P. Cipollini, and J. Benveniste (Eds.) (2010), *Coastal Altimetry* (680 pp.). Springer. [ISBN 978-3-642-12795-3, due August 2010.]

Vogelzang, J. (1997) Mapping submarine sand waves with multiband imaging radar, 1: Model development and sensitivity analysis. *J. Geophys. Res.*, **102**(C), 1163–1181.

Vogelzang, J., G. J. Wensink, C. J. Calkoen, and M. W. A. van der Kooij (1997), Mapping submarine sand waves with multiband imaging radar, 2: Experimental results and model comparison. *J. Geophys. Res.*, **102**(C), 1183–1192.

Wang, D. X., Y. Liu, Y. Q. Qi, and P. Shi (2001), Seasonal variability of thermal fronts in the northern South China Sea from satellite data. *Geophys. Res. Letters*, **28**, 3963–3966.

Weeks, A. R., and J. H. Simpson (1991), The measurement of suspended particulate concentrations from remotely sensed data. *Int. J. Remote Sensing*, **12**, 725–737.

Weeks, A. R., J. H. Simpson, and D. G. Bowers (1993), The relationship between concentrations of suspended particulate material and tidal processes in the Irish Sea. *Continental Shelf Res.*, **13**, 1325–1334.

Yanagi, T., S. Igawa, and O. Matsudat (1995), Tidal front at Osaka Bay, Japan, in winter. *Continental Shelf Res.*, **15**(14), 1723–1735.

Yang, X. (Ed.), (2009), *Remote Sensing and Geospatial Technologies for Coastal Ecosystem Assessment and Management* (Lecture Notes in Geoinformation and Cartography, 561 pp.). Springer-Verlag, Berlin.

14

Putting ocean remote sensing to work

14.1 SATELLITES AND APPLIED OCEANOGRAPHY

14.1.1 Introduction

So far, in its objective of presenting the applications of satellite oceanography methods, this book has tended to focus on the results of scientific research. It has sought to emphasize new insights and improved understanding of ocean phenomena which have been discovered by observing and measuring from Earth-orbiting satellites. Each chapter has brought the perspective of a satellite remote-sensing view to bear upon a different aspect of ocean science. A clearer and more holistic view has emerged of processes like mesoscale eddies, large-scale dynamic features such as Rossby waves, air–sea interaction phenomena like El Niño or climatological events (e.g., the unexpected rapid summer retreat of north polar sea ice cover). It has also been pointed out along the way how remote-sensing methods make it easier to monitor the behavior of the ocean, not only to assist scientific research and discovery but also to support those who make a living at sea. This penultimate chapter picks up this theme of applied satellite oceanography, how ocean remote sensing can be put to work to serve the operational needs of human civilization in mankind's interactions with the sea.

The chapter contains a number of individual, but related, strands illustrating aspects of "Putting ocean remote sensing to work". There are two further subsections of Section 14.1 offering a reflective discussion on the importance of ocean monitoring for applications beyond mere scientific curiosity, and reasons why ocean scientists should contribute to such work. Section 14.2 explains what is involved in operational oceanography, noting in particular the important role that satellite data play in the ocean-forecasting systems now coming into operation. Section 14.3 is devoted to the particular challenge of how ocean color satellite data can be used to make marine ecosystem models more operationally effective. Section 14.4 discusses

how satellite data products need to be adapted before they can be usefully ingested into ocean-forecasting systems and to be suitable for other applied tasks distinct from scientific research applications. This includes a case study of an international collaborative program that has revolutionized the way SST data are treated. Section 14.5 outlines how oil spill monitoring from space, once primarily a topic of research, has now transitioned into an operational service. Finally Section 14.6 introduces the role of satellite ocean data in climate monitoring.

14.1.2 The fundamental importance of ocean monitoring and forecasting

Although this is the final substantive chapter, it is not simply an afterthought to suggest that scientific research should be shown to have some beneficial spinoff for society in general. On the contrary, the importance of the application of ocean remote sensing is intended to be the narrative culmination of the book. Where earlier chapters have told the story of how unique insights and understanding of ocean phenomena and processes have emerged from the use of satellite remote sensing, we now reach the exciting conclusion that such knowledge is transferable to a much wider user base that has requirements beyond mere scientific curiosity. The endeavor of many remote-sensing scientists over four decades has given us the tools to describe reliably the present condition of the upper ocean. We have acquired the scientific and technical capability needed to monitor ocean behavior and thereby attempt to predict it, using a combination of computer-based models, observations from satellites, and *in situ* sensors.

Once such ocean-forecasting systems (OFSs) are fully established and have matured we will have more comprehensive knowledge of the existing ocean state in near-real time, and the capacity to forecast with some confidence how it is likely to change in succeeding hours or days. This will make it much safer for mariners and all other users of the sea to go about their business better prepared for the extreme ocean phenomena which make it a hazardous environment in which to live and work. Moreover, having continuously updated knowledge of the ocean state should enable us to exploit the ocean sustainably, learning, for example, how to avoid overfishing or damaging critical ecological balances by what we put into or take out of the sea. While *in situ* measurements are also essential, especially for sampling the deeper ocean, it has to be recognized that without the unique perspective of satellite oceanography we would lack the global and holistic view of oceans that makes OFS possible. Ocean remote sensing now offers mankind a realistic opportunity to contemplate living in harmony with the ocean environment.

Improved knowledge of the current ocean state is already leading to better weather forecasting through, for example, better quality sea surface temperature data. Because the ocean covers 70% of the planet's surface, monitoring it is essential for understanding the whole Earth system. The perspective provided by satellite oceanography helps to clarify the ocean's role in climate and climate change, which is arguably the most challenging issue facing human civilization in the 21st century. It is interesting to reflect on how we have reached this point. Had there not been the "space race" in the 1960s leading to the growth of remote sensing in general,

which was supported by a group of creative and visionary ocean scientists who dared to dream about the far-fetched possibilities of "Oceanography from Space" (Ewing, 1965), ocean science might today lack both the observational tools and the scientific framework that are essential for engaging in Earth system science and climate change research. Today, operational oceanography from space is a reality. It may seem an inflated claim that the growth of satellite oceanography and the important applications summarized in this book have been essential to enable marine science as a whole to face the challenge of how mankind can survive on a finite planet. Yet without the capacity for global ocean monitoring we would lack the global perspective and observational inputs which convert numerical ocean models from computational tools to operational simulations of actual ocean state, the basis of an OFS.

Governments now acknowledge that systems delivering regular ocean measurements from space are essential tools for managing the impact of modern civilization on our finite planet. Substantial investments have been made in programs such as the European Union's Global Monitoring for Environment and Security (GMES). This initiative has committed expenditure on space hardware for EO systems over the next two decades, to ensure continuity of today's satellite ocean measurements. The concept of "operational oceanography" monitoring key ocean variables as a "public good" comparable with weather forecasting has become a reality. The existence of firm, funded programs, no longer merely the dreams of scientists, provides evidence that the everyday applications of ocean remote-sensing research have come of age.

14.1.3 Motivation for scientists to engage in applied ocean remote sensing

At the start of a chapter promoting the importance and value of the operational applications that have emerged from scientific and technical advances of satellite oceanography, it is worth a short digression to reflect on the distinction between "pure" and "applied" research. The term "pure" is generally assigned to research that is typically curiosity-driven, developing the fundamental intellectual structure of an oceanographic topic, refining its theoretical framework, or using probing experiments to characterize a particular ocean phenomenon. Applied research refers to the intellectual and problem-solving effort invested in creating systems that harness the results of pure research for the benefit of activities in the wider public arena, including environmental management, operational forecasting, industry, and commerce. Ocean remote-sensing research falls into both categories; the earlier chapters of this book show many examples of both classes of research while the rest of this chapter particularly addresses the emergence of important areas of applied research dependent on ocean remote sensing.

The reason for raising this subject at all is to reflect on the unfortunate tendency within the academic oceanographic community to rank achievements in pure research higher than those in applied research. It may not be a deliberate bias but it can certainly be detected, for example, in the way that journals publishing mostly applied research tend to be ranked somewhat lower than those which publish more

pure research results. At a time when it has become the practice within scientific management to emphasize the use of bibliometric indicators for evaluating scientific performance, this can lead insidiously to undervaluing scientists who spend more effort on applied research, even though their intellectual creativity may be just as strong as those engaged in more "fashionable" topics of pure research. To draw attention to this unfairness is not to undermine the importance of striving for excellence in all scientific research, both pure and applied. On the contrary, if excellent applied research is to be rewarded and thus encouraged, it is important that excellence is measured with a wider set of criteria that avoids the careless use of bibliometric accounting which could risk promoting an unjustified bias against applied research activity.

Ironically, this is happening at a time when national, science-funding agencies are encouraging scientists to demonstrate the beneficial impact of their fundamental research. This makes a lot of sense, since governments who are the major funders of science need to be able to show taxpayers that their investment is for the benefit of the wider public and not just to support the sometimes narrower intellectual interests of the academic community. For this reason ocean scientists should welcome the opportunity which OFS and other operational oceanographic activities provide to demonstrate the societal benefits to be gained from the previous decade's pure research. Satellite oceanographers can take pride in presenting these as "payback" to the public for the investment of their taxes in research, an important consideration since the costs of space infrastructure have been high. Moreover, successful applications make the case for investment in pure research. Underpinning every operational system, innovative fundamental research is needed to maintain and improve operational ocean forecasts and is an essential element of all successful systems.

There is another reason it makes sense not to discourage ocean scientists from engaging in applied remote sensing. The more we are able to develop applications of genuine public good using ocean monitoring and forecasting systems, the more justification there is for governments to invest in the continuity of ocean-monitoring satellite programs such as GMES. In practice this benefits the science community, since even those at the "pure" end of the research spectrum now increasingly make use of ocean observations routinely delivered by satellites, although the science sector could not by itself afford to fund all those programs. By investing intellectual effort in developing operational oceanography systems, scientists are benefiting both the public good and their own interests.

It therefore seems important that our best scientists should be encouraged to engage in both pure and applied science, which in fact has been the norm for much of the history of modern science. But beyond the rather self-serving reasons outlined in the previous paragraphs, there is an even more compelling reason to encourage practitioners of fundamental scientific research in ocean science to become better connected to the development of real-world applications of that research. This is to do with the recruitment into the field and ongoing motivation and retention of the best young scientists. As both a researcher and an academic teacher, the author has encountered several new generations of scientists: undergraduates, postgraduate researchers, and young postdoctoral workers, especially those who have been capti-

vated by the subject of ocean remote sensing. Their vision of new scientific possibilities and applications is often what drives the next stages of research, not to mention refreshing the enthusiasm of their teachers! When young scientists commit to a project, a Ph.D. program, or a career in satellite oceanography, what motivates them? First, of course, is the attraction of an interesting topic with lots to learn and wide, open possibilities for new ways of advancing oceanographic science, coupled with the intellectual challenge of grappling with scientific and technological problems, often crossing the boundaries of traditional science disciplines. But what also seems to persuade many young scientists to enter this field instead of another is an aspiration that what they study should have a wider human benefit than simply providing a scientific problem on which to prove their own intellectual capacity.

The sense that "my scientific work has the potential to improve life for other people" can be a strong and healthy motivating factor for someone entering a research career. The desire to "make the world a better place" is often very strong in students, and hopefully it still survives in scientists approaching the end of their career! It is a noble ambition to serve one's fellow citizens by exploiting for the public good the scientific knowledge that one acquires. Unfortunately, if achievements in applied ocean remote sensing are not adequately recognized, we risk turning away those bright scientists who are also motivated by high ideals. It is therefore the author's intention in this chapter to demonstrate to a new generation of scientists that satellite oceanography has reached the stage where it offers challenges in applied as well as pure science.

14.2 INTEGRATED OCEAN-FORECASTING SYSTEMS

14.2.1 What is operational oceanography?

The question of whether, and how, satellite ocean data can meet the requirements demanded by operational marine applications is at the core of this chapter. First, it is important to clarify what is meant by the adjective "operational" in the context of marine environmental monitoring, management, and forecasting. Broadly speaking, across many fields including industrial management, commerce, military operations, transport logistics, and environmental forecasting, the operational tasks are those parts of a system which are essential if the whole system is to function normally and fulfill its purpose successfully. They are generally routine and repeated tasks, defined and executed according to a prescribed specification, but if they are not performed at the right time the system as a whole will malfunction. Typically, operational systems have multiple dependences on their outputs, and the extent to which the failure of an operational element will lead to a disastrous outcome will depend on the end purpose of the system and on how much resilience is built into system design.

In operational oceanography we are considering the delivery of measurements of ocean variables, or of quantified information about the ocean state, necessary for the safe and efficient undertaking of particular tasks that are of humanitarian, environmental, industrial, commercial, economic, or military importance. There are many

systems in which such data may be used, including management decisions about ship routes, the timing of offshore engineering operations, or canceling a yacht race because of forecast high seas, warning of impending toxic algal blooms, of coastal flooding, or of dangerous sea ice, or directing search and rescue operations. Weather and sea state forecasting are also important systems relying on satellite ocean data. Less time-critical, but equally important, outcomes of operational ocean systems would be archives of reliable global data of diverse kinds, needed for management, for scientific applications, or for monitoring climate change.

It is characteristic of operational data that they must be delivered regularly, on time, with an adequate spatial and temporal resolution to detect the variability of the ocean and over as wide an area of ocean as the system objective requires. They also need to have a known degree of accuracy. Although each separate measurement may have little individual impact on the system overall, there may be particular circumstances when it is critical. As ocean measurement technology evolved in the latter part of the 20th century, it became feasible to measure the ocean routinely and more comprehensively using *in situ* platforms such as buoys, drifters, floats, and gliders as well as ships, and exploiting the increasing reliability of satellite measurements. Thus oceanographers conceived the possibility of regularly measuring the global ocean so that up-to-date information would be available to those responsible for managing a wide range of human activity at sea. Their inspiration was the World Weather Watch (WWW) which was established in 1963 to provide an international framework of meteorological measurements and which now underpins all weather forecasting. Thirty years after the WWW was set up, the concept of "operational oceanography" began to take shape under the auspices of GOOS, the Global Ocean Observing System (Alverson, 2005, 2008; Alverson and Baker, 2006). While improved scientific knowledge for climate has been a strong motivation for global GOOS, regional ocean-observing systems established in the last 15 years under the GOOS umbrella have focused more specifically on promoting operational oceanography systems to meet specific ocean management needs.

Operational oceanography envisages regular monitoring of the ocean in order to be able to provide critical knowledge of the ocean state when needed in emergency situations. At its core is a computer-based capability for numerical ocean prediction (NOP); it should include an element of forecasting (to provide predictions of future ocean states), nowcasting (which allows interpolation to estimate the ocean state at locations where no measurements are currently available), and it should also contribute to establishing a long-term archive of historic data. The European Global Ocean Observing System (EuroGOOS), a body set up to promote such a system, states:[1]

> "Operational Oceanography can be defined as the activity of systematic and long-term routine measurements of the seas and oceans and atmosphere, and their

[1] See the EuroGOOS website at *http://www.eurogoos.org/index.php* Select "what is EuroGOOS" and then "Operational Oceanography".

rapid interpretation and dissemination. Important products derived from operational oceanography are:

- nowcasts providing the most usefully accurate description of the present state of the sea including living resources
- forecasts providing continuous forecasts of the future condition of the sea for as far ahead as possible
- hindcasts assembling long term data sets which will provide data for description of past states, and time series showing trends and changes.

Operational Oceanography proceeds usually, but not always, by the rapid transmission of observational data to data assimilation centres. There, powerful computers using numerical forecasting models process the data. The outputs from the models are used to generate data products, often through intermediary value-adding organizations.''

In this definition it is worth noting the emphasis on rapid interpretation, transmission, and dissemination of data. Although the assembly of hindcast data for scientific and climate purposes is not time-critical, nowcasting and forecasting need measurements to be delivered in near-real time if they are to serve the needs of urgent management decisions. ''Near-real time'' is a somewhat loose phrase, but typically means within 3 to 24 hours of observation, varying between systems and according to the particular operational issue being addressed.

One other important aspect of ocean-observing systems is that they should be generic and holistic, offering a broad range of observations and information about the current state of the ocean, on behalf of a wide range of clients. Hitherto the collection of ocean data for operational needs has tended to be fragmented across different types of data (e.g., wave measurements, SST records, and algal bloom sampling being performed in isolation from one another) and duplicated across and within user sectors so that separate measurement services have been established for different client organizations. Thus different government departments, the navy, commercial shipping interests, the fishing industry, offshore mineral extraction companies, etc. have in the past arranged to meet their own particular requirements for critical information by individually procuring their own separate monitoring service, even though there may be considerable overlap with the data requirements of other user sectors. Because much ocean-monitoring work is contracted to commercial companies there has been an understandable reluctance to pool measurements and make them available to other marine user organizations. The evident inefficiency of such a situation is compounded when several countries individually address their own needs for observations of almost identical sea areas.

Recognition that such a fragmented system for operational ocean measurements needs to be improved has provided the political incentive to establish more co-ordinated marine-monitoring services, making use of the emerging concept of an integrated ocean-observing system (OOS). Within the European Union, this has taken the form of marine core services (MCSs), one of the first main elements of the GMES initiative to come to fruition. A single OOS is being established for all

European seas and the adjacent oceans. The NOP component of the MCS uses nested ocean-forecasting models at global to regional scales. Fed by satellite and *in situ* observations in near-real time, it generates basic ocean data representing the current ocean state, updated day by day. This is intended to serve the needs of all users of regional seas for quantitative information about the basic variables that describe the physical state of the sea and its biogeochemical contents. Following the European Union's principle of subsidiarity, it delivers core data free of charge as a public good, like basic weather services. The designation of "core" includes global data at 5 km to 10 km resolution and shelf seas data at 1 km to 2 km, which are of interest to a very wide range of downstream users, but does not in general include coastal and estuarine high-resolution data which typically have a narrow and more specific user base. The general provision of core data should remove the need for individual nations or user sectors to maintain their own separate monitoring service. In practice previously established national services making *in situ* measurements, and satellite data from the European Space Agency and Eumetsat will provide much of the "upstream" data supply into the OOS. This concept is illustrated in Figure 14.1.

Sectors with specialist needs to analyze and interpret core data will be served by specialist "downstream" services, such as ship routing or algal bloom monitoring, which "add value" to core data by interpreting them to clients and in some cases converting them into specialist data products. Downstream services will be provided by a mixture of fully commercial specialist consultancies and publicly supported marine management agencies, whose work may be contracted out to private companies. It is intended that the MCS will serve the governments of individual nations by supporting them in meeting their obligations under international marine environmental and pollution treaties to monitor conditions in their own local seas. The scientific and operational principles of the MCS were developed and tested through the European Union's MERSEA research project (Brasseur *et al.*, 2005). The prototype integrated NOP component of the MCS is now being delivered by a consortium called MyOcean which combines scientific (research and operational) expertise from many European public organizations and private companies working in the fields of ocean measuring, remote sensing, and numerical modeling.

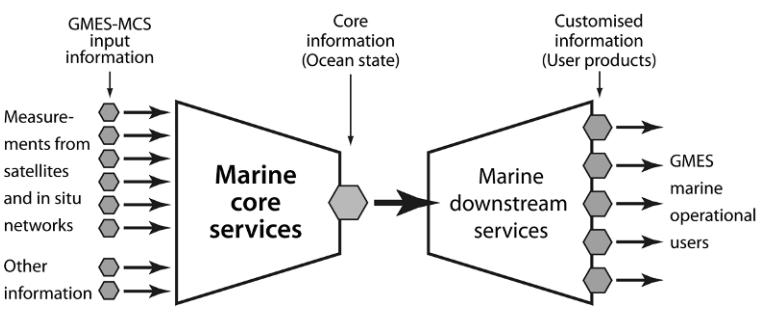


Figure 14.1. Schematic of the GMES Marine Core Service, showing its scope and its role for assimilating satellite and *in situ* observations from several suppliers and feeding integrated ocean information to end users, sometimes through downstream value-adding agencies.

14.2.2 Combining satellite oceanography and ocean models for operational tasks

Although it would be hard to contemplate operational oceanography in the absence of any satellite ocean sensors, in many applications numerical ocean models must also be used to facilitate the contribution of satellite ocean data products. In order to explain why satellite data alone are not usually adequate for operational tasks and why ocean numerical models have a role to play, it is worth outlining the way that delivery of ocean data products has evolved over the last 35 years.

In the early stages of satellite oceanography, the space agencies would process data to level 0 and level 1 (see Figure 2.8 and Section 2.3) but then leave it to scientific users to convert these raw data into estimates of useful ocean variables. During the 1980s and 1990s, as reliable algorithms were developed for atmospheric correction and to retrieve ocean variables, the agencies increasingly took responsibility for producing level 2 data products, and then for generating composite, level 3 products.

As the quality of the products stabilized and it became possible to quantify their errors, the agencies gained sufficient confidence to encourage applications of ocean data products in operational tasks. In practice, it was applications of SST to weather forecasting and of radar sensors to sea state forecasting, that matured quite quickly because meteorological agencies were already experienced in using near-real time satellite data for their numerical weather prediction (NWP) systems. Applications of altimetry and SST data to ocean circulation models have developed more recently (see Section 14.2.3). The least operationally developed field of application is the use of satellite ocean color data, which still has a long way to go before it meets the operational needs of those with a responsibility to manage marine and coastal ecosystems and resources, and to mitigate damage from natural hazards and pollution (as discussed in Section 14.3).

With the growth of possibilities for applied use of satellite data, the priorities of the space agencies shifted from a focus purely on sensor technology. For example, during the planning in the late 1980s of payloads for major Earth-observing platforms, such as ADEOS, Terra, Aqua, and Envisat, the agencies were also concerned to meet the needs of the users of intended ocean data products. New sensors for satellite oceanography were proposed, assessed, and justified not only to promote new technology but also in relation to the intended benefits of new ocean data products. By the turn of the century, the requirements of operational users had become the drivers of space agency procedures for designing sensors and data delivery systems. Applications dictated sensor design and choice of platform.

Having embraced the importance of user requirements, the space agencies at first tended to promote satellite measurements as the unique answer to the challenge of ocean monitoring. Little thought was given to the role of *in situ* observations, other than for validating satellite-derived products. Ocean models were recognized to be users of satellite data products, but not at first seriously acknowledged as having a part to play in generating those products. Neither was the idea of blending data from different satellite sensors uppermost in the minds of sensor specialists. New sensors were proposed in order to meet a particular ocean measurement need, such as

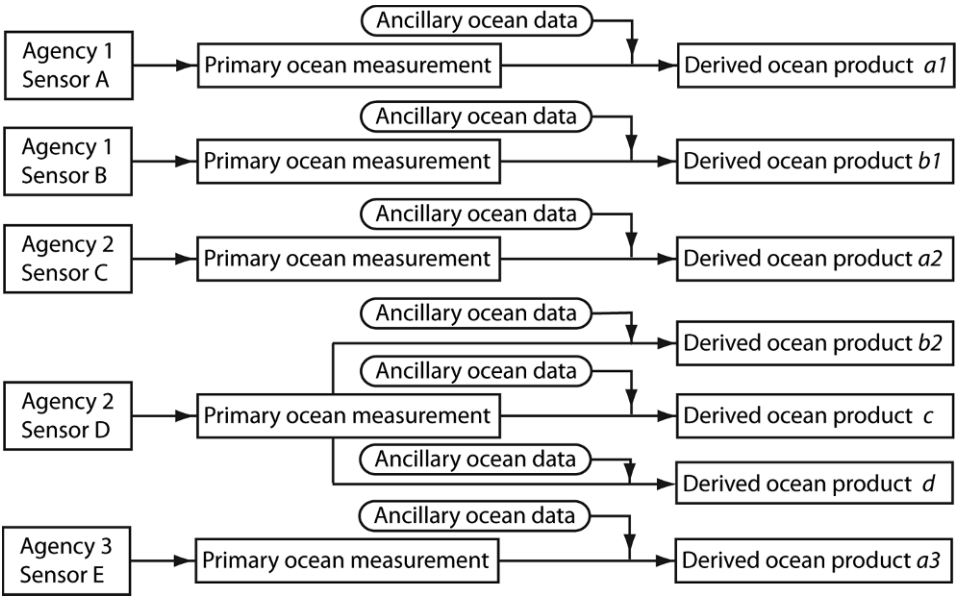


Figure 14.2. Schematic representing the narrow, direct approach of deriving a particular ocean product (*a*, *b*, etc.) from a specific sensor (A, B, etc.) of a particular space agency (1, 2, 3, etc.), independently repeated many times for different products. "Primary ocean measurement" corresponds to images of water-leaving radiance (color sensors), brightness temperature (IR or MW radiometers), surface roughness (radars), or surface height (altimeters). "Derived ocean product" could include such diverse data as a time series of global SST distribution or surface ocean currents, a map of internal wave patterns, the probability of the existence of a harmful algal bloom, detected oil spills, sea ice concentration, or diffuse light attenuation coefficient. Different agencies may produce different versions (*a1*, *a2*, etc.) of products given the same name but sometimes defined differently.

measuring ocean circulation or detecting harmful algal blooms, which it was supposed could be derived from the sensor's primary measurements. Figure 14.2 illustrates this approach schematically. A list of intended derivable products would be attached to the proposal for a new sensor, providing welcome evidence that space agencies now took user requirements seriously. However, it is questionable whether such a linear approach can work effectively in the complex environment of ocean remote sensing. First, it does not acknowledge that most ocean properties needed for operational oceanography must be derived from the sensor's primary ocean measurement by some inversion procedure that requires additional information about other aspects of ocean state (the ancillary data shown in Figure 14.2). Second, it ignores the reality that the frequency of sampling required for operational applications is almost always higher than can realistically be achieved from satellites without operating multiple satellites as part of a dedicated constellation.

Considering the capacity of satellite systems alone to measure an ocean variable, the most obvious weakness is that satellites can typically measure only at the sea

surface. The demands of operational oceanography require knowledge of the ocean state and how it varies in depth as well as horizontal space and time. Moreover, most useful ocean properties are not directly measurable by remote sensing. For example, if ocean color retrievals of near-surface chlorophyll concentration are to be reliable, more must be known about the relative proportions of phytoplankton, mineral suspended sediments, and land-derived dissolved organic material under Case 2 conditions. It is even harder to go a step further to provide operational warnings of harmful blooms, or to quantify the risk of anoxic conditions, which marine environmental managers are calling for. It is evident that, by itself, ocean color measurement is limited in how it can be interpreted to deliver quantitative measures, whereas if its interpretation can be combined with other observations the utility of ocean color data will be greatly enhanced. The challenge of how to provide other relevant data, which may come from *in situ* measurements or may be based on biogeochemical modeling, led to the realization that the most effective way to exploit satellite observations is in the context of a much wider system, embracing measurements from many sources and a continuously updated model that represents the best estimate of the ocean state at any instant. The same is true for many other ocean variables needed by an operational oceanography system.

Considering the sampling capability of satellite sensors, the requirement for daily, and preferably subdaily, samples is very hard to meet. The physical laws constraining satellite orbits impose constraints on when and where platforms can be placed (see section 3.2 of *MTOFS*—Robinson, 2004). Without incurring the high costs of launching several satellites in complementary orbits, it is not feasible to deliver high-resolution, Level 2, global ocean data products several times a day from a single sensor. Where several agencies were attempting to produce their own versions of regularly mapped products, such as SST, SSHA, or chlorophyll, it became evident that combining data from different agencies and sensors could improve overall sampling frequency. Unfortunately, for visible and infrared sensors the obstacle of cloud cover creates a sampling gap that even unlimited numbers of platforms would not be able to overcome. The reality is that for most types of ocean variables it is not logical to pursue a goal of generating purely satellite data products, at least in the context of operational applications. The paradigm implicit in Figure 14.2 needs to be replaced.

Figure 14.3 outlines the alternative which is now more widely accepted as the way forward. Instead of feeding primary measurements from satellite sensors into the linear processing chains of Figure 14.2 to generate purely "satellite-derived" data products for operational users, they are incorporated into a numerical, model-based representation of the contemporaneous ocean state. In general, operational users can be pointed to the model-based ocean-observing system, often referred to as a numerical ocean prediction system (NOP), to find the best estimate of ocean state at a particular location and point in time. This can only be justified provided the ocean model used for nowcasts and forecasts is forced by the same factors (e.g., wind stress, solar radiation) that drive the real ocean, and that it is effectively constrained by assimilation of contemporaneous satellite data products and *in situ* observations so that the model is not allowed to diverge from reality. A major benefit is that,

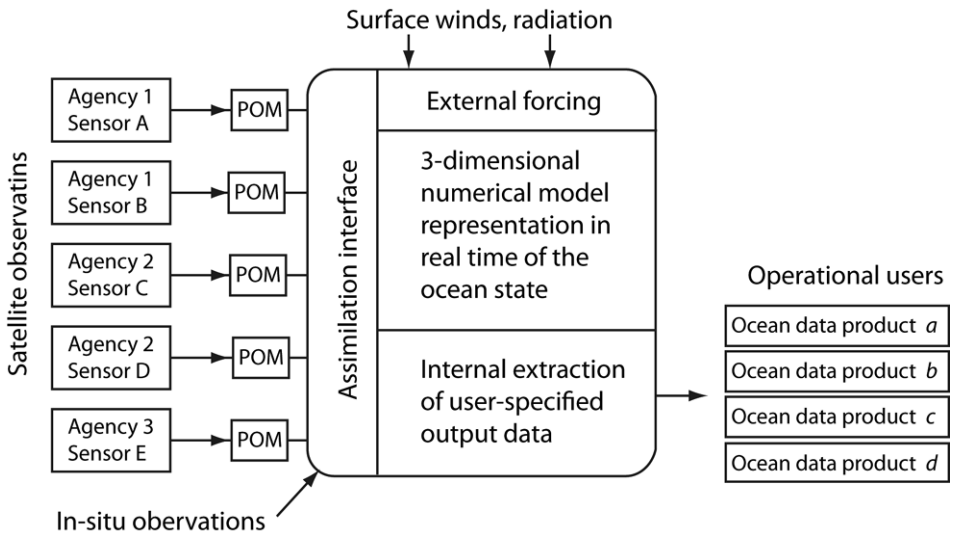


Figure 14.3. Schematic of a model-based approach in which primary ocean measurements (POMs) from satellite sensors are fed into a model system from which user-specified ocean data products are extracted for operational and other applications.

provided the model properly simulates the physical, dynamical, and other scientific processes that govern the real ocean, it allows the state of the ocean to be sampled at any location or time, irrespective of whether direct observations were made of that variable at that location. In principle, it provides a means of ensuring consistency between a variety of different types of observational input (thereby improving observational data quality) and offers a more efficient way of providing operational knowledge of the ocean required for near real–time management decisions.

Arguably, it offers by far the most effective means by which satellite ocean data can be made to serve the public good, delivering rich benefits to those whose livelihood comes from the sea and repaying the investment that has supported 40 years of developing ocean remote sensing. Nonetheless, providing observations to users through a model environment is a fundamental step for a remote-sensing agency to take. Observational marine scientists will tread cautiously before they finally accept output from a model system in preference to a direct measurement, whether *in situ* or remote sensing. There is still healthy skepticism between ocean-observing and ocean-modeling scientists. Yet the only way to retrieve specific information on some ocean state variables is from the NOP system.

The development of NOP systems for operational oceanography follows the same path that meteorologists trod some 25 years ago when they embraced numerical weather prediction, relying on the assimilation of observed atmospheric variables. It has only become feasible for the ocean because advances in ocean modeling have occurred in the last decade, in particular through the Global Ocean Data Assimilation Experiment (GODAE). This helped to develop the necessary

assimilation tools for the ocean, and proved the concepts of constraining global ocean circulation models to simulate the same mesoscale structures as the real ocean (Chassignet and Verron, 2006). As this book goes to press, the development stage has been reached where NOP systems have started to operate while being extensively validated. Through the GMES program, the space agencies within Europe are now committed to deliver satellite data needed for assimilation into the NOP systems of the Marine Core Service. Thus ESA and Eumetsat can start to think of NOP as a tool for refining their satellite observations and extending their usefulness. Soon oceanographers in general should be able to turn to NOP outputs to obtain data for scientific analysis, in the same way that meteorological scientists use NWP data.

Nonetheless, we should remain cautious. Although it appears that ocean dynamics can be successfully simulated by NOP models at several scales (as discussed in the next subsection), it is less clear how readily the complexities of ocean biogeochemistry will be modeled and constrained by observations (as discussed in Section 14.3). Moreover, although the adoption of the NOP concept seems to imply that any observational data can be fed into the model melting pot to make a contribution, it has already been discovered that a lot needs to be done by remote-sensing specialists to prepare data on SSHA, SST, or ocean color in order that they can be assimilated more effectively by NOP. This is the topic covered by Section 14.4. Neither should we lose sight of the need to maintain a pure satellite observation record of certain variables in order to allow independent validation of model output to be performed, and also to provide the basis for the model-independent climate records discussed in Section 14.6. Finally, no matter how valuable a tool NOP becomes for the application of observational data, oceanographers must never give up the opportunity for direct examination and interpretation of satellite and *in situ* observations, where scientific intuition and insight derived from experience can never be completely replaced by computer models.

14.2.3 Assimilating satellite data into ocean-dynamical models

The subjects of ocean modeling and data assimilation fall outside the scope of this book, but it is important for the satellite oceanographer to be aware of the issues involved in assimilating satellite data, especially if this is to become a route through which ocean remote-sensing data flow to meet the needs of operational applications. This section therefore outlines the concept of assimilation of satellite data into ocean-forecasting models, without attempting to explain the methods of ocean modeling (see, e.g., Griffies, 2006) or technicalities of different types of data assimilation, for which there is a growing literature (Brasseur, 2006).

Assimilation is a procedure used in ocean models to allow their predictions of the ocean to be compared with observations in such a way that the model can be adjusted to more closely resemble observations. One type of assimilation procedure uses a representative archived set of observations for the purpose of tuning model parameters so that the general behavior of the model is shown to be statistically similar to the real world. That is needed when models are being developed, but is not

the type of operational assimilation used in NOP models. Because the ocean is naturally turbulent, if a model has fine enough resolution it will spontaneously generate mesoscale variability. It is the stochastic nature of turbulence that, although wind and other forcing terms are the same, different runs of a model will generate different turbulent perturbations, and these will typically differ from each other and what happens in the actual ocean. Unless it is forced to follow what is happening in the actual ocean, the model will not otherwise be able to simulate the details of mesoscale eddies. Since these in fact represent the "weather of the ocean" it is important that they are faithfully reproduced in NOP used for operational applications. Therefore NOP needs to use the procedure of sequential assimilation in which theoretical, short-range forecasts of the model are systematically compared with what is measured in the real ocean, in order to constrain the model in subsequent runs to develop mesoscale perturbations that closely resemble reality. Observations are said to constrain the trajectory of the model through data assimilation techniques.

The aim of this type of assimilation is to generate a more complete and accurate description of the state of the modeled system than that obtained by either observations or model simulations alone.[2] The rationale behind that assertion is that the model generates fields of dynamically dependent variables such as temperature, salinity, density, etc. defining the ocean state, which should be self-consistent to the extent that the model code properly represents the physical processes, although it may not correspond to the true ocean state. Observations, if made with sufficient accuracy, can provide more reliable knowledge of the ocean state at a number of specific locations and times, but are not able to describe ocean state at other times and places. In principle the observations provide accuracy and the model offers completeness. The model steps forward in time by itself to predict the evolution of ocean state, responding only to prescribed external forcing factors, such as wind stress, which drive the model. Then, after an interval called the assimilation cycle, which for NOP is typically 6 hours, the assimilation procedure is performed in order to reset or "initialize" the model by adjusting the ocean state closer to what has been measured by observations during that assimilation cycle.

Figure 14.4 shows this schematically. The "first guess" is the model's prediction of ocean state variables at a given instant, and the "observations" are an array of measurements of state variables. "Objective Analysis" is the core of the assimilation procedure which compares these two inputs and seeks to reduce the differences between them in a consistent way. The new initialization that results from objective analysis represents the best description of present ocean state. This is the nowcast. The model is then run forward from initialization to make future predictions up to several days ahead. The prediction for a short period ahead (typically 6 hours ahead) forms the first guess for the next cycle of assimilation, which is compared with observations that are acquired during that intervening 6-hour period. Bell *et al.* (2000) describe an early example of an ocean-forecasting model intended for operational application and assimilating various types of data. The literature now reports

[2] See the GODAE website at *http://www.godae.org/Data-Assimilation.html*

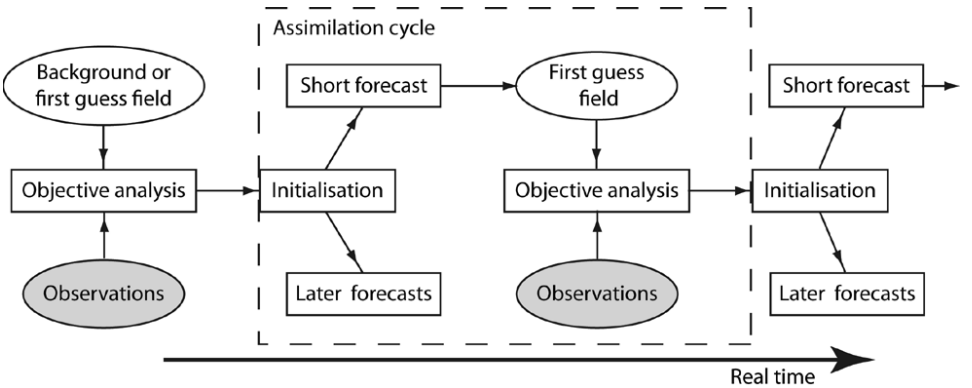


Figure 14.4. Schematic of sequential assimilation scheme for physical variables in an ocean general circulation model.

a variety of different ocean-forecasting models designed for assimilation of observations (e.g., Chassignet *et al*., 2007).

There is a variety of different techniques for performing the objective analysis component of the assimilation cycle, such as Kalman filters and variational methods, but these need not concern us here. For the ocean-observing scientist it is important to note those factors that promote an effective confrontation between observations and models, leading to more accurate nowcasts and forecasts.

First, we should note that an NOP model typically describes ocean variable fields over a large ocean region in three dimensions. The more representative the observational data are of the whole domain, the better we can expect the model to be constrained. Satellite observations provide an efficient and effective way to cover the ocean surface, but there is evidently an equally essential requirement for subsurface measurements. This is why the global array of Argo floats, with other subsurface observations from buoys and gliders, are equally necessary for NOP to be reliable.

Second, it is important to identify which observable variables are most effective at constraining modeled ocean dynamics. For operational applications we are more interested in making sure the model gets relatively short-term mesoscale variability right. Thus being able to regularly update surface currents, through SSHA measurements from altimetry, is a fundamental requirement without which NOP is ineffective. SST can also provide a good tracer of mesoscale dynamics, and also helps to constrain the flow of thermal energy through the sea surface. NOP also relies on radar measurements of wind fields, but note that these are part of model forcing rather than the assimilation process. In practice, forcing fields are typically provided by an NWP system that normally assimilates SST and surface wind speed measurements.

Third, it is useful to note that different variables need to be treated in different ways in the assimilation process. It may seem straightforward and effective to simply nudge a model variable towards the observed value at each assimilation cycle, and this may make for an acceptable nowcast. However, assimilation that does not

respect conservation of integral properties of state variables could distort the ongoing evolution of the model, by upsetting the internal equilibrium of forces, or denying continuity conditions and thus pointing the model forward in the wrong direction. For example, if surface temperature is arbitrarily increased to match observations, but subsurface temperatures are not changed, this could falsely stabilize vertical stratification and reduce vertical mixing in the model. It would also breach physical laws of thermal energy conservation. Comparable problems can arise when assimilating other variables. It is advisable to devise variable adjustment schemes which are conservative and maintain consistency with internal model physics. Careless assimilation can introduce more errors and uncertainty into the model than it reduces.

Fourth, it is important that error bars can be assigned with confidence to all observational data provided for NOP assimilation schemes. Put at its simplest, if uncertainty in an observation exceeds the difference between that observation and the initial guess, there is nothing to be gained by allowing it to influence the model at all. To nudge the model closer to a bad measurement can inject additional error into the model. Typical assimilation schemes assess the model's internal uncertainties for comparison with observation uncertainties, to decide how much weight to give to each in objective analysis. If observational errors are overestimated it may result in the model completely ignoring assimilation data, although this may not be very apparent to those running the model. Thus reliable error statistics are just as important for satellite data as keeping errors as low as possible.

Fifth, the archive of satellite observations provides a good resource for estimating the characteristic variability in space and time of the field variable used in assimilation. This is expressed in what is called the *background error covariance matrix* which itself may vary with location and day of the year. The smaller the observational errors are, compared with the background error, the more influence the observations are allowed to have on the analysis. Reliable knowledge of the background error covariance effectively controls the radius of influence of an isolated observational point during the assimilation process. If the covariance is anisotropic (different in different directions) it also controls the directions along which the effect of observations spreads most strongly through the analysis field.

These aspects of ocean assimilation must be kept in mind as satellite oceanographers start to prepare datasets for use in operational oceanography systems. Simply placing satellite data on ftp sites for modelers to download is not enough. Ocean remote sensors should engage scientifically with modelers in order to understand how satellite data can best be used to constrain models. There is a need for ocean-observing scientists to gain more knowledge of the details of assimilation than have been presented here. A good place to start is to read the proceedings from summer schools on issues of ocean assimilation (Chassignet and Verron, 2006). Equally modelers need to grasp what are the strengths and limitations of satellite measurements, if a partnership in assimilation is to be forged between observers and modelers that benefits both fields of science and leads to strong operational applications based on both satellite data and ocean-forecasting models.

14.3 ECOSYSTEM MODELING

14.3.1 How can satellite ocean color data support operational applications?

Moving from the previous section's outline of how physical variables are assimilated into ocean circulation models to consider whether something similar can be done with satellite ocean color data, we find exciting opportunities but at the same time are faced by considerable challenges.

Previous chapters presented evidence of valuable scientific applications of ocean color data (e.g., as a tracer for ocean mesoscale dynamics, the signature of upwelling events, and most importantly as the basis for providing ocean biologists with a more complete view of global distribution of primary production). However, when it comes to the more operational applications of ocean color, such as monitoring water quality or warning of harmful algal blooms in coastal seas, the theoretical potential has not yet been confirmed by consistent demonstrations of reliable operational systems. The discussion in Section 13.2.6 noted that the high probability of cloud cover in many locations means that satellite ocean color data cannot be relied on, by themselves, to fulfill a near real–time operational monitoring role. Neither can satellite data reveal biogeochemical processes occurring at depths below the upper photic zone. In principle these problems can be mitigated by assimilating chlorophyll concentration data into ecosystem models embedded into an NOP dynamical model. This is the common objective of those around the world developing integrated ocean-forecasting systems, such as the MyOcean project in Europe. However, for a number of reasons it is proving harder to achieve success with ocean color assimilation than with SSHA or SST.

The question of how to confront marine ecosystem models with satellite ocean color data is more complex than the equivalent physical oceanography problem. Although we do not yet have all the answers, this section aims to introduce the reader to the issues involved in what is one of today's most challenging fields for applied satellite oceanography research. It first outlines the basic principles of marine ecosystem modeling. It explains four distinct ways in which satellite chlorophyll observations can be used to improve ecosystem models and also discusses how ocean color data can be used directly to define the optical environment which determines the light level in an ecosystem model. Finally it considers alternative ways in which assimilation of ocean color data could be performed.

14.3.2 Marine ecosystem models, scientific principles, and operational purpose

At the core of ecosystem models is the concept of a single well-mixed tank of seawater containing chemicals, plants, and animals, illuminated periodically by sunlight. This is represented mathematically by a set of chemical-engineering process equations which simulate interactions between different components as they are observed to occur in the ocean. For example, one equation defines how plants (phytoplankton cells) grow by photosynthesis in sunlight, converting certain chemicals in solution into plant biomass.

Because the real ocean is not a simple tank but an extensive fluid, biogeochemical constituents must be treated as variables that are advected and diffused as tracers within a three-dimensional ocean circulation model. At each time step of the ecosystem model, biogeochemical interactions over the time interval are calculated separately for every grid cell of the ocean circulation model, treating it locally as a well-mixed tank. Ecosystem components are then redistributed by the circulation model before the next calculation of biogeochemical processes. It is evident that there is a massive amount of additional computation required when an ecosystem module is embedded into the basic ocean circulation model.

The biogeochemical process model itself can be constructed with varying amounts of complexity, depending on the number of independent state variables used to represent the various ecosystem components. Here for simplicity we outline the simplest and the most complex types, two extremes of a continuum of models with differing complexity. The final part of this subsection outlines the types of operational tasks intended for such ecosystem models within an NOP system.

Simple open-ocean models with few biogeochemical partitions

For application in the open ocean, relatively simple, pelagic ecosystem models are used, based largely on the approach pioneered by Fasham *et al.* (1990) and Fasham (1993). In this the ecosystem is separated into a small number of separate compartments. Using nitrogen as the basic currency of the model, each compartment is defined in terms of the amount of nitrogen per unit volume (mmol N m^{-3}). The seven compartments of a typical model are phytoplankton (P), zooplankton (Z), bacteria (B), nitrate, ammonium, dissolved organic nitrogen (DON), and detritus as nonliving particulate organic nitrogen (PON). Model equations represent the applicable biological processes within a single model element for P, Z, and B, such as primary production, respiration, grazing, excretion, and mortality. There are many model variants involving more or fewer than seven nitrogen pools. In some cases additional variables are introduced to represent carbon partitioning and alkalinity (Drange, 1996). Such models are represented schematically by boxes corresponding to separate nitrogen pools, and joined by arrows, each of which relates to an equation defining a flux between the two pools. The example shown in Figure 14.5 (Hemmings *et al.*, 2008) is for a four-compartment NPZD model where N is inorganic nitrogen, P is phytoplankton. Z is zooplankton, and D is detritus.

When the ecosystem model is embedded within a three-dimensional ocean circulation model, additional equations represent the transport of different ecosystem components between model grid cells by advection, diffusion, or sinking under gravity. A variety of different basin-scale models have been developed, using slightly different ecosystem and chemical models attached to different types of ocean circulation models, some with fine enough spatial resolution to resolve mesoscale eddies. They have been created for the North Atlantic (Gunson *et al.*, 1999; Oschlies *et al.*, 2000), the tropical Pacific (Christian *et al.*, 2002), and the World Ocean (Gregg, 2001; Palmer and Totterdell, 2001).

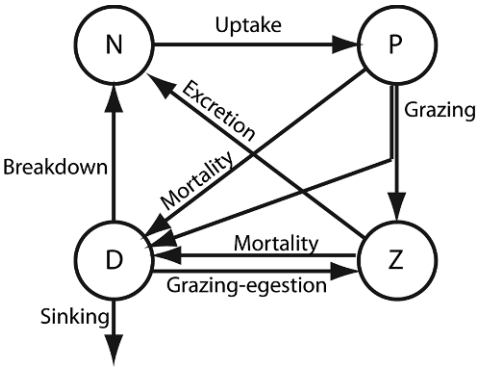


Figure 14.5. Dominant inter-compartmental nitrogen flows in a four-compartment, NPZD, ecosystem model. The compartments representing separate nitrogen pools are dissolved inorganic nitrogen (N), phytoplankton (P), herbivorous zooplankton (Z), and detritus (D) (based on figure 1 of Hemmings *et al.*, 2008).

Complex ecosystem models for shelf seas

In shelf seas, where there is greater heterogeneity of the pelagic ecosystem over shorter lengthscales and where benthic processes must also be taken into account, modelers have used more complex designs, such as the European Regional Seas Ecosystem Model (ERSEM) (Baretta *et al.*, 1995), which has been applied to the North Sea (Siddorn *et al.*, 2007), the Mediterranean Sea, the Adriatic Sea, and the Arabian Sea. The marine ecosystem is represented by a network of physical, chemical, and biological processes organized within pelagic and benthic components.

In these more complex models, biota are necessarily segregated into functional groups that are intended to represent particular types of behavior rather than species lists. Each functional group is defined by a number of explicitly modeled components: carbon, nitrogen, and phosphorus and, in the case of diatoms, silicon. In ERSEM, phytoplankton are represented by four functional types: picophytoplankton (0.2–2 μm), small autotrophic flagellates (2–20 μm), large autotrophic flagellates (20–200 μm), and diatoms (20–200 μm). When embedded into a three-dimensional physical model of a shelf sea, there are 36 pelagic state variables to be advected and diffused by hydrodynamics.

A debate continues within the ecosystem modeling community about the relative merits of simple and complex approaches (Anderson, 2005). It might seem obvious that a model with more variables would be better able to represent the rich diversity of the real ocean, but the purpose of models is also to forecast how an ecosystem will develop in future. The higher number of degrees of freedom, which allows a complex model to be better fitted to a known reality, also seriously reduces the model's predictive capacity. The simpler models are able to predict quite well how the balance between broad ecosystem components (nutrients, phytoplankton, zooplankton, etc.) will progress, even though they are unable to distinguish between different functional groups that make up those components. There is value in both approaches, as long as their individual shortcomings are recognized. It is also worth noting that ecosystem models embedded in ocean circulation models are sensitive to physical forcing (Friedrichs *et al.*, 2006). It is prudent to make sure the physical constraints on such a model are properly tuned before adjusting ecosystem complexity.

Models to meet operational needs

Some of the models referred to above were constructed primarily for scientific applications, in order to explore how ecosystems work. However, the interest of this chapter is in operational models that are run to simulate how the real ocean ecosystem is behaving, in order to provide critical management information. Take, for example, the operational responsibility of warning coastal holiday resorts about the occurrence of algal blooms which interfere with maritime leisure activities. Satellite data alone might be used for this task, but would be adequate only if daily cloud-free cover were available. Since this can rarely be guaranteed, authorities with operational responsibility are unable to rely exclusively on satellite data. In these circumstances the use of a model capable of predicting the evolution and movement of algal blooms, regularly updated by satellite observations when skies are clear, appears to have potential as an effective solution.

Tourism is only one of several sectors that need such integrated observing systems; others include the marine aquaculture industry, fisheries, and those regulatory authorities with a statutory responsibility for water quality monitoring. As international legislation imposes greater demands on nations to monitor the quality of the water in their own exclusive economic zones (EEZs), ocean-observing systems are increasingly seen as an essential tool, and it is evident that these must contain an ecosystem-modeling component if they are to provide the required management information.

The main difference between an ocean ecosystem model used for scientific analysis and one used for operational applications is found in the way it is operated, how outputs are delivered, and observations are prepared for the model. Operational models need to provide a best estimate of the present state of the ocean, and therefore only contemporaneous observational data are useful for constraining model state variables, whereas scientific models can afford to wait until the best observational data are available, allowing complex preprocessing of satellite data to be performed if necessary and allowing the use of *in situ* measurements that may be several days old when made available. The operational modeler does not have the luxury of being able to wait for such inputs. There is a tradeoff to be made between the reliability and timeliness of operational model forecast outputs. Thus it should be very much the concern of ocean forecasting to ensure that numerical ocean ecosystem models are provided not only with the most relevant information content from satellite ocean color data, but also that this is done as close to real time as technically possible.

14.3.3 Ways in which ocean color data are used in ocean modeling

Notwithstanding the discussion in the previous paragraph, not all observational data presented to operational models has to be supplied in near-real time to be useful. Some can be used in retrospect to validate the performance of model hindcasts, or to provide a climatology against which model output can be compared, including the development of background error covariance estimates (a measure of typical spatio-

temporal variability of chlorophyll in different locations and seasons). This subsection considers three distinctly different ways in which ecosystem models can be confronted with ocean color observations so as to improve model performance, but not explicitly to constrain state variables.

Parameter estimation in biogeochemical models

A critical challenge for all types of ecosystem models is to determine appropriate values for the parameters used in each model expression to represent a biological or chemical process. This requires a calibration procedure which optimizes model parameters in order to match the modeled state to an independent set of observations. Although local ecosystem models can be tuned for a particular place and season using a suite of *in situ* measurements of main ecosystem components, parameter sets determined in this way cannot be applied with confidence to a basin-scale model running throughout the year. In that case, chlorophyll data obtained from satellite ocean color can provide the global geographical coverage, spatial detail, and most importantly multiple annual cycles needed to represent the variety of conditions which such a model must be capable of simulating.

The use of satellite-derived chlorophyll data in this way for model parameter estimation is sometimes referred to as assimilation, although it must be clearly distinguished from sequential assimilation used to constrain operational models (as discussed in Section 14.3.4). In fact much ocean color assimilation literature to date has been concerned with parameter estimation to tune marine ecosystem models. For example, Hemmings *et al.* (2003, 2004) used SeaWiFS-derived chlorophyll data from locations across the North Atlantic to calibrate a zero-dimensional plankton ecosystem model (i.e., one that is not embedded within a circulation model but which attempts to describe the seasonal cycle separately at a number of different locations). The aim of these authors was to achieve this using a single set of parameters. They separated their SeaWiFS observations into two independent sets; using one for calibrating model parameters and the other for validation of resulting model performance. They concluded that derived parameter sets are improved by using satellite data because the high volume of observed data can supply examples of a wide range of possible biogeochemical responses to different physical conditions. For the same reason they could demonstrate that model performance is improved if the North Atlantic is divided into two provinces for which parameters are independently calibrated. However, they introduce a note of caution by commenting that despite the advantage of using satellite chlorophyll for calibrating the model, in this case the parameters remain poorly constrained unless other ecosystem elements are known, such as the distribution of nutrient concentrations over winter. Satellite data alone are inadequate to tune model parameterizations.

Validating model simulations

Another way in which satellite data can be used in relation to numerical marine ecosystem models is to compare model forecasts with satellite observations for validation purposes (analogous to how the operational ocean wave-forecasting com-

munity use satellite SAR and altimeter data for validating wave models). Note that the specific purpose of validation is not to improve the model, but to objectively quantify its performance by generating error statistics, the availability of which is essential if model results are to be used within an operational management framework. A model may be performing very well but, until validation error statistics are provided to support that claim, an operational manager would not be advised to use the model's predictions. Error statistics can be based on the extent to which model prediction agrees with the satellite-derived field of an appropriate relevant variable (e.g., the concentration of chlorophyll-*a*). To preserve the integrity of such a process, it is important that the same satellite data should not already have been used to tune model parameterization (as in the previous subsection) or for direct assimilation (as in Section 14.3.4), even though this disadvantages the model because not all available data are being used to constrain its trajectory.

Other difficulties can be envisaged with this approach which could lead to a model's performance being underrated and criticized inappropriately. When differences are detected between a modeled and an observed variable, they arise not only from a failure of the model to predict the true value, but also from the failure of the observation to represent the true value of the variable being modeled. For a start the observational measurement error may be relatively large (as discussed in Section 13.2.6) which straightaway imposes a level of uncertainty on the validation process.

Moreover, the observational sampling method may not allow comparison of like with like. Typically the modeled surface chlorophyll concentration represents an average concentration throughout the volume of a model grid cell in the surface layer. An *in situ* measurement typically represents a single point at a particular depth only. Any subcell-scale horizontal and vertical variability will lead to mismatches. Of course one of the advantages of satellite measurements over *in situ* observations of ocean color is that, like numerical models, they integrate in horizontal space, although their grid may not match the model grid. Satellite observations also, to some extent, integrate over depth, although the outcome is strongly weighted to surface values and they do not penetrate much below the depth at which solar illumination falls to about 1/3 of the surface value.

Consequently there remains uncertainty about how much error the horizontal patchiness of chlorophyll distribution, and its vertical profile, introduces to the attempt to match like with like, since it must depend on subscale processes which are neither measurable by satellites nor explicitly forecast by the model. While some general conclusions can be drawn from the assumption of a log-normal distribution of chlorophyll concentration (Campbell, 1995), there is little reported work on the magnitude of mismatch errors which may arise from this source. It is important to distinguish these as *sampling errors*, in contrast with *measurement errors* arising from the uncertainty of algorithms when retrieving chlorophyll concentration from water-leaving radiances, especially in Case 2 waters.

In the context of model validation, there should be little difficulty in ensuring coincidence in both space and time between a satellite observation and its corresponding model value, thus avoiding one of the pitfalls encountered when *in situ* measurements are used to validate satellite observations. There remains the question

of whether comparison should be made on a pixel-by-pixel, cell-by-cell basis or whether both satellite observations and the model field should be smoothed before any comparison is made. From a given cloud-free satellite overpass, there may be many thousands of cloud-free pixels available for comparison with the modeled variable field at a particular time. Should each pixel be treated as an independent measurement? Or should any fine-scale spatial variability be suppressed by smoothing because the model is not expected nor intended to reproduce that variability? If so what should the smoothing lengthscale be?

These questions relate to the issue of whether validation should require detailed agreement between satellite image and modeled field, or whether it should be based more on the model's ability to present the same broad patterns of distribution as satellite data. The matching and comparison methodology that is adopted could make a large difference to quantification of the mismatch between model and image data. For this type of application of satellite data to models, it may be more appropriate to use an analyzed field of chlorophyll derived from one or more overpasses rather than individual single-overpass data. In this way certain degrees of smoothing and quality control could be applied to data in order to reduce spurious mismatch errors that have nothing to do with numerical model performance.

Finally, the insistence on making comparisons only between temporally coincident model and satellite data could fail to do full justice to a model that is broadly successful but gets the timing of a bloom wrong by a few days. It would be beneficial for the comparison method to be capable of detecting lagged correlation between model and satellite data.

From this discussion it can be concluded that, although the principle of model–satellite data comparisons seems straightforward as a means for validating models of phytoplankton populations, in practice it requires further analysis and careful testing before it is used as a metric for the performance of ecosystem models.

Initializing forecasting/nowcasting model runs

Another way in which satellite data can be applied to support numerical modeling of ecosystems is to use them by occasionally re-starting operational biogeochemical models. To do this the chlorophyll distribution is derived from a largely cloud-free ocean color image of the modeled region, and is then entered as the starting values of model chlorophyll concentration. The ecosystem model, embedded in a physical circulation model, is then run forward in time to provide forecasts of the development of plankton blooms, and these may be used to issue warnings of harmful algae blooms (HABs) or the risk of eutrophication in vulnerable sea areas.

One example of how such a model makes use of satellite ocean color data is provided by Durand *et al*. (2002) in a case study describing operational monitoring for a HAB of *Chattonella* spp. This species started to bloom in Scandinavian waters in 1998, returning in 2000 and 2001, killing thousands of tons of farmed salmon along the southern Norwegian coast. The study describes real-time monitoring of the bloom in 2001 using SeaWiFS data.

At the heart of the monitoring process is an ecosystem model embedded in a hydrodynamic model of Norwegian coastal seas, relying on observational data inputs to achieve a good representation of the real ocean, but still able to provide estimates of ocean state at times and locations when no observations are available. Every cloud-free chlorophyll image is compared with model phytoplankton prediction. When a bloom event is detected by the satellite but is not predicted by the model, the model phytoplankton field is manually reinitialized to concentrations observed from the satellite image. The model then continues to predict further development, advection, and final decay of the bloom, and warnings are issued if the bloom approaches the coast. Subsequent satellite observations are used just for validation, and further reinitialization is performed only if the model deviates strongly from satellite estimates.

The model used in that study included as state variables the abundance of two functional groups, diatoms and flagellates, but did not specifically represent a particular species of harmful algae. Therefore when a bloom is detected from space, *in situ* samples are urgently needed to determine whether it is a harmful species or not. Despite the obvious shortcomings that HAB species are neither represented explicitly in the model nor directly distinguishable from non-harmful species in satellite ocean color data, this case study demonstrated the effectiveness of combining a model, satellites, and *in situ* sampling for critical operational tasks. Although in this example the way in which satellite data were used to confront the model was fairly crude, it was reasonably effective and adequate for the operational task. Such an approach would not be justified in a scientific study. By forcing phytoplankton concentration to match the satellite, without making compensatory adjustments to other related constituents of the model, biogeochemical behavior is being seriously distorted, and this is likely to make forward evolution of the model less reliable.

14.3.4 Sequential assimilation to constrain ecosystem state variables

Although the uses of satellite ocean color data, and particularly satellite-derived chlorophyll-*a* (discussed in Section 14.3.3), are sometimes referred to as assimilation techniques, they differ from assimilation of observations in order to constrain some of the ecosystem model's state variables directly. Each newly acquired field of satellite-derived chlorophyll observations is compared with model-predicted chlorophyll distribution for the same region at the same time. The assimilation process is then used to drive model predictions closer to observations (as outlined in Section 14.2.3 for assimilation of physical variables). This may be a simple nudging or relaxation process, or a more complex procedure may be applied, such as three-dimensional VAR or the ensemble Kalman filter (EnKF). The latter methods require reliable estimates of errors associated with the observed data field, preferably on a pixel-by-pixel basis. These "observation error" estimates should include both direct measurement errors and sampling errors (as discussed in Section 14.3.3). Also required is information about the autocorrelation lengthscale for the assimilated variable, appropriate for that region of the sea at that time of the year. This can be obtained from satellite ocean color–derived climatologies of chlorophyll distribu-

tion. It provides what is called the "background error covariance matrix" which indicates the extent to which chlorophyll distribution is typically expected to change. Only if observational errors are comparable with or smaller than the background error will the assimilation scheme allow observations to significantly influence model reinitialization during the assimilation procedure.

A number of models have been developed to include this type of assimilation. One of the first published demonstrations was by Natvik and Evensen (2003a, b) who developed techniques to assimilate SeaWiFS-derived chlorophyll data into a full three-dimensional ecosystem model embedded in a North Atlantic hydrodynamic model. This work used an ensemble Kalman filter (EnKF) approach to assimilation. They were able to demonstrate that satellite data measurably improved model response. Assimilation reduced variance fields for all the variables, including those partitions of the ecosystem that were not directly observed. They concluded that use of a multivariate analysis scheme enabled phytoplankton information supplied by satellite data to influence analyses for other ecosystem partitions beneficially, even though the observations of chlorophyll could be related to only one of the state variables (phytoplankton component).

It is worth emphasizing here, although it is equally true for the other types of ocean color assimilation mentioned in the previous section, that to allow assimilation to be performed the model must first generate a prediction of the same variable as that retrieved from the satellite. In most cases it is chlorophyll, but it could be suspended particulates or CDOM. In practice, many ecosystem models would not include any of these explicitly as state variables, so that an extra step must be added in the ecosystem model to produce what we can refer to as the "assimilation variable". To be effective, assimilation must constrain one or more of the model's state variables. If there is a well-defined correlation between the assimilation variable and a single-state variable (as, e.g., might be locally defined between chlorophyll concentration and a state variable corresponding to phytoplankton biomass), assimilation can proceed effectively, although even this example raises the question as to how certain we can be about the universality of the Chl–phytoplankton relationship. If the connection between the assimilation variable and the state variables is less well defined, or shared unequally across a number of state variables (as when several state variables consist of different size classes of phytoplankton), careful research must go into designing the assimilation scheme. The question to be answered is: "When model-predicted chlorophyll differs significantly from assimilated chlorophyll, how is (are) the state variable(s) to be adjusted to reduce that difference?"

A good example of addressing this issue can be found in Hemmings *et al.* (2008) who discusses ways in which to distribute assimilation changes across different partitions of the ecosystem model. In this particular case the model's purpose was to estimate the air–sea flux of CO_2, driven by the pCO_2 difference across the air–sea interface. As a plankton bloom develops it is expected that total dissolved inorganic carbon (DIC) should reduce as it is taken up into organic compartments, thus lowering the pCO_2. The main reason for assimilating ocean color data was so that the model would track phytoplankton growth effectively. In that context, the ocean

color assimilation scheme had to be designed so that it would correct not just phytoplankton biomass but all components of the system affecting DIC and/or alkalinity, without forcing relationships between the various state variables beyond what is naturally observed. In this example a computationally efficient scheme was created for balancing daily surface phytoplankton increments in an NPZD nitrogen cycle model.

It is evident that progress is being made towards operational sequential assimilation schemes for ecosystem models, but much of it is still based on one-dimensional model test beds and the stage has not been reached where ecosystem models can be considered reliable for nowcasting the state of the ecosystem. The most promising results are in the open ocean where chlorophyll concentration retrieved from satellites is fairly reliable. However, large-error estimates on Case 2 retrievals of chlorophyll make it difficult for assimilation schemes to give much weight to satellite observations in coastal and shelf seas.

14.3.5 Characterizing light penetration in numerical models

It is important not to overlook another way in which satellite ocean color data make a significant contribution to ecosystem modeling. This is by providing detailed information about the optical attenuation of light in the sea, normally represented by K_d, the diffuse attenuation coefficient. Knowledge of K_d is essential, first, for characterizing the penetration of light and hence the depth distribution of solar heating in physical/dynamical elements of numerical ocean models. This can strongly influence the vertical distribution of density, buoyancy, and hence vertical mixing in the water column, impacting on mixed layer depth and evolution of a diurnal thermocline. This in turn can have a significant impact on the behavior of the ecosystem when biogeochemical elements are added to the ocean model. Although K_d varies in space and time over the ocean, some models treat it as uniform while others rely on climatology of the geographical distribution of K_d and in some cases its seasonal distribution. However, it is now possible to make use of satellite-derived global fields of K_d, updated regularly, in order to supply physical/dynamical models with realistic estimates of K_d.

It should be noted that this application of ocean color data is simply to improve definition of a physical parameter within the model. K_d is not a state variable within a physical model, and so there is no question of using K_d to confront the model and adjust its predictions in one of the ways outlined in Sections 14.2.3, 14.3.3, and 14.3.4. Rather it can be seen as one of the drivers that force the model, providing refinement to the way in which solar heating is distributed with depth in the water column.

The second requirement for information about light penetration is to characterize the availability of light for photosynthesis. Photosynthetically available radiation (PAR) is an essential component of all ecosystem models, whether complex or simple. It is the driver for photosynthesis and controls the rate of primary production. PAR at the sea surface can be estimated climatologically, or derived

directly from satellite observations (see Section 7.3.3), but to estimate how it varies with depth requires knowledge of K_d.

From this it is clear that the use of satellites to update regularly the value of K_d used in models, and to provide the spatiotemporal distribution of PAR at the sea surface can make an important contribution to the performance of both physical and biogeochemical elements of ocean models.

14.3.6 Alternative approaches to ocean color assimilation

As a summary of this subsection on ecosystem models, Figure 14.6 provides a schematic representation of various ways in which it is perceived that satellite ocean color data can be used to influence ecosystem models. Satellite-derived water content concentrations, principally chlorophyll but also possibly suspended particulate material, may be assimilated into the model, although to do this requires the corresponding variable to be derived within the model from the model's state variables. Optical parameters such as K_d are used to set the model's optical attenuation conditions. In addition, if a model output variable corresponds to a satellite-derived variable, an analyzed field based on satellite data can be used for model validation.

However, probing beneath the surface of this simple data flow diagram reveals potential sources of error in assimilation data and uncertainty in various processes that must be performed to allow satellite data to confront model data. These are identified in Figure 14.7. Such errors need to be managed and minimized before the type of assimilation scheme that works for SSHA or SST in an ocean-dynamical model will permit chlorophyll data (with typical errors of 30% in the open ocean) to have any influence at all on the ecosystem model. Judging from the recent literature (referred to by Hemmings *et al.*, 2008), it seems that workable chlorophyll assimilation schemes have been achieved in the open ocean where Case 1 conditions reduce the uncertainty of satellite retrievals.

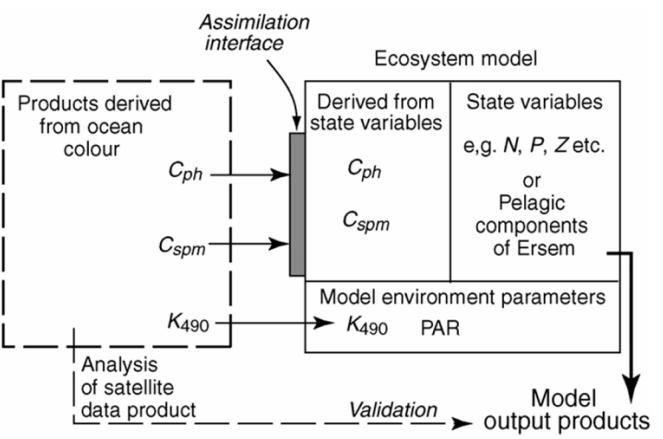


Figure 14.6. The conventional view of how satellite ocean color data can interface with an ecosystem model.

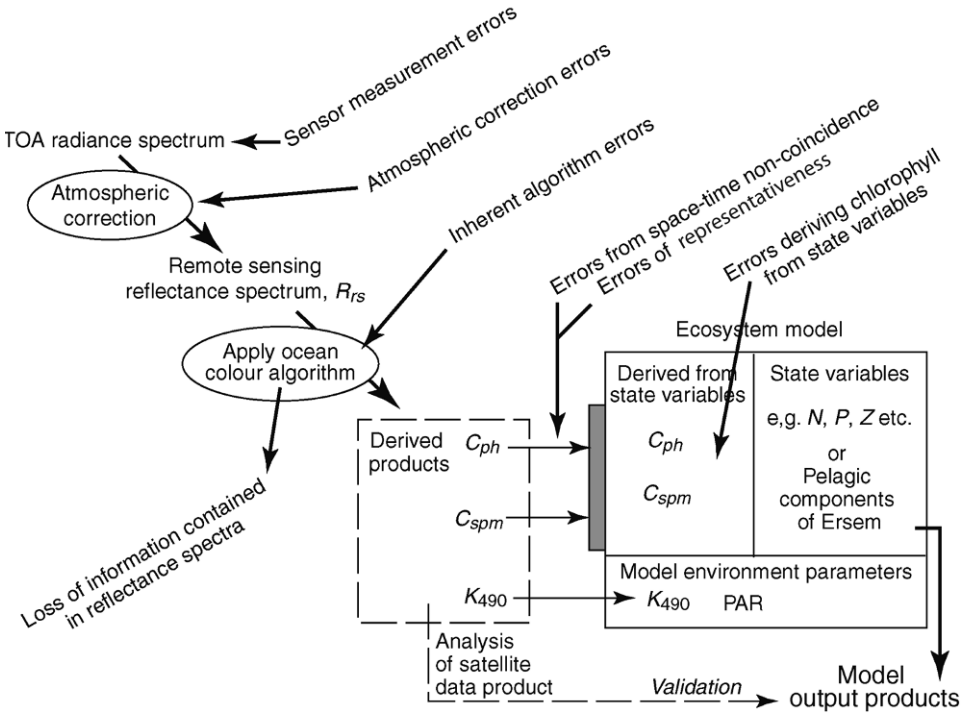


Figure 14.7. Sources of error and potential loss of information in the conventional assimilation of satellite-derived chlorophyll data.

In Case 2 waters the problems are compounded by additional optical constituents, causing higher errors in satellite retrievals (of order 100%) and creating more connections between satellite ocean color products and various ecosystem model state variables. While solutions may possibly be found in local seas if the optical characteristics are very well understood, the prospect of a universal assimilation approach for a variety of Case 2 waters seems highly improbable. This is unfortunate because the very places where precise nowcasting, if not forecasting, of algal blooms and the onset of eutrophication is most needed for critical operational applications are shelf seas where optical conditions are typically Case 2. This poses a real challenge to NOP for biogeochemical variables. Will we be able to serve the public need for water quality information at the point where it is most crucial, but also most difficult?

Faced with this challenge, it is worth recalling the rationale for the integrated ocean-forecasting systems presented in Section 14.2.2, that numerical models can help to mitigate limitations of observations and *vice versa*. Perhaps it is time to take some fresh approaches to the optical problems of Case 2 waters and the complexity of shelf sea ecosystems. Here are two suggestions among others that are being considered at present, although there is little published evidence yet of whether they will bring workable answers.

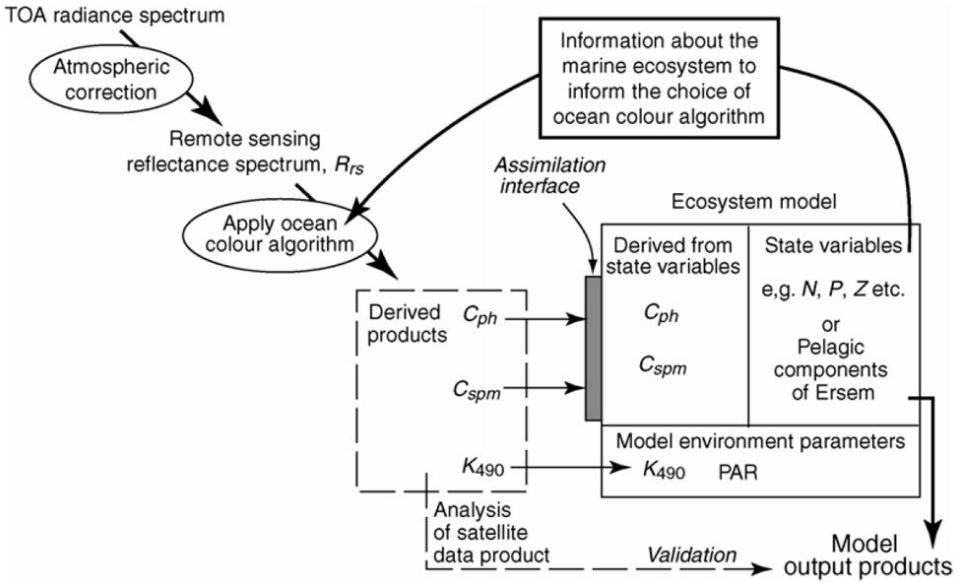


Figure 14.8. Schematic showing how ecosystem information can be fed back from the model to inform ocean color data-processing choices before derived products are assimilated into the model.

The first arises from noting that the model already contains time- and space-variable information about the properties of a particular ecosystem, such as the relative quantities of phytoplankton and mineral particles, or the balance between dissolved organic carbon and phytoplankton. Both of these could be crucial pieces of evidence in solving the Case 2 problem at a given location and time. The first indicates whether to interpret higher satellite-measured reflectance in the green as increased phytoplankton or increased suspended sediment. The second could be related to the relative blue light–absorbing effects of CDOM and chlorophyll. Why not use information from a mature ecosystem model to assess the character of the Case 2 conditions found at a particular place and time? From this assessment the system could select the most appropriate specialist Case 2 algorithm to use in deriving useful assimilation variables from satellite ocean color data. This changes the original information flow diagram of Figure 14.7 into the version shown in Figure 14.8. It shows some information flowing from the model back up the satellite processing chain to help resolve algorithm choice and thus reduce uncertainty. One attraction of this approach is that it overcomes a justifiable criticism of specialist Case 2 algorithms that they are not very helpful unless it is already known which of several possible choices is appropriate. The model may offer the extra information needed to make that choice.

The second suggestion is for a more radical look at the assimilation process and redefining the assimilation interface. Revisiting Figure 14.7, two apparently unconnected facts can be noted. The first is that a lot of satellite ocean color information is thrown away by the use of simple Case 1 algorithms, based on the ratio of two spectral bands, when in fact satellite reflectance contains much richer spectral in-

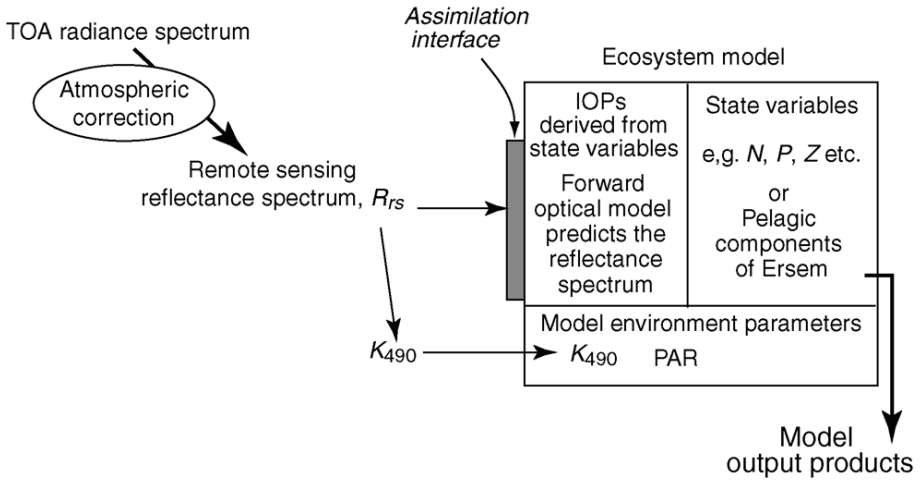


Figure 14.9. Schematic showing an alternative approach to assimilating ocean color data into an ecosystem model. If the ecosystem model incorporates an optical model which calculates the remote-sensing reflectance spectrum, R_{rs}, associated with state variables, then R_{rs} can be used as the assimilation variable.

formation. The second is that to achieve an interface between basic satellite measurement (which is the reflectance spectrum after atmospheric correction) and the model state variable(s) we have to convert both into a third variable (chlorophyll or suspended sediment concentration). This introduces errors from both sides to quantities that are compared at the assimilation interface, reducing the capacity for observations to constrain the model state effectively. So why not reposition the assimilation interface to somewhere else in the processing chain between raw satellite data and model state variables? An attractive alternative assimilation variable would be the water-leaving reflectance spectrum itself.

This is shown schematically in Figure 14.9. It immediately removes much of the uncertainty associated with algorithms used for retrieving variables such as chlorophyll from satellite data. It has the merit that the reflectance spectrum is a variable that is much more tightly definable, as a physical quantity, than the biogeochemical variables used in the present mainstream approach. Of course it introduces the completely new requirement that the ecosystem model must use internal information about the biogeochemical state of the local sea to predict the surface reflectance spectrum. It would require a forward optical model to be included as an additional module. But the physics of forward optical modeling is well understood. What holds back such an approach most is the lack of confident knowledge about inherent optical properties (IOPs) that characterize the optical behavior of water content (see section 6.2.6 of *MTOFS*—Robinson, 2004). However, improved measurement of IOPs and SIOPs, which relate IOPs to the concentration of optically influential water constituents, is already a fairly high priority among optical oceanographers seeking to improve Case 2 algorithms (Robinson *et al.*, 2008). Using the reflectance spectrum as the assimilation variable offers a multichannel comparison (tailored each time to match the spectrum of the particular ocean color sensor providing

observations). This holds out the potential for increased information flow across the assimilation interface and therefore better model constraint. Nonetheless, it is fair to warn the reader that this suggestion is speculative. In contrast with the more mature ideas, information, and knowledge presented in the rest of this book, the effectiveness of assimilating spectral reflectance has not yet been properly tested and awaits evaluation by peer review.

The application of ocean color remote-sensing data to improve the reliability of operational marine ecosystem models remains a field of ongoing research, one that is both intellectually and technically challenging and where success is eagerly awaited by those with responsibility for managing shelf sea water quality and the ocean's living resources.

14.4 PREPARING SATELLITE DATA FOR OPERATIONAL USE

During the first 30 years of satellite oceanography's development, the major limiting factor in determining how readily a new type of remote-sensing observation would be exploited by the scientific community was ease of access to data. That is now changing. Difficulties that preoccupied early satellite oceanographers, such as the cost of acquiring data, the medium of transmission, software to read special file formats, etc. have largely been swept away by widespread use of the Internet. It is taken for granted that most agencies producing ocean data products from their satellites and sensors will provide online access which is generally free. Sometimes commercial users are charged but increasingly most satellite data products, apart from very high–resolution mapping images, are treated as core ocean information provided by national or international agencies for the common good (as outlined in Section 14.2.1). This is good news for ocean scientists.

However, just because data are available does not necessarily mean that they will be used for operational applications. This section of the chapter considers the factors that determine whether datasets are considered fit to be taken up readily for applications or are ignored by the user community. The first part explores generic issues, particularly in relation to how similar data products from different satellites and agencies are combined or harmonized to make them more useful for operational purposes. The second part provides a case study of how application to operational tasks of one particular type of product, in this case SST, has been transformed by developing a data-processing framework that enables data from all sources to be used in a complementary approach. This happened when an international group[3] of scientists working in different sectors of the SST data chain (spanning satellite sensors, SST retrieval algorithms, dataset production and distribution, operational assimilation into NWP and ocean-forecasting models, creation of climate datasets, long-term data stewardship, etc.) and associated with each of the major remote-sensing agencies, responded to the challenge to produce an SST data system to meet the needs of 21st century operational oceanography.

[3] GHRSST = Group for High Resolution Sea Surface Temperature (pronounced *grist*).

14.4.1 Providing merged data from multiple sensors/satellites

The case has already been made in Section 14.2.2 for avoiding the rather narrow, competitive approach in which an agency promotes the benefits of data products from its own sensor without reference to other products (Figure 14.2), and the user plays the same game, selecting which product best suits their requirement and rejecting the others. The rationale of using an integrated approach is that all available data should be presented to the NOP model for assimilation in order to achieve the best output (Figure 14.3). The quality (as measured by the error estimate and confidence indicator) of individual pixels will determine whether they are actually used to constrain ocean state variables in the model nowcast. However, it would be misleading to imply that the NOP approach is always fully equipped to cope with the different characteristics of observations of the same variable made by different sensors, or with the assimilation of all available data (which is often impractical due to the enormous volume of available data). It would also be wrong to assume that the use of NOP models eliminates the need to go on producing satellite-only datasets of particular ocean variables that combine data acquired using different sensors and produced by different agencies. Such merged datasets are essential in their own right to serve particular applications: they are often the best way of preprocessing satellite inputs to models, they are useful for scientific studies, and they also contribute to climate records. Here we address the issues to be taken into account when generating merged satellite datasets, with emphasis on making the resulting dataset as useful as possible for operational users.

There are three main reasons operational users may prefer to use an ocean data product that combines measurements from several sensors into a single analysis field of a particular ocean variable. These are improvements in sampling frequency or coverage, increased security against catastrophic loss of data, and improvement in data quality (reduced errors) which comes from the use of different methods to produce different data products. At the same time users must not overlook the risk that, if data from different sources are combined carelessly without regard to the particular character of each input or the application for which output will be used, the diversity and differences between individual datasets will generate additional errors in the combined product. This could defeat the object which is to obtain a combined data product that is better than any individual input.

This is the main reason discerning users have been reluctant to use data from more than one satellite data source. They rightly judge that the task of combining data from mixed sources is best performed by those with expertise in remote-sensing methodology, if unanticipated errors are to be avoided. Indeed it is probably best done by collaboration between the different agencies that produce the datasets and who understand their individual characteristics. For the same reason, those in charge of NOP models may prefer to use a combined data product for assimilation in which anomalies of particular inputs have already been harmonized, rather than having to program into the assimilation process a separate subroutine that deals with those anomalies.

Improvements in sampling and coverage need little explanation. The availability of data from two sensors (on different platforms, of course) instead of one can double the number of measurements. Moreover, different ocean remote-sensing missions make different choices in the tradeoff between spatial and temporal resolution. For example, combining the data from a geostationary and a polar-orbiting sensor delivers the limited coverage, medium-resolution, and 30-minute sampling interval of the first with global coverage, finer resolution, but only daily samples of the second. For the operational user looking for new observations at regular intervals (e.g., every 3 or 6 hours) the use of combined datasets means having wider and more dense spatial coverage for each assimilation time. However, there are pitfalls from combining data from different orbits, which must be avoided. An obvious one is the use of data from two sensors in Sun-synchronous orbits having different Equator crossing times. Any diurnal variability in measurements is aliased by a Sun-synchronous orbit, producing a bias which differs with overpass time. This is not very apparent from the record of a single sensor, but can appear as variable bias between two sensors when their data are combined.

The advantage of increased security from catastrophic loss of source data arises because if one of the sensors providing data has a sudden failure the input stream to the operational application is reduced but is not cut off entirely. An operational service that is entirely dependent on a single input stream is extremely vulnerable. Where the service is set up to use multiple data sources, the quality of the service is likely to be reduced when one of them fails but it still keeps going. This makes for a more robust and resilient service, which is a more important performance criterion for some types of operational application: those where poorer quality information is better than none at all. Nonetheless a well-designed system for combining data should make provision for adjustments to be applied in case one of the inputs is lost for days, weeks, or months while a replacement sensor is provided. Such capability is essential in any case since data from new sensors need to be brought online while other sensors are retired from supplying data at the end of their operational life.

The improvement in data quality arising from the diversity of sensor methods may seem a more obscure benefit from combining data. It derives from consideration of the sources of error in a particular ocean measurement method and whether those errors can be detected or estimated. For example, a certain sensor might produce systematic errors in a particular geographic region, or at a particular time, but they are never detected because the dataset does not happen to be validated at that location or time. Independent methods of measuring the same variable will have different sources of error and therefore comparisons between the two datasets would reveal the systematic error. Moreover, such comparisons also yield objective measures of data uncertainty, which is an essential ancillary field for many operational applications.

Of course, different errors associated with diverse data sources could combine to reduce its overall quality unless treated properly. Analysis of differences between input datasets, when and where they overlap, reveals both the bias and standard deviation of the differences. Before combining data, they can be adjusted

relative to each other to remove the bias. While this does not necessarily improve the absolute accuracy of the combined dataset, it prevents additional noise (spatial and temporal variability) being introduced, which would occur if the different datasets have different mean values. Sometimes there may be known differences because different sensors measure subtly different properties of the ocean, or may suffer from a different diurnal bias (as mentioned above). Where processes are understood, then adjustments can be made to allow for these differences, aiming to convert each dataset to represent a standardized version of the observed ocean variable.

It should be clear from the above discussion that the process of combining datasets of the same variable from different sensors generally needs to be more intelligent than simple merging and averaging within space and time windows. Depending on the application a combined data product may be required every 3, 6, or 12 hours, 1 day, or 7 days. In most cases there are also gaps in sampling which need to be filled to meet a typical user requirement that every pixel of the combined dataset contains a valid value. This requires objective analysis of all available data using a method like optimal interpolation to fill spatial gaps to produce what is called the "analyzed product". The use of objective analysis implies that error estimates need to be supplied for all input data so that more reliable data carry more weight. For near real–time analyses, analysis from the previous time window will also be used. If subsequent reanalyses are performed, analysis of the following time window can also be used as input. Because a pixel value in the final analysis product cannot uniquely be tracked back to any particular pixel measurement in input (level 2 or level 3) datasets, it is described as a level 4 data product (see Section 2.3.6).

As long as no other information (e.g., from a model or *in situ* measurements) have been included in the analysis procedure the level 4 product can be considered as the "best" field for that ocean variable as measured by satellites. But the meaning of "best" depends on the particular analysis model that was used and the degree of spatial and temporal smoothing that has been applied. Potentially there may be several different analysis products generated from the same level 2 and level 3 satellite data sources, each designed for a different purpose, such as a near real–time input to an NOP model, or a high-quality climate record, and each would be produced with a different analysis configuration. Analysis should also assess the uncertainty attached to the value in each output pixel, to produce pixel-by-pixel output errors. Level 4 analysis datasets (what we have previously called combined data) should be interpreted with regard to their analysis configuration and the error field.

The foregoing description of data analysis methods was written generically so as to be applicable to a range of ocean variables. Sea surface height from altimetry and products derived from ocean color are variables for which quite a lot of work has been done to develop analyses that combine data from multiple sensors. They are discussed in the following subsections. In the case of SST, the emergence of an international collaborative approach to producing merged SST datasets is described in Section 14.4.2.

Altimetry

As mentioned in Section 2.4.5 sea surface height data since the early 1990s has been acquired by several satellite altimetry missions, principally the TOPEX/Jason series run by a French–U.S.–European consortium, and the European Space Agency's ERS Envisat series. These individually produce geophysical data records (level 2 data) for individual sensors, including SSHA, SWH, and wind speed along-track. The U.S. Navy's Geosat Follow-On (GFO) Mission has also produced comparable data. For the last 10 years, combined products from each of the altimeters operating at any time have been produced by a processing system, named SSALTO-DUACS after the French group (*segment sol multimissions d'altimétrie, d'orbitographie et de localisation précise*) which runs it and the EU project "Data Unification and Combination System" which first developed the concept in 1997.

After careful cross-referencing has been performed between different altimeters, the main combined products are sea level anomaly (SLA) and absolute dynamic topography (ADT) globally gridded on a $1/3 \times 1/3^{\circ}$ Mercator grid and integrated over 7 days. In addition higher resolution data are produced for the Mediterranean and Black Seas. A near real–time product is generated to serve many operational oceanography needs around the world, including the MyOcean marine core service for Europe mentioned in Section 14.2.1. Note that to produce the ADT field requires a best estimate of the ocean geoid, which is continually being revised following research and new information received from satellite gravity missions such as NASA's GRACE and ESA's GOCE. In addition high-quality reference datasets are produced after reprocessing to serve the needs of the climate-monitoring community, along with long-term mean fields of sea level and dynamic topography. As the products and services provided by DUACS are regularly updated, readers should refer to the website[4] for latest product details.

The DUACS service of combining data from all altimeters demonstrates how the preparation of merged products in near-real time for operational users also facilitates subsequent production of high-quality reprocessed products, since the scientific knowledge and processing skills needed for both are very similar. It addresses many of the issues discussed generically in the previous few pages. In some respects merging altimetric height data has fewer pitfalls than color or temperature products, because it is straightforward to specify the analyzed variable. On the other hand, the extreme accuracy sought in altimetry for operational applications, resolving down to millimeters, makes the success of the DUACS service a remarkable achievement. Widespread use of the service by agencies around the world confirms that satellite altimeter data for assimilation are now an essential element of ocean forecasting and planetary management.

Ocean color data

Sections 13.2.6 and 14.3 in this book have drawn attention to the difficulty of providing reliable chlorophyll or other color-derived satellite data products that

[4] *http://www.aviso.oceanobs.com/en/data/product-information/duacs/index.html*

can meet the need of operational ecosystem modeling and forecasting systems. A central problem is the scarcity of satellite ocean color observations. Even with three ocean color sensors in orbit, the problems of cloud cover, sunglint, and low illumination at mid to high latitudes in the winter hemisphere restrict ocean coverage to less than 25% in a single day and less than 65% in a 4-day period (IOCCG, 1999). For a single sensor the coverage is about 2/3 of these percentages. Ideally three or four ocean color missions need to be in orbit at any time to maximize retrievals from cloud-free regions. Consequently there is a need to combine data from several sensors.

Given the differences between color sensors (different spectral band sets, viewing angles, overpass times relative to local noon, sensor calibration methods, atmospheric correction models, retrieved ocean variables, product retrieval algorithms, etc.) the data products from different ocean color missions are not immediately intercomparable. Even the primary measurement—atmospherically corrected water-leaving spectrum—is expressed differently, in one case as normalized water-leaving radiance, in the other as remote-sensing reflectance. Moreover the basic chlorophyll product produced by different missions is defined in subtly different ways. This presents a challenge to the operational user intending to start assimilating ocean color data products into ecosystem models.

Nonetheless there is now available a time series of data from ocean color missions (see section 6.4.3 in *MTOFS*—Robinson, 2004) including OCTS (1996–1997), SeaWiFS (1997–present [2009]), MODIS (2001–present), and MERIS (2002–present). In recognition of the importance of blending these data into a coherent and consistent long-term record for climate applications, and to help prepare for near real–time merging of color data to meet the anticipated needs of NOP systems in the near future, ESA supported the GlobColour program.[5]

A fundamental task for GlobColour was to assess different merging strategies already proposed by the ocean color research community (IOCCG, 2007). These included methods that start from both radiance and derived bulk properties (such as surface chlorophyll concentration). The final choice to merge at the level of water-leaving radiance was made following an algorithm intercomparison and tradeoff analysis against *in situ* data. The GlobColour service has produced global data sets of chlorophyll concentration, water-leaving radiance, diffuse attenuation coefficient, colored dissolved and detrital organic material, total suspended matter or particulate backscattering coefficient, turbidity index, and several other variables. These products are generated on a daily, 8-day, and monthly basis, on a grid projection of 4.63 km equal area bins on a sinusoidal grid.

To produce combined water-leaving radiances, different inputs are merged by weighted average, based on sensor characterizations. Most other products are produced using a semianalytical ocean color–merging model (Maritorena and Siegel, 2005) applied to water-leaving radiances. This pioneering work should provide the basis for developing a real-time service for operational oceanography to serve the needs of GMES marine core services. This is expected to deliver on a daily basis a

[5] See the GlobColour website: *http://www.globcolour.info/index.html*

global ocean color dataset derived from MERIS and MODIS, intended for use as an input to forecast models (as discussed in Section 14.3).

14.4.2 GHRSST: A case study on preparing SST data for operational applications

In 2000 a challenge was presented to the SST community to rationalize various satellite-based SST products in order to meet the anticipated needs of ocean-forecasting models. It came when the Global Ocean Data Assimilation Experiment (GODAE) called on the community to

> "Develop a global, high-resolution sea surface temperature analysis with proper consideration of the skin effect and sufficient temporal resolution to resolve the diurnal cycle, that is available in real-time for all environmental and climate applications."

"High resolution" implied a spatial resolution of 5 km or finer. A target accuracy for SST of 0.2 K was the goal for climate applications although it was acknowledged that this accuracy would be harder to achieve for SST data delivered in "real time" which was to be interpreted as within a few hours of the satellite overpass.

Meeting the GODAE SST challenge

In response to this challenge a group of interested scientists came together from across the different stages of SST data flow, from sensor technology through algorithm development and data management to ocean modelers, meteorologists, and climate scientists, to form the GODAE high-resolution SST pilot project (GHRSST-PP). The group included academic oceanographers, applied scientists, and space agency managers; it spanned many nationalities and included the main space agencies and several ocean/atmosphere-forecasting agencies. At that time the broad assumption in mainstream oceanography was that SST measurement from space was a mature technique with little need for improvement and no longer a cutting edge research topic. What the GHRSST participants had in common, despite their different perspectives, was sufficient understanding of the subject to know that, at that time, they could not meet the GODAE challenge to deliver operational or climate quality data. They were aware of the shortcomings of each different data product available, but also recognized that the strength of some products or methods overcame the weaknesses of others, and *vice versa*. This gave them a realistic hope of meeting the GODAE challenge if different interests could work together in a complementary way instead of in isolation.

The story of how GHRSST-PP developed (Donlon *et al.*, 2007) is a demonstration of how the potential of satellite data to make a strong contribution to the public good has been fulfilled. After 3 years of discussion and debate, the group had identified a clear way forward with enough credibility that by 2004 ESA was prepared to support an international project office, while several agencies around the world invested in SST data-processing and product developments that

were aligned with GHRSST principles. By 2005 the first data products to comply with the GHRSST data specification (GDS) started to be produced routinely by ESA's Medspiration project, to be followed soon afterwards by GHRSST versions of SST products from U.S. satellites delivered by the NOAA/NASA-funded MISST project. By 2006 the routine availability in near-real time of GDS-compliant SST data products gave confidence to a number of operational players, mainly meteorological agencies, to start producing their own SST analysis products. For the first time these made use of all available SST sensors and improvements in resolution, reliability, and stability were soon very clear. By 2007 the U.K. Met Office had adopted a GHRSST-based SST analysis product to provide a boundary condition for its numerical weather-forecasting models. Satellite-derived SST is now embedded in operational meteorology and is assimilated into ocean-forecasting models, facilitated by adoption of the GHRSST approach.

Interestingly, basic level 2 SST products produced by the agencies have hardly needed to change. What GHRSST has done is to provide a scientifically underpinned framework of product definitions which has enabled all the different products to work with each other to deliver the information requested by the users. From the outset GHRSST was committed to meeting the requirements of the SST user community. Once users responded by ingesting GHRSST-compliant products into their operational systems the data-providing agencies needed no persuasion to start aligning their products with the GDS. This trend continues as GHRSST facilitates dialogue between major data providers and operational users.

To add substance to this short history of a successful project, we must look more closely at what characterizes the GHRSST approach. What are the issues that it has had to face? The following discussion assumes the reader is aware of the scientific basis of SST remote sensing using both infrared and microwave sensors (as presented by chapters 7 and 8 of *MTOFS*—Robinson *et al.*, 2008). The essence of limitations of SST measurements from space, highlighted by the GODAE challenge which provoked the GHRSST initiative, can be illustrated in Figure 14.10 which shows global maps of SST produced by several different sensors for a single day. Each sensor can be characterized in terms of one of the four broad classes of SST systems defined in Table 14.1.

The figure shows at a glance that no individual sensor provides global daily coverage at high resolution, partly because of data loss caused by cloud (IR sensors) or heavy rain and sidelobe contamination (MW) and limited swath width. The 15–30 min sampling by geostationary sensors provides more opportunities for cloud-free views within a single day. Note that although cloud does not obstruct a microwave radiometer, its spatial resolution is too coarse to meet operational requirements and it does not deliver reliable data within 100 km of the coast. Combining all products would considerably improve coverage and is the obvious way to proceed to reach the GODAE target. However, there are a number of other factors which complicate the measurement of SST from space and which would introduce additional errors if the various datasets were simply merged without further care. These are outlined below, along with a brief explanation of how the GHRSST approach seeks to mitigate the problem.

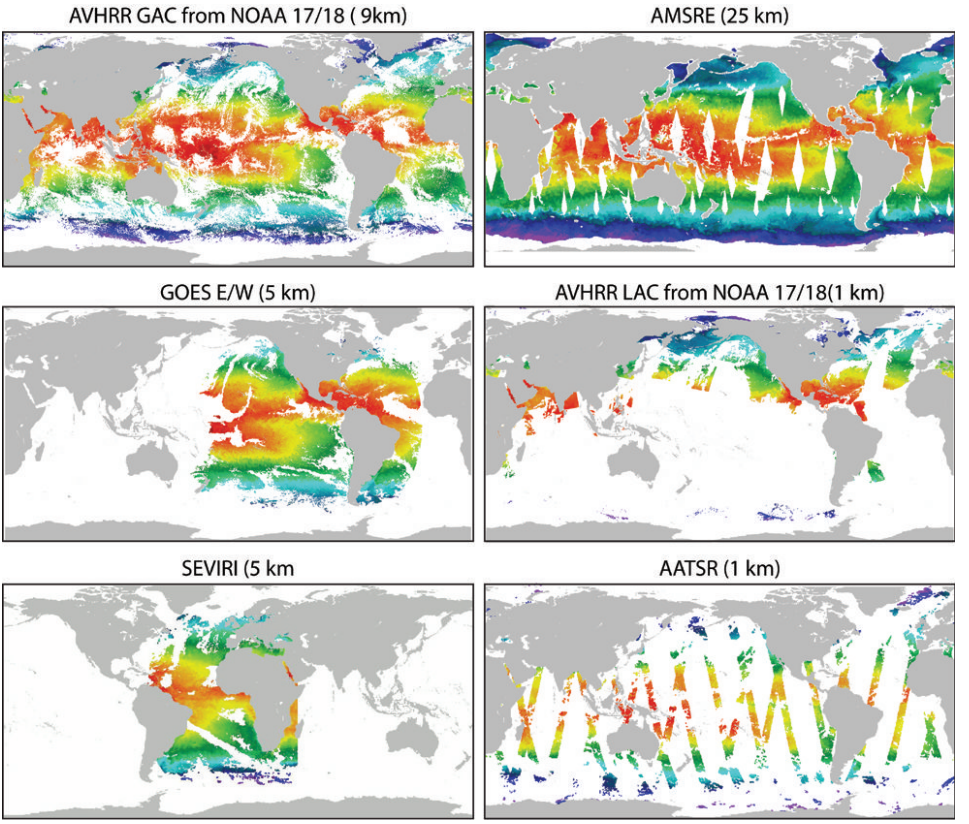


Figure 14.10. Typical example of daily coverage of SST from six different SST data products (see also Table 14.1).

Table 14.1. The main classes of SST sensor systems on satellites.

Class	*Sensor type*	*Platform orbit*	*Pixel size* (km)	*Revisit interval*	*Limit of cover*	*Accuracy* (K)
1	Wide-swath infrared (e.g., NOAA-AVHRR)	Polar LEO	1–2	12 h	Global	~0.4
2	Microwave (e.g., AMSR-E)	Polar LEO	25–60	~24 h	Global	> 0.5
3	Spin-scan infrared (e.g., SEVIRI)	GEO	3–4	30 min	Limited	0.5
4	Dual-view, conically scanning narrow-swath infrared (e.g., AATSR)	Polar LEO	1	~3 day	Global	< 0.3

Thermal structure and definitions of SST

The near-surface thermal structure of the sea is not uniform with depth (as illustrated in Figure 2.17). Different sensor types measure temperature at different parts of this structure. Microwave radiometry penetrates the cool skin of the ocean and detects the slightly warmer subskin SST. All infrared sensors inherently detect the skin SST although some IR-derived SST products are calibrated against *in situ* buoy measurements of SST which sample at an unspecified depth and may differ from the subskin if there is a diurnal warm layer present (see Figure 2.17b). This implies there is already considerable uncertainty in these products although, if care is taken to avoid buoy samples under conditions when a diurnal thermocline is present, these could be taken as subskin SST.

The GHRSST response is to insist that producers of SST datasets make it clear which skin or subskin their product represents. If buoy samples at unspecified depths have been used for calibration, the associated uncertainty should be reflected in the error statistics for that dataset. When datasets are to be combined they should first all be converted to either skin or subskin using an agreed adjustment. At present the accepted relationship between the two is

$$\mathrm{SST_{skin}} = \mathrm{SST_{subskin}} - 0.17\,\mathrm{K} \tag{14.1}$$

(Donlon *et al*., 1999, 2002). but this could be refined in light of further observational evidence since the amplitude of skin deviation decreases slightly in higher winds, and *vice versa*.

The type of SST typically represented by ocean models is that of the upper wind mixed layer, above the seasonal thermocline but below the cool skin or diurnal warm layer. To clarify what could otherwise be a confusing relationship between model temperatures and satellite measurements, GHRSST has defined the "foundation" temperature, $\mathrm{SST_{fnd}}$, which is the temperature in the water column just below any diurnal heating effect. By definition it is equivalent to $\mathrm{SST_{subskin}}$ at dawn, when the previous day's diurnal structure has been erased by convection and before any new structure has developed.

Dealing with diurnal variability

Diurnal warm-layer events occur whenever wind stress is insufficient to mix solar heating through the whole mixed layer down to the seasonal thermocline. Some warming events can have a temperature amplitude of up to 5 K, but there are probably widespread cases of low-amplitude heating that is hard to detect or to predict, but still introduces errors of a few tenths of a Kelvin. Because the phenomenon impacts the calibration or validation of many satellite SST products, and also confuses the interface between satellite data and model temperatures, GHRSST considers this to be a central issue that must be recognized to affect both the accuracy of SST products and the way they are applied or interpreted operationally.

Previously it was either quietly ignored or the draconian measure was taken of using SST from only nighttime satellite overpasses. The long-term goal is to seek reliable ways of estimating the magnitude of solar warming of SST compared with

the most recent dawn measurement, so that any satellite measure of SST_{skin} or $SST_{subskin}$ could be converted to the best estimate of SST_{fnd}. To facilitate that approach, the GHRSST recommendation is that level 2 SST data products have ancillary fields attached to every pixel, to include local wind speed and solar radiation. This provides data for further research on diurnal variability and the factors that control it (e.g., using subhourly samples of SST from Class 3—geostationary—sensors to provide case studies of diurnal variability). Until a diurnal modeling scheme has been proved reliable for operational tasks, ancillary data can be used to identify high-wind, low-insolation daytime data for which diurnal variability can be assumed negligible, thus reprieving some daytime data that would otherwise be rejected.

Facilitating use of level 2 SST inputs to operational systems: L2P products

In order to facilitate operational assimilation of SST into models, GHRSST specified new forms of level 2 SST products (such as those illustrated in Figure 14.10) produced by the various agencies. The important thing to note first is that no change is made to actual SST values delivered for each pixel. It is not appropriate for an intermediary body between producers and users of SST data (which is what GHRSST has effectively become) to interfere with agencies' own efforts to deliver the most accurate SST products. Instead, the GDS for level 2 products (called level 2 preprocessed or L2P products) requires addition of fields that are considered essential for effective assimilation or for generating free-standing SST analysis (level 4) products. These ancillary fields should also be beneficial for many other uses of L2P SST products, including scientific applications.

Ancillary fields include error estimates (bias and standard deviation) from every pixel, fields such as: wind and solar radiation needed for assessing the likely impact of diurnal warming on the skin or subskin observation from the satellite; measures of departure from a reference field to provide an easy way of testing whether data represent abnormal variability; fields such as aerosol optical depth that are helpful in understanding anomalous atmospheric effects; measures of sea ice concentration where appropriate; and a confidence flag, specified uniquely for each type of L2P product, which can be used to stratify data into best, acceptable, unreliable, etc. The confidence flag can provide a way of summarizing what is in other ancillary data, such as indicating a high probability of diurnal warming. It may include factors such as proximity to cloud that, for some SST products, is reckoned by the producer to imply a high risk of cloud contamination, or proximity to the coast in the case of a microwave SST product. Figure 14.11 shows examples of some of the fields that are incorporated within an SST L2P product based on the GHRSST data specification.

The rationale for including all this extra information, attached to each pixel, is to enable and encourage users to critically assess the data as they use them. The other simple but fundamentally important specification is that all L2P products for datasets from different suppliers have a common format, for which NetCDF was chosen, together with the NetCDF Climate Forecast (CF-v1.4) convention. This ensured that once an operational user had decided to use one of the SST level 2 products

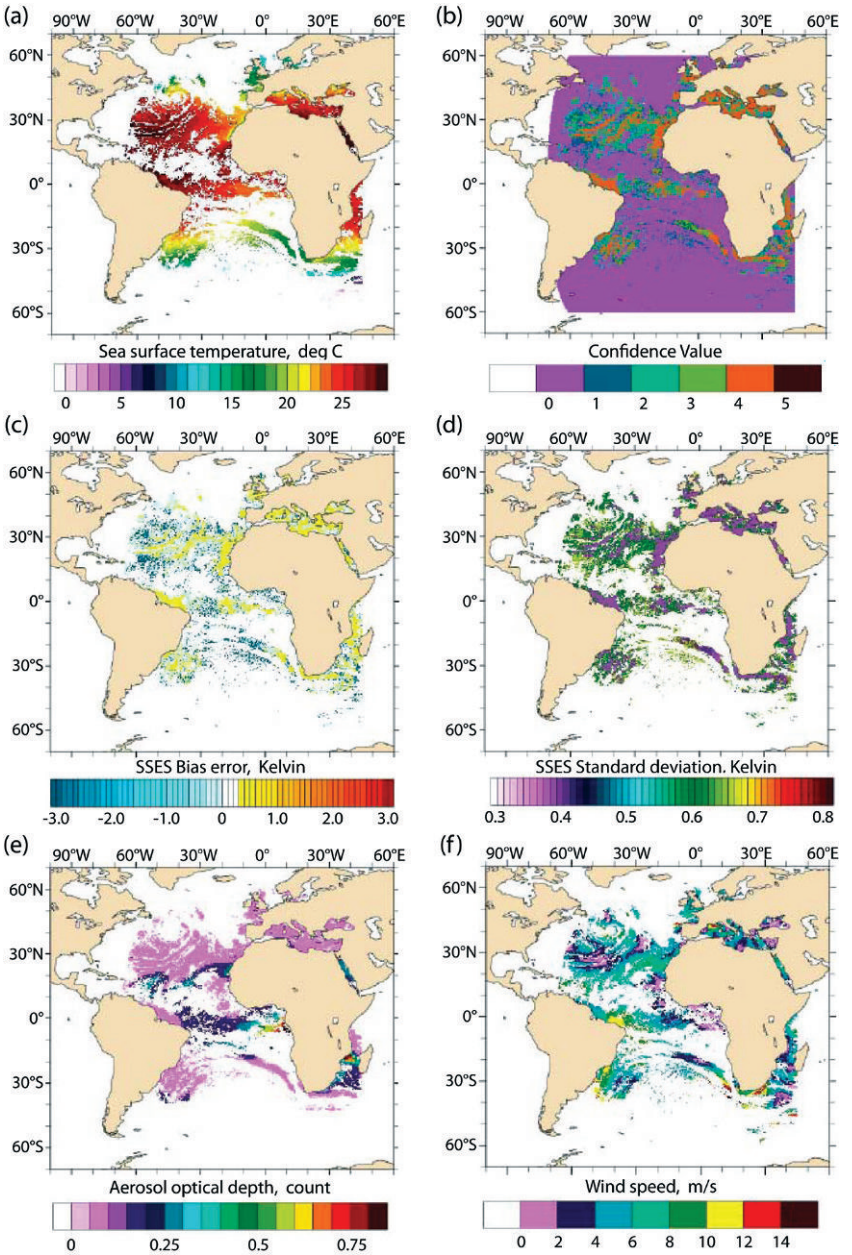


Figure 14.11. Example of the content of a GHRSST L2P file, in this case showing data from the SEVIRI sensor on the Meteosat second-generation geostationary satellite stationed over the Equator and the Greenwich Meridian. (a) SST. (b) Confidence value. (c) Bias. (d) Standard deviation. (e) Aerosol optical depth. (f) Wind speed. Note that fields (a)–(d) represent the mandatory contents of an L2P core file. The remaining ancillary fields are included in the full L2P files.

in GHRSST form, no further work was needed to ingest all the other L2P products. Previously it had not seemed worth the effort for operational users to write new code to ingest more than one product, when it was not clear to them that there would be any benefit. However, once L2P products began to be produced (first by projects such as Medspiration and MISST to kick-start the process but latterly by the originating agencies of SST products) and ocean-forecasting groups in meteorological agencies decided to switch to the new L2P products, they discovered for the first time how valuable it was to have access to more than one source of SST data.

L2P products needed to be produced within 3hours of the original level 2 data being released by the provider, in order to keep them current for assimilation into operational systems. Now that SST producers are themselves generating L2P versions of their products, timeliness of delivery should improve, and delays should reduce. If ancillary fields are not available in time, it is still important to get the SST product out by the promised regular delivery time which matches the schedules of operational users for ingestion to models or analyses. Because ancillary data are also very useful for preparing high-quality SST records for climate, there is provision for ancillary data to be upgraded at the time of any subsequent reprocessing of L2P datasets.

Level 4 products

The Medspiration project that initiated the generation of L2P products also produced experimental level 4 analysis products. This was the catalyst for other agencies interested in SST analyses to start developing their own level 4 products. GHRSST has specified formats for SST level 4 analysis products since these also are used by operational users who need to have confidence that any data they ingest will be recognized by their system. However, GHRSST has not sought to constrain the configuration of objective analysis or optimal interpolation techniques used in level 4 production, since these may be optimized for specific applications. General and scientific users of SST outside the operational forecasting sector are benefiting from the diversity of new analyses. Without streamlining L2P product delivery, it is doubtful that this activity would have developed.

Most analysis systems use all available L2P sources to produce a level 4 product. One of the sources, the Class 4 dual-view radiometer, the AATSR, can be seen in Figure 14.10 to have poor global coverage in a single day because of its narrow swath. This is an inevitable consequence of its conical scan, but that methodology is necessary to deliver the highly accurate atmospheric correction which is the unique capability of dual-view infrared sensors. This underpins the very stable and accurate SST fields delivered by AATSR (O'Carroll *et al.*, 2008), which are now being exploited within new level 4 analysis engines to provide a reference against which to adjust the bias of other input datasets (with better coverage), leading to a more stable and accurate analysis product. Thus, although by itself the narrow swath of AATSR has put off potential operational users, its unique strengths are now feeding into level 4 products and being widely used.

This synergetic outcome is an example of precisely what GHRSSST hoped to achieve in facilitating the complementary combination of diverse SST data types. As soon as this new role for AATSR was appreciated, a strong request from the international user community was voiced, through GHRSST, for ESA to provide a successor dual-view sensor of this type when AATSR reaches the end of its life. They stressed its importance for operational applications as well as climate monitoring for which it was designed. This request and the evidence of widespread operational use of the data allowed the European Commission (EC) and ESA to justify operational continuation of AATSR data by developing a new instrument called the Sea and Land Surface Temperature Radiometer (SLSTR) to be flown on the Sentinel-3 satellite series for a 20-year period starting in 2013. This seems to be an excellent outcome since it is evidence that GHRSST has created an active SST data producer–user community with a life of its own, encouraging the diversity of multiple players while enforcing a unity where it matters to operational users, in the stability of data production. delivery, and transmission.

Quality analysis

One other activity that GHRSST has sought to promote is data quality, not only through validation of SST products but also in assessing error and confidence information supplied within the L2P data package. It is the responsibility of those who provide SST data to the operational community to validate their products, although it is also incumbent on users to monitor the quality of data they ingest as well as that of their own derived products. GHRSST provides a forum in which producers and users can debate quality issues, an activity that is essential for healthy operational exploitation of satellite observations.

To facilitate that activity, GHRSST has spawned two tools for supporting quality analysis. One is a matchup database, in which *in situ* measurements of SST are matched with coincident satellite observations. This can then provide a basis for evaluating or monitoring error statistics included within L2P data, which operational users rely on as a means of weighting the influence of each pixel of data.

The other tool is the high-resolution diagnostic dataset (HR-DDS).[6] This facility underpins research to test and improve the merging of SST data from different sources and serves as an operational tool for comparing different L2P and level 4 products at the same locations. Nearly 200 DDS sites are specified, distributed around the World Ocean, and typically $2 \times 2^{\circ}$ in extent. At each of these sites data are extracted from every L2P data product containing valid SST pixels over the site, as well as level 4 and other SST products. The extracted data in each granule are resampled by nearest neighbour substitution onto a 0.01° grid, archived, and made publicly available. There is also a web portal to the HR-DDS, which allows users to access pre-evaluated statistical information about every data

[6] Web access at *http://sst.hrdds.net/*

granule, stored in a relational database, making this a powerful tool for data intercomparison.

14.5 OIL SPILL MONITORING

14.5.1 Introduction

It is more than 30 years since the early enthusiasts for satellite oceanography started to discuss the feasibility of monitoring oil spills from space, but only very recently has this finally become an operational reality. This section first briefly reviews the different remote-sensing techniques, several of them available only from aircraft, used to detect and measure the properties of oil slicks on the sea surface. Section 14.5.2 considers the strengths and weaknesses of using satellite data, and in particular images from synthetic aperture radar (SAR), for detecting and monitoring oil spills, adding to the previous discussion of the subject in section 10.11.3 of *MTOFS* (Robinson, 2004). In this chapter's context of operational applications, emphasis is placed on how satellite data can contribute to detecting, policing, and preventing polluting oil spills arising from anthropogenic causes. Finally Section 14.5.3 outlines a new operational service established for this purpose in European seas.

Oil slicks on the sea surface can be observed by a variety of remote-sensing methods (Trieschmann *et al.*, 2003). Damping of capillary and short gravity waves by any surface film material (natural or man-made) can be sensed by imaging radars (SAR or SLAR) because the consequent removal or strong reduction of Bragg backscatter produces dark signatures in SAR imagery. Airborne systems for local pollution control therefore locate oil discharges using SAR or SLAR. Combined infrared/ultraviolet (IR/UV) scanning is used to quantify the extent of the film, the UV showing the boundaries of the fine sheen due to light oil fractions. The IR detector is sensitive to the thermal signature of an oil spill, which can be quite complex: the lower IR emissivity of oil compared with water means a thin oil layer has a lower brightness temperature than the surrounding sea, although a thick oil spill absorbs solar radiation and may show up brighter by its higher temperature. A microwave radiometer can also be used to quantify the thickness. A preliminary classification of oil type, before physical samples have been obtained, can be made using a laser fluorosensor. Deployment of a combination of sensors from aircraft has proved to be an effective operational approach.

A known oil spill can thus be investigated, characterized, and its movement monitored while steps are taken to minimize its environmental impact. Aircraft continue to offer the best platform for observing individual slicks, but only after the existence and general location of an oil slick contaminating the sea surface has been reported so that the aircraft know where to go. An effective operational monitoring system must be able to detect oil slicks over a very wide area, soon after they occur and without reliance on reports from mariners. The objective of such a system is to provide an alert of previously unreported slicks when they are still a long way offshore, in time to organize remedial action from ships and aircraft, if necessary.

Only satellite-based sensors can deliver such routine wide-area surveillance (Brekke and Solberg, 2005).

14.5.2 How can oil spills be monitored routinely from space?

There are two candidates for sensor types capable of detecting oil spills from satellites: color imagers which are passive radiometers using reflected sunlight in the visible and near-IR waveband, and SARs. Color sensors detect oil films either by the contrast in apparent coloration of the spill compared with the rest of the water surface, or by using sunglitter patterns to distinguish between low surface roughness in the slick region and the greater roughness outside. Color sensors can also detect dispersed oil by water discoloration even when the surface slick is less pronounced. Sensors like SeaWiFS, MODIS, and MERIS have been shown to detect large oils slicks under favorable cloud-free and solar illumination conditions, but the effectiveness of color sensors for oil slick monitoring has been demonstrated mainly by airborne sensors. The approximately 1 km spatial resolution of satellite color sensors is very limiting; reports of oil spill detection from satellites have mostly been based on high-resolution mode images available from MODIS (250 m) and MERIS (300 m). From space, it is cloud cover that most seriously limits the capacity of ocean color to offer a credible alternative to SARs, which presently provide the basis for wide-area operational oil spill–monitoring services.

Following the launch of ERS-1 in 1991 and the steady provision of SAR ocean data since then, several research projects were initiated during the 1990s to develop image analysis tools and management decision systems that would exploit satellite SAR systems. The capacity to distinguish oil slicks as sharp-edged regions of low radar backscatter on an SAR image makes it feasible to develop image analysis systems for automatically scanning large numbers of SAR images in order to detect possible oil slicks. Figure 14.12 shows an ASAR wide-swath mode image over the North Atlantic Ocean off the northwest coast of Spain, acquired on November 17, 2002 while the drama of a major oil pollution event was unfolding when the tanker *Prestige* broke up. However, detection of oil slicks is not a precise art; situations occur where dark patches on SAR images can generate false-positive identifications of oil slicks where none is present, while high wind conditions may prevent oil spills from showing up on SAR images, resulting in false negatives.

The causes of oil spill lookalikes, dark patches on SAR images that can lead to false positives, include localized regions of low wind below the threshold of about 2 m/s to 3 m/s at which ripples at the appropriate Bragg wavelength are formed, grease ice, and heavy rain that can dampen Bragg ripples. Near-surface, short-scale hydrodynamic processes, such as internal waves and shear zones can also suppress Bragg ripples, as do naturally occurring biogenic surface films associated with phytoplankton blooms or fish (Espedal and Johannessen, 2000; Brekke and Solberg, 2005). Whereas these signatures can generally be distinguished from anthropogenic oil spills in high-resolution, airborne radar imagery, it is more difficult using satellite SAR images. Nonetheless reliable oil spill detection algorithms have been developed (see, e.g., Solberg *et al.*, 1999; Fiscella *et al.*, 2000; Solberg and Brekke,

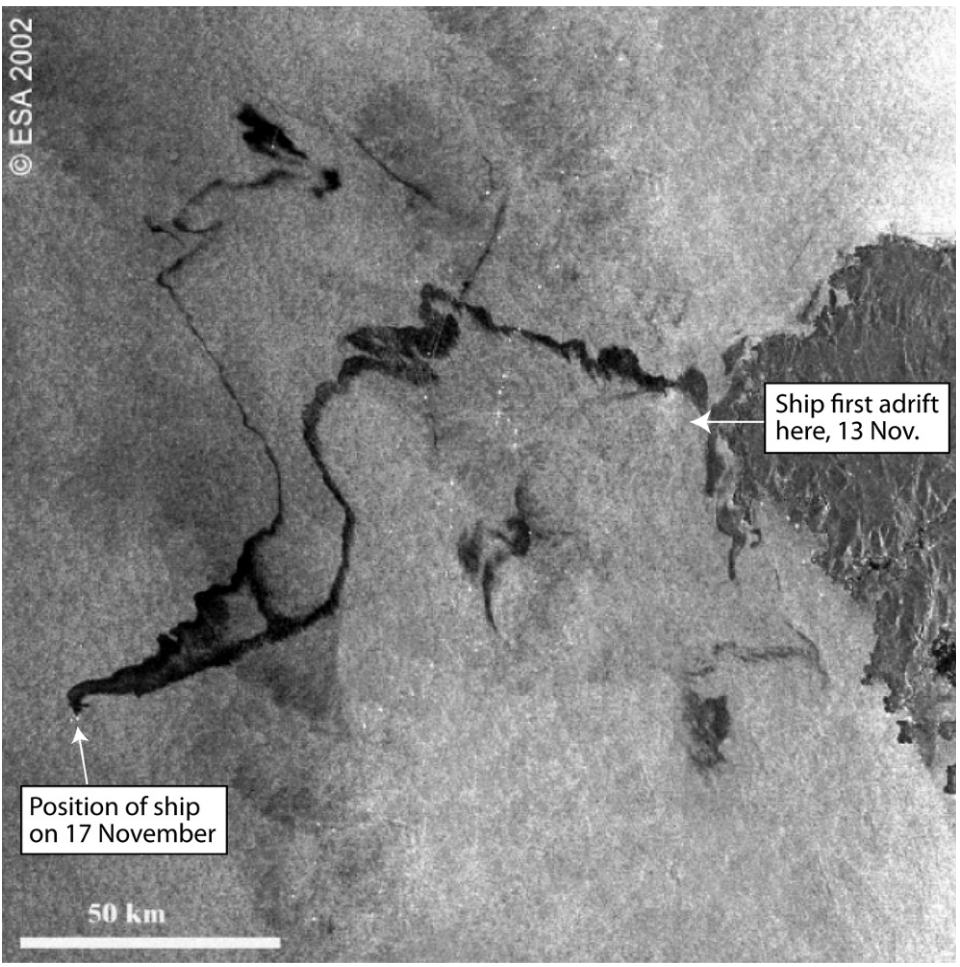


Figure 14.12. On November 13, 2002, following a heavy storm off the Atlantic coast of Spain, the oil tanker *Prestige* suffered a puncture in its oil tank, a crack in its hull, and engine loss. It was towed away from the coast and abandoned by its crew on November 15. On November 14 the International Charter on Space and Major Disasters was activated, leading to acquisition of all available synthetic aperture radar (SAR) images of the region for several weeks following the incident. This image is an ESA, Envisat, ASAR, wide-swath mode image acquired on November 17 when the ship was still being towed away from the coast. Two days later the vessel split in two and sank. This image shows the extensive oil slick produced by continuous seepage from the leaking tanks. It is revealed by the dark zone of very low radar backscatter resulting from the damping action of oil on waves of around 10 cm wavelength, which normally contribute to Bragg scattering for this C-band SAR. The exceptional clarity of the slicks in this image is somewhat fortuitous because the light-to-moderate wind is optimal for slick detection. Under high-wind conditions the damping action of oil is overcome by wind stress and spilled oil loses its strong SAR signature even though it is still present.

2008) which proceed by first detecting dark spots, then extracting these as individual features whose shape, orientation, and edges are defined, and then classifying them as spills or lookalikes in the context of wind speed and direction, local oceanography, previous history of oil spills in the region, etc. Classification can also be refined by tracking the history of slicks and relating them to the recent wind history in the region (Espedal and Wahl, 1999).

14.5.3 CleanSeaNet, a European service for oil spill detection

As long as there were only one or two narrow-swath SARs in operation (e.g., using just ERS-2 and Radarsat-1 in the late 1990s), the revisit interval for SAR images was several days so that an operational service was not feasible except at very high latitudes, where sampling frequency increases as orbits converge. However, once automatic oil spill detection software had been proven, it could be applied to the archive of SAR data back to the launch of ERS-1 in 1991, enabling statistical studies to determine the probability of anthropogenic oil spills occurring in different geographical regions, such as the Mediterranean (Ferraro *et al.*, 2008). These studies identified locations at high risk, such as along the routes of tanker traffic, or in regions of offshore oil production. They could also be used to test models that predict the trajectory of identified oil spills, based on knowledge of ocean currents and wind vectors.

The accumulation of this type of information demonstrated the effectiveness of using SAR images for oil spill detection, encouraging further investment in operational systems which can take advantage of improved coverage by satellite SARs. An important example of a major, new operational development is the establishment of the CleanSeaNet service by the European Maritime Safety Agency (EMSA). EMSA is an organization established by the EC with the task of enhancing overall maritime safety system across Europe. One of its goals is to use satellite monitoring to reduce the risk of marine pollution and to assist member states of the EU in tracing illegal discharges at sea. CleanSeaNet was established in 2007 as a service which effectively delivers for the single issue of oil pollution what the GMES Marine Core Service (see Section 14.2.1 and Figure 14.1) does more generally for marine management. Thus CleanSeaNet acquires from its "upstream partners" (the appropriate space agencies) all available SAR data for the seas within Europe and the surrounding ocean. On behalf of all the member states it analyzes these data and issues oil spill warnings to the appropriate "downstream user" (i.e., the country with oversight of the particular sea area where the spill is identified. It also co-ordinates cases where oil spills threaten the waters of several nations. This management approach, based on the principle of subsidiarity, removes the need for duplication of the SAR data-processing effort by several member states.

SAR images are acquired routinely from ASAR on Envisat and SAR on Radarsat 1 and 2. Figure 14.13 shows the enhanced revisit frequency possible individually from these SAR products. Wide-swath data mode (images that are 405 km square in the case of ASAR) provide the most frequent sampling but are not always available depending on the scheduled requirements for other SAR data modes.

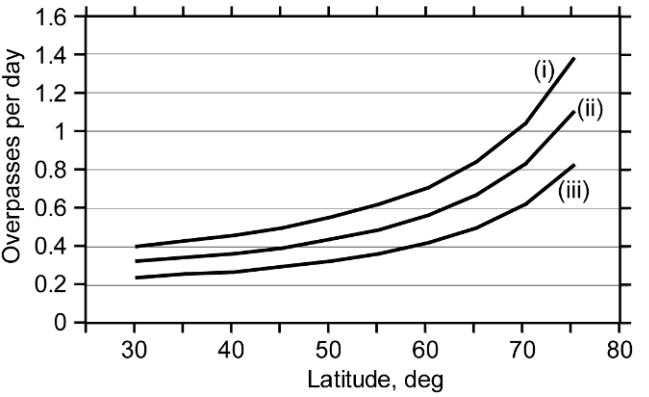


Figure 14.13. Approximate revisit interval for SAR acquisitions showing the dependence on latitude of SAR coverage for oil spill monitoring. (i) Radarsats 1&2 ScanSAR wide mode. (ii) Envisat ASAR wide-swath mode. (iii) Radarsats 1&2 ScanSAR narrow mode.

However, the combination of Envisat and Radarsat, plus any additional image data that may be obtained from the Phased Array L-band Synthetic Aperture Radar (PALSAR) on the Japanese Advanced Land-Observing Satellite (ALOS), and from the German X-band SAR on TerraSAR-X, can result in a sampling frequency better than once per day for North European waters, and better than once every 2 days for South European waters and the Mediterranean Sea. SAR raw data are transmitted to the nearest ground station, where they are immediately processed and interpreted by experienced image analysts on behalf of EMSA. Within 30 minutes of a satellite overpass, information about a detected oil spill and the image itself are sent to the pollution control authorities of the member states responsible for the area of interest. In many cases, the technology of the automatic identification system[7] (AIS) is combined with SAR data to link ships to potential pollution events. At this stage locally managed aircraft surveillance and vessel patrols will be sent to the area to verify the spill and, when confirmed, to identify the polluter if possible. The better the coverage and the faster the response, the greater the likelihood of prosecuting offenders, leading to stronger deterrence of potential polluters.

Regional and local task sharing between EMSA and member states also provides for feedback about verification of possible slicks, which should lead to improvement of detection algorithms. For major spills the outputs of regional ocean circulation–forecasting models can also be used to assist in forecasting expected trajectories of the oil. The CleanSeaNet service has only been implemented relatively recently, and as yet no statistics are available about SAR coverage available, reliability of oil spill event detection based on satellite SARs, effectiveness of the service in reducing the impact of pollution, and any evidence of an improved deterrence effect. Readers are encouraged to check future reports on the service

[7] AIS is a short-range coastal tracking system used on ships and by vessel traffic services (VTSs) for identifying and locating vessels by electronically exchanging data with other nearby ships and VTS stations.

produced by EMSA in order to ascertain whether this operational service achieves its goals and justifies continued investment in maintaining the service indefinitely.

14.6 USING SATELLITE DATA FOR CLIMATE MONITORING

14.6.1 Introduction

The climatology of the ocean and the ocean's role in climate change are topics of such great importance for the sustainable future of human civilization on our planet, as well as being fields of research so rich in scientific interest and challenge, that it deserves a chapter, if not a whole book, to fully explore how satellite observations can be used to monitor ocean climate change. However, in order to complete this volume before the earlier chapters are out of date, and to contain its size, the topic of oceans and climate has been restricted here to just a few pages. Nonetheless, readers of previous chapters will have noted the ways in which ocean remote-sensing methods provide the larger space and timescale perspective of oceanographic phenomena which allows them to be studied in the climatological context. Thus for a scientist wanting to pursue research on the ocean's role in climate and wanting to make use of satellite data, the rest of this chapter is intended as an introduction.

In practice, given the relatively short time span of global satellite data presently available, typically between one and two decades for ocean measurements of reliable quality with quantified errors, there has been little opportunity yet for a lot of results to emerge from climate research based on remote sensing. That is about to change, and this section aims to prepare readers for the expansion of satellite-based climate research expected in the coming few years, building on planned production of climate variables derived from satellite data (GCOS, 2006). Presently the Global Climate Observing System (GCOS) serves the requirements of the World Climate Research Program (WCRP) to provide measurements of space-time variability of the Earth's climate system, in order to gain knowledge and understanding of climate change processes (GCOS, 2004). Such research has been driven by more than scientific curiosity since the Intergovernmental Panel on Climate Change (IPCC) was set up to advise governments through the United Nations Framework Convention on Climate Change (UNFCCC). Reliable observations of ocean elements of the climate system, stable and continuous over many decades, are therefore considered as an essential international public good, which needs to be provided by operational procedures comparable, but not necessarily identical, with operational monitoring for short-term ocean forecasting and management (discussed in Sections 14.2 and 14.4). Space agencies are now starting to set up systems of satellites, sensors, and data-processing services to meet that need.

In the rest of this section, Section 14.6.2 considers why ocean observations are so important for being able to characterize and understand climate change in general, identifying the aspects of ocean science expected to be more important in this context. Section 14.6.3 will identify the formal structure of definitions established by GCOS for global climate records (GCRs) and essential climate variables (ECVs).

Finally Section 14.6.4 will focus on those ECVs that are specifically related to the ocean and derivable from satellite data, discussing the options and opportunities for new avenues of research that make use of the versatility of satellite-derived datasets.

14.6.2 The ocean's role in the climate system

The term "climate" is used by meteorologists to characterize typical meteorological conditions encountered at a particular geographical location and the way they vary over an annual cycle. It is not the same as meteorological conditions actually occurring at a given time which are referred to as "weather". This varies from day to day because of the turbulent, chaotic behavior of the atmosphere. Loosely speaking, climate can be thought of as the average of weather conditions over a suitably defined period of time. The distinction between climate and weather has been colloquially expressed in the statement: "Climate is what we expect, weather is what we get!" If climate is to be defined, say, as typical weather conditions in each month of the year, then climate statistics must be compiled from the average of several years of records of the weather during a given month.

However, a more complete description of climate will record how average conditions for a given month change from one year to the next, referred to as "interannual variability". When a record of climate statistics has been built up over many years, it becomes possible to explore longer period oscillations of the climate. For example, if climatology is based on the rolling average of a 5-year window over a 40-year span of observations, then oscillations having a period of 10 years or longer can be detected in the climate record and are referred to as interdecadal climate variability. Given the steady upward trend in atmospheric concentration of greenhouse gases, caused largely by burning of fossil fuels and deforestation, one of the tasks of climate science is to measure the secular trend of certain variables such as air temperature against a background of interdecadal variability. Now that global warming has been unambiguously detected (IPCC, 2007), another task is to determine whether the amplitude and periodicity of climate variability are themselves consequences of global warming of air temperature.

The study of climate has expanded from a purely meteorological concern to embrace all parts of the Earth system because it is now recognized how interdependent they are. The ocean, especially, interacts with the atmosphere, cryosphere, and land in ways which both mediate and moderate climate changes. If we are properly to understand how the climate system works in order to predict its future behavior we must be able to estimate the relative distribution of key climate variables, such as heat, carbon, freshwater, and so on between the atmosphere, ocean, cryosphere, and land, and also to measure fluxes of properties within and between these domains. This is a daunting measurement task.

For example, the ocean's large mass compared with the atmosphere makes it a large reservoir for carbon and heat; thus small changes in ocean pCO_2 and temperature can represent relatively large changes of carbon and heat storage compared with the carbon and heat content of the atmosphere. This places challenging demands on the precision of ocean measurements if the ocean's contribution to carbon and heat

budgets of the whole Earth are to be determined with sufficient accuracy to be able to predict future behavior. Variability of the ocean at lengthscales of a few hundred kilometers (as discussed in Chapter 3) also demands high spatial sampling frequency for ocean observations across the whole Earth if global integration is to avoid aliasing, which would lead to unacceptably high uncertainties when compiling the ocean component of global inventories of climate variables.

The reliable estimation of air–sea fluxes of heat and CO_2 places even tighter demands on the accuracy of surface oceanic and sea level atmospheric variables (as discussed in Chapter 10). Fluxes within the ocean which transfer heat, salt, and nutrients with the potential to fix carbon through primary production are moderately well understood, but simply monitoring them requires the use of assimilating ocean GCMs in the type of forecasting systems discussed in Section 14.2. Changes in ocean and atmosphere variables resulting from global warming may change the driving forces for ocean circulation, or induce nonlinear dynamic responses, that significantly alter the meridional fluxes important for climate. In order to be able to factor these possibilities into climate-forecasting models that anticipate major global warming, it is desirable to find ways of detecting the ocean's sensitivity by examining the much smaller changes within signals of today's natural climate variability. Similar considerations apply to the study of ocean–cryosphere interaction and fluxes. For example, can we learn from today's observations about how either sea ice or grounded ice sheets respond to slightly higher sea temperatures during an anomalous warm year, so that this observed response can be factored into global models? These are the types of question which should influence the design of the ocean-observing system for climate, if ocean measurements are to lead to improved climate-forecasting models.

There is another related set of questions which are less about the coupled climate system and more to do with the ocean's internal response to climate change, and how that will impinge on human civilization. One of the most obvious is the measurement of mean sea level rise and its geographical distribution. The capacity of altimetry to measure sea level changes (as discussed in Section 11.5.2) is likely to be exploited increasingly as governments start to prepare realistic, long-term contingency plans for dealing with coastal inundation. Another very serious consequence of the ocean's uptake of anthropogenic CO_2 into the ocean is the increase in alkalinity which must therefore be monitored as a climate variable. Related to this is the importance of monitoring how primary production and water turbidity change in response to both enhanced levels of pCO_2 in the upper ocean and changes in ocean circulation, upwelling, frequency of El Niño/La Niña events, and changes in summer sea ice distribution in polar regions.

Faced with these challenging questions and their implied need for accurate, precise, and spatially detailed knowledge of how the ocean changes over a range of climatological timescales, it is reassuring to note that the methods of satellite oceanography provide an appropriate set of measurement tools. Experience with applying satellite data to the fine spatial–resolution, high-frequency, global-sampling challenge of operational oceanography (as discussed in Sections 14.2 and 14.4) has not only given us confidence and vision for establishing a satellite-based ocean

climate–monitoring system, it has also started to equip us with the data-processing and management infrastructure that can implement such a system in practice. For example, in the field of SST (discussed in Section 14.4.2) the new, level 4, high-resolution, global SST analysis data products for operational application are also being adapted to provide a climate quality record of SST. The same is happening in relation to monitoring of other ocean variables. As this chapter is being written it is interesting to note that the team within ESA that has fostered the development of operational applications of ocean satellite data for the last 10 years has recently announced its intention to promote production of satellite-based ECVs in a new climate change initiative.

Thus the prospects are promising for developing an effective ocean climate–monitoring system based on remote sensing, as long as the present inventory of satellites and sensors is maintained. However, it would be misleading not to point out that space-based measurements by themselves are not enough to identify and understand the ocean's role in climate change without a matching program of well-planned *in situ* ocean measurements, which are needed to monitor the subsurface distribution of ocean variables. As noted above, our understanding of how the ocean moderates the response of atmospheric climate to enhanced concentrations of greenhouse gases requires both a detailed knowledge in space and time of air–sea fluxes of particular properties such as heat and CO_2, and an inventory of reservoirs of those properties in the ocean. Satellite data are essential for monitoring ocean surface conditions that constrain air–sea fluxes, but they need to be complemented with a comprehensive system of subsurface floats, moorings, and other *in situ* methods for monitoring the distribution of properties of the ocean throughout its depth.

The force of the argument in support of the necessity for subsurface ocean observations can be illustrated in a simple way by considering the ocean's thermal heat capacity. As the greenhouse effect traps more heat inside the Earth's system, the climate response depends on how that extra heat is distributed. If somehow the excess heat could all be absorbed in the deep sea by very slightly raising the temperature of a large mass of the ocean's coldest and deepest waters, while the temperature of the atmosphere and upper ocean hardly changed at all, then we would not be facing the expectation that the planet's weather systems will change significantly. Of course the climate record of atmospheric warming is evidence that this simplistic scenario is not the case; we still do not know with confidence what proportion of excess heat is being captured in the deep ocean. That is why it is so important to gain reliable direct measurements of how the ocean is changing at depth in relation to a variety of different climate-relevant variables. Satellite data alone are insufficient to provide this knowledge. It is therefore essential that a complete ocean climate–monitoring system combines satellite-based data with *in situ* observations for measuring subsurface conditions.

14.6.3 Essential climate variables

An essential climate variable (ECV), as its name implies, is an environmental variable characterizing some aspect of the global climate system which needs to be

measured or quantified regularly in order to monitor the way in which the climate system varies. In practice, climate scientists have established a list of required ECVs (GCOS, 2004) selected for meeting the criteria (i) that they have high impact for meeting the information requirements of the UNFCCC and (ii) that it is currently feasible to monitor them globally. Table 14.2 lists these variables, distinguishing between atmospheric, oceanic, and terrestrial domains. As scientific understanding of the processes of climate variability develops, prompting new questions, and as measurements of other variables become feasible, the list of ECVs can be expected to grow. Readers should be able to obtain an up-to-date list from the Global Observing Systems Information Center (GOSIC).[8]

It is important to recognize that the study of climate change has now acquired a significance that extends well beyond the curiosity-driven interests and problem-solving endeavors of academic research. Over the last 20 years the warnings of climate scientists, that anthropogenic emissions of CO_2 and other greenhouse gases is causing global atmospheric warming and climate change, have been taken seriously by politicians and the general public in all parts of the world. Once the UNFCCC was established, information about climate change started to become potent knowledge in the realms of politics, economics, industry, international relations, human rights, and much more, including personal ethics and an individual's lifestyle choices. What scientists discover about cause and effect in the processes controlling climate reverberates far beyond the environmental scientific community; it carries weight, influences opinions, and changes behavior in society as a whole. Therefore, like the other topics covered in this chapter, it is entirely appropriate to consider the monitoring of climate variables to be an application of science, responding to the requirement by agencies and interests outside the scientific community for science-based understanding of how to manage the impact of human civilization on the planet. Those engaged in climate monitoring must therefore accept that their work is driven not simply by the internal compulsion of scientific curiosity to find out how the world works, but is constrained by international institutional users of climate information. In order that politicians can make informed decisions when negotiating international treaties they need to know answers to the wide variety of important questions raised within the UNFCCC (IPCC, 2007).

The consequence of the very high significance now accorded to climate data, and the intense scrutiny to which they are subjected, is that great care must be taken to ensure the accuracy, provenance, and consistency of ECVs and all measurements on which they are based. It is especially important to be able to demonstrate that the methods used for measuring climate variables should not themselves be the source of trends discovered in long time series. This is problematic. Because the most useful climate data are those which extend over many decades, we are dealing with time series that start with observations made by obsolete measuring systems that were less reliable or comprehensive than those achieved by modern observing systems. Each different climate variable must be individually subjected to appropriate quality controls in order to minimize any errors and inconsistencies when assembling a long

[8] The GOSIC webpage is at *http://gosic.org/default.htm*

Table 14.2. Essential climate variables identified by GCOS (GCOS, 2004).

Domain	*Essential climate variables*	
Atmospheric (over land, sea, and ice)	Surface	Air temperature Precipitation Air pressure Surface radiation budget Wind speed and direction Water vapor
	Upper air	Earth radiation budget (including solar irradiance) Upper-air temperature (including MSU radiances) Wind speed and direction Water vapor Cloud properties
	Composition	Carbon dioxide Methane Ozone Other long-lived greenhouse gases Aerosol properties
Oceanic	Surface	Sea surface temperature Sea surface salinity Sea level Sea state Sea ice Current Ocean color (for biological activity) Carbon dioxide partial pressure
	Subsurface	Temperature Salinity Current Nutrients Carbon Ocean tracers Phytoplankton
Terrestrial	River discharge Water use Ground water Lake levels Snow cover Glaciers and ice caps Albedo	Permafrost and seasonally frozen ground Land cover (including vegetation type) Fraction of absorbed photosynthetically active radiation (fAPAR) Leaf area index (LAI) Biomass Fire disturbance

Table 14.3. GCOS climate-monitoring principles (GCOS-Secretariat, 2009).

1	The impact of new observing systems or changes to existing systems should be assessed prior to implementation.
2	A suitable period of overlap for new and old observing systems is required.
3	The details and history of local conditions, instruments, operating procedures, data-processing algorithms, and other factors pertinent to interpreting data (i.e., metadata) should be documented and treated with the same care as the data themselves.
4	The quality and homogeneity of data should be regularly assessed as a part of routine operations.
5	Consideration of the needs for environmental and climate-monitoring products and assessments, such as IPCC assessments, should be integrated into national, regional, and global observing priorities.
6	Operation of historically uninterrupted stations and observing systems should be maintained.
7	High priority for additional observations should be focused on data-poor regions, poorly observed parameters, regions sensitive to change, and key measurements with inadequate temporal resolution.
8	Long-term requirements, including appropriate sampling frequencies, should be specified to network designers, operators, and instrument engineers at the outset of system design and implementation.
9	The conversion of research observing systems to long-term operations in a carefully planned manner should be promoted.
10	Data management systems that facilitate access, use, and interpretation of data and products should be included as essential elements of climate-monitoring systems.

climate time series from heterogeneous observations. In order to underpin this, GCOS has set out a set of generic climate-monitoring principles (GCOS-Secretariat, 2009) which are summarized in Table 14.3. These need to be applied whenever an ECV is being generated. In pursuit of rigor in the way ECVs are defined, GCOS documents refer to the concept of a *fundamental climate data record* (FCDR) and distinguish this from what is meant by a climate data *product*. These are explained in the following subsections.

Fundamental climate data record

In GCOS documentation, the term fundamental climate data record (FCDR) is used to denote a long-term data record of the fundamental measurement from which a particular ECV may be derived as a product (GCOS, 2006). For an example related

to ocean remote sensing, if the ECV were SST the appropriate FCDR would be top-of-atmosphere (TOA) radiance measured in a particular infrared or microwave waveband. Irrespective of how the SST product is derived by atmospheric correction and other procedures, its long-term integrity is considered to depend on the accuracy and stability of measured TOA radiance, which is thus considered to be the fundamental constraint on the reliability of the derived ECV. Typically the FCDR is acquired over a long span of time from a series of instruments/satellites, with potentially changing measurement approaches. In this case there must be overlaps and intercalibration between radiances measured by different instruments, sufficient to allow generation of homogeneous products to provide a measure of the intended variable that is accurate and stable enough for climate monitoring. Thus FCDRs must include not only primary measurements (e.g., radiances) but also ancillary data used to calibrate them.

Where direct overlap of different satellite missions has not occurred, or where measurements from a "one-off" research spacecraft are used to contribute to the FCDR, there must be sufficient supporting measurements and information available to safeguard the continuity of calibration between one instrument and the next (Ohring *et al.*, 2005). These would include rigorous prelaunch instrument characterization and calibration, onboard calibration, and complementary surface-based observations.

Another important element of the FCDR must be the spatial and temporal sampling characteristics of each sensor/platform type used, including the overpass time for a Sun-synchronous orbit. In the case of SST, the same issues (such as diurnal variability) are relevant for generating an ECV dataset as for producing an operational SST analysis (as discussed in Section 14.4.2), although quality criteria differ between operational and climate applications. Sampling complementarity and conflicts between different instruments/platforms may also be relevant when constructing the FCDR. Despite possible limitations arising from the sampling characteristics of satellite-based observations, they are greatly outweighed by benefits from maintaining global coverage consistently by the same sensor for many years. This is why remote sensing has such an important role to play in climate monitoring. To encourage a rigorous approach to ensuring the quality of FCDRs, GCOS has set out a further set of climate-monitoring principles applicable to satellite systems (GCOS-Secretariat, 2009), as summarized in Table 14.4.

Climate data product

Whereas the FCDR can be considered as a directly measured quantity, even though several different observing instruments may be used, the term "product" should be used to refer to values or fields of geophysical variables that have been derived from FCDRs (GCOS, 2006). Such products, normally referred to as integrated climate products,[9] are typically produced from a combination of satellite and *in situ* observa-

[9] In NOAA documentation these are sometimes described as thematic climate data records (TCDRs).

Table 14.4. Specific GCOS monitoring principles applicable to satellite systems (GCOS-Secretariat, 2009).

1	Constant sampling within the diurnal cycle (minimizing the effects of orbital decay and orbit drift) should be maintained.
2	A suitable period of overlap for new and old satellite systems should be ensured for a period adequate to determine intersatellite biases and maintain the homogeneity and consistency of time series observations.
3	Continuity of satellite measurements (i.e., elimination of gaps in the long-term record) through appropriate launch and orbital strategies should be ensured.
4	Rigorous prelaunch instrument characterization and calibration, including radiance confirmation against an international radiance scale provided by a national metrology institute, should be ensured.
5	Onboard calibration adequate for climate system observations should be ensured, and associated instrument characteristics monitored.
6	Operational production of priority climate products should be sustained, and peer-reviewed new products should be introduced as appropriate.
7	Data systems needed to facilitate user access to climate products, metadata, and raw data, including key data for delayed mode analysis, should be established and maintained.
8	Use of functioning baseline instruments that meet the calibration and stability requirements stated above should be maintained for as long as possible, even when these exist on decommissioned satellites.
9	Complementary *in situ* baseline observations for satellite measurements should be maintained through appropriate activities and cooperation.
10	Random errors and time-dependent biases in satellite observations and derived products should be identified.

tions, relying on a physical model framework as the basis for blending data from different sources. Table 14.5 lists all the ECVs in the oceanic domain that are largely dependent upon satellite observations, specifying the type of data product that must be generated and the FCDRs appropriate for each.

The importance of distinguishing products from FCDRs is that climate data products are not exclusively dependent on FCDRs. Their accuracy and stability depends not only on space agencies' validation of FCDRs but on a variety of other agencies responsible for the reliability of *in situ* data. It also depends on the validity of the assumed physical model. Such products can be the result of a complex interplay between different measurement and operational agencies, and research groups, each of which may be expected to refine their procedures over time. Indeed, experience with satellite data records has shown that regular reprocessing

Table 14.5. ECVs for the oceanic domain, largely dependent on satellite observations, showing corresponding data products and FCDRs required (GCOS-Secretariat, 2009).

ECV	*Global products requiring satellite observations*	*Spatial and time-sampling resolution*	*Target accuracy*	*Target stability*	*Fundamental climate data records required for product generation*
Sea ice	Sea ice concentration	12 km daily	5%	5% per decade	Microwave and visible imagery
Sea level	Sea level and variability of its global mean	25 km daily	1 cm	0.5 mm/ decade	Altimetry
SST	Sea surface temperature	1 km 3 hourly	0.25 K	0.1 K/ decade	Single and multiview IR and microwave imagery
Ocean color	Ocean color, chl-*a* concentration derived from ocean color	1 km daily	5%	1% per decade	Multispectral visible waveband imagery
Sea state	Wave height, wave direction, wavelength, wave period	25 km 3 hourly	10 cm (SWH)	5 cm/ decade	Altimetry
Ocean salinity	Research towards measurement of changes in sea surface salinity	100 km weekly	0.05‰	0.05‰/ decade	Microwave radiances

of historic data is desirable as growing knowledge permits improvement of the quality of datasets and products. Consequently it is essential that strong collaboration is maintained between different players, coupled with disciplined documentation, strict version control of data, and carefully orchestrated introduction and testing of improved products, in order to ensure that long time series of such products do not acquire artifacts which are mistaken for evidence of climate variability.

In order to ensure that attention is paid to these points, the GCOS Steering Committee has issued a set of recommendations (GCOS-Secretariat, 2009) related to the planning of procedures to generate ECV satellite datasets and products. These are summarized in Table 14.6.

14.6.4 Ocean datasets used for climate

This section briefly identifies some of the outstanding issues for each satellite-based ECV in the ocean domain (as listed in Table 14.5), where preferred sampling resolution, target accuracy, and stability are also noted. At the time of writing, ESA is about to embark on a 5-year climate change initiative to develop satellite-derived

Table 14.6. GCOS requirements for datasets and products used for ECVs, summarized in tabular form from GCOS-Secretariat (2009).

Information to be attached to the data product	1	Full description of all steps in the generation of datasets and products, including algorithms used, specific FCDRs used, and characteristics and outcomes of validation activities.
	2	Information on publications in peer-reviewed journals, covering both the description and application of datasets and products.
	3	Statement of expected accuracy, stability, and resolution (time, space) of the product, including, where possible, a comparison with requirements stated in the Satellite Supplement (or any subsequent revision).
	4	Arrangements for access to datasets, products, and all documentation.
Required characteristics of the data product	5	Version management of datasets and products, particularly in connection with improved algorithms and reprocessing.
	6	Long-term stability and homogeneity of the product.
	7	Full application of all appropriate calibration/validation activities that would enhance product quality.
	8	Global coverage where appropriate.
	9	Timeliness of data release to the user community to enable monitoring activities.
	10	Application of a quantitative maturity index if possible.
Other factors	11	Facility for user feedback.
	12	Publication of a summary (preferably online) documenting point by point the extent to which the GCOS guideline has been followed.

ECVs. Readers should therefore expect significant advances in systems to generate these ECVs during the next few years.

Sea ice

Sea ice concentration and how it varies seasonally is an important indicator of climate change at high latitudes. It has been measured daily using microwave radiometers on polar-orbiting satellites since 1978 (as discussed in Section 11.4). This gives a 30-year climate record although there is scope for consolidating this record by improving or confirming the consistency of calibration between different algorithms,

frequency bands, and sensors that have been used over the years to construct it. It will be important to ensure continuity in the provision of passive microwave radiometers beyond the lifetime of the present SSM/I and AMSR-E sensors.

The focus of current interest is to observe trends in the reduction of summer ice cover of the Arctic Ocean following the unprecedented low coverage that occurred in September 2007 (as mentioned in Section 11.4). It will also be interesting to monitor any emerging trends in winter maximum extent of sea ice, not only along the coast of Greenland but also in northern coastal seas, such as the Baltic, the Sea of Okhotsk and the Weddell and Bellingshausen Seas to either side of the Antarctic peninsula. As changes to ecosystems are studied in response to changing ice cover, it will be important to monitor the detailed distribution of sea ice at finer spatial resolution, for which the growing availability of SAR images will make a valuable contribution. It will also be revealing to monitor SST, SSH, and ocean color where previously there has only been ice. It may be necessary in the Arctic Ocean, where solar elevation is quite low in late summer when the ice has melted, to improve chlorophyll retrieval from ocean color at low illumination levels.

Assessment of total ice volume also requires estimates of ice thickness which should be available from the synthetic aperture altimeter (SIRAL) instrument on the ESA Cryosat mission launched April 8, 2010. This coupled with sea ice coverage will allow estimates of freshwater fluxes associated with the ice.

Sea level

Sea level is an important ECV because it represents one of the main human impacts of climate change. Sea level is rising in response to increasing sea temperature and the melting of glaciers and continental polar ice sheets. As discussed in Section 11.5, it is not only average sea level rise that matters but also the spatial distribution of the increase. Low-lying coastal and island nations will be watching emerging scientific results with particular concern. This ECV is likely to be one that influences international political discussions about the protection and long-term future of coastal communities around the world.

Knowledge of the absolute dynamic topography will also be important for detecting changes in ocean circulation patterns, and for this the improved geoid expected from the recently launched GOCE mission will make a strong contribution. This will also facilitate better initialization of coupled ocean–atmosphere global circulation models for climate predictions.

Sea surface temperature

The recent progress towards internationally co-ordinated monitoring of SST as a result of the work of the GHRSST (discussed in Section 14.4.2), has prepared the ground effectively for delivery of high-quality SST datasets for climate application in the near future. Key to the high accuracy and stability of SST records needed for climate analysis is the ATSR series of satellites first launched in 1991. The most recent of these, the Advanced Along-Track Scanning Radiometer (AATSR) has well-validated accuracy and stability (see, e.g., Wimmer *et al.*, 2010) and work has

been done to harmonize data across all three ATSR class sensors from 1991 in order to generate a 19-year, climate quality global SST dataset (Merchant *et al.*, 2008).

The one weakness of ATSR data is the relatively poor spatiotemporal sampling resulting from its narrow swath. This is where the GHRSST infrastructure makes its most significant contribution by facilitating the use of ATSR data for bias adjustment of other SST datasets from a variety of different sensors. As outlined in Section 14.4.2 this has enabled new, global, high-resolution SST analyzed datasets to be produced, such as the U.K. Met Office's operational sea surface temperature and sea ice analysis—OSTIA (Donlon *et al.*, 2010). Such an approach allows the accuracy and stability of ATSR data to stabilize combined SST products from all available data which produces a better sampled dataset than is possible from ATSR alone.

The GHRSST Long Term Stewardship and Reanalysis Facility (LTSRF) based at the U.S. National Ocean Data Center, provides a repository for all the different SST datasets. This will facilitate iterative reprocessing of records from individual centers and their blending by reanalysis to deliver climate quality, high-resolution SST datasets. These promise more than existing climate SST datasets that typically have a coarse resolution of monthly averages at 1° latitude $\times$ 1° longitude, which was adequate for detecting SST global trends but did not reveal information about high-frequency features. Instead the new analyses will be resolved daily at 5 km to 10 km, which means they will contain information about mesoscale variability. This may open an opportunity to create climatologies of mesoscale processes such as fronts, eddies, or upwellings and how these are modulated over longer periods. Such climatologies can then be analyzed to determine whether they are sensitive to climatological indices (e.g., the North Atlantic Oscillation) which are known to characterize some aspects of ocean response to atmospheric forcing.

Continuity of the ATSR class of sensors will be maintained by a new SST sensor to fly on the ESA Sentinel-3 series (see Section 15.2.2) while wide-coverage infrared sensors for routine SST monitoring are planned to fly on polar and geostationary meteorological satellites for the foreseeable future. A possible weak link for sustaining a high-quality SST climate dataset is the uncertain continuity of passive microwave radiometers with frequency bands at 6–7 GHz and 10–11 GHz needed for SST. The Japanese GCOM-W1 satellite, expected in 2013, will carry an upgraded and refined AMSRE radiometer and provide some continuity over a number of missions. Although products from such a sensor have coarse spatial resolution, they measure through cloud. Thus not only do they make a valuable contribution to the analyzed SST data product during cloudy conditions, but should help to keep the SST record free of erroneous modulations caused by changed IR sampling associated with climatological change in cloud distribution.

Ocean color

The ocean color ECV is required to contribute to carbon cycle–monitoring requirements of the UNFCCC. Main derived products are the spectrum of normalized water-leaving radiance and depth-averaged concentration of chlorophyll-*a* (Chl)

in the surface photic zone of the ocean, although (as discussed in Chapter 7) measures of the diffuse light attenuation coefficient and estimates of PAR are also valuable for estimating total primary production in support of climate change monitoring.

Satellite observations from ocean color sensors provide the only means for global monitoring. Since no single sensor can cover the whole globe every day, and because cloud is a seriously limiting factor, data merging from several platforms and different types of ocean color sensor will be necessary. To achieve the long time span necessary for climate change analysis it will be necessary to patch together datasets from SeaWiFS (1997–2006), MODIS (2001–present), MERIS (2002–present), and future sensors such as Sentinel-3 OLCI and NPOESS VIIRS (see Section 15.2).

Blending data from different sensors to produce daily, high-resolution global maps of Chl and, if possible, applying optimal interpolation methods to fill time gaps caused by cloud cover, presents considerable challenges to ocean color scientists. Thorough calibration and validations of different sensor systems is essential if TOA visible waveband radiances are to meet the accuracy and stability requirements to serve as effective FCDRs. However, some progress has already been made by the ESA Globcolour Project,[10] which has produced a 10-year merged demonstration dataset of ocean color and chlorophyll products combined from SeaWiFS, MERIS and MODIS data.

Sea state

Sea state, representing the roughness of the sea surface, is considered to be an important variable for climate science because of the effect it has on air–sea fluxes (as discussed in Chapter 10). Meteorologists in the climate community have tended to assume that waves predicted by wave models from NWP forecast winds would be adequate for developing a climatology of sea state. However, as wave forecasters look increasingly to satellite data to help refine operational wave forecasting it is also being recognized that the 17 years of altimetry since the launch of TOPEX/Poseidon in 1992 offers a valuable directly measured record of sea state over the World Ocean. As explained in Chapter 8, altimetry yields a record of significant wave height, wave period, and possibly other wave statistical properties such as the PDF of wave height and wave age, which can be used in assessing air–sea flux parameters (as mentioned in Chapter 10). These can now provide an independent climate record of sea state, from which extreme wave height statistics may be determined, and also provide a means of evaluating the reliability of sea state climatology based on wave models. Continuity of altimeters is expected into the foreseeable future because of their requirement for operational ocean-forecasting systems.

Recent advances in measuring directional wave spectra using SAR point to the possibility of deriving global-scale records of wavelength and direction by this means, although it would be difficult to extend backwards a climatological record

[10] Consult the Globcolour website at *http://www.globcolour.info/*

of such data to earlier than 2002 when ASAR was launched on Envisat. Ongoing continuity of such data is expected from European plans for future SAR missions including the Sentinel-1 satellite series expected in 2012.

Sea surface salinity

There are at present no satellite-derived sea surface salinity (SSS) data products. At the time of writing the SMOS radiometer is in orbit but the performance of the L-band microwave radiometer is waiting to be evaluated before attempts are made to retrieve surface salinity measurements. In addition, NASA's Aquarius mission to measure SSS is due for launch in 2010. Between them it is likely that these two missions will achieve a capacity for coarse-resolution monitoring of SSS variability. Despite limited spatial and temporal resolution, if the target performance is reached it should provide more information than is presently available and may reveal new patterns of SSS variability that will contribute to our understanding of climate variations in the ocean. Even limited success for either or both missions will increase the demand for subsequent missions with improved specifications, since regular monitoring of SSS is essential information for understanding climate variability of both ocean circulation and the hydrological cycle.

14.7 REFERENCES

Alverson, K. (2005), Watching over the world's oceans. *Nature*, **434**, 19–20.

Alverson, K. (2008), Filling the gaps in GOOS. *Journal of Ocean Technology*, **3**(3, An Eye on Poseidon), 19–23.

Alverson, K., and D. J. Baker (2006), Taking the pulse of the oceans. *Science*, **314**(December 15), 1657.

Anderson, T. R. (2005), Plankton functional type modelling: Running before we can walk? *J. Plankton Res.*, **27**, 1073–1081.

Baretta, J. W., W. Ebenhöh, and P. Ruardij (1995), The European Regional Seas Ecosystem Model: A complex marine ecosystem model. *Netherlands J. Sea Res.*, **33**, 233–246.

Bell, M. J., R. M. Forbes, and A. Hines (2000), Assessment of the FOAM global data assimilation system for real-time operational ocean forecasting. *J. Mar. Syst.*, **25**, 1–22.

Brasseur, P. (2006), Ocean data assimilation using sequential methods based on the Kalman filter. In: E. P. Chassignet and J. Verron (Eds.), *Ocean Weather Forecasting* (pp. 271–316). Springer-Verlag, Dordrecht, The Netherlands.

Brasseur, P., P. Bahurel, L. Bertino, F. Birol, J. M. Brankart, N. Ferry, S. Losa, E. Remy, J. Schröter, S. Skachko *et al.* (2005), Data assimilation for marine monitoring and prediction: The MERCATOR operational assimilation systems and the MERSEA developments. *Quart. J. Roy. Meteorol. Soc.*, **131**, 3561–3582.

Brekke, C., and A. Solberg (2005), Oil spill detection by satellite remote sensing. *Remote Sens. Environ.*, **95**(1), 1–13.

Campbell, J. (1995), The lognormal distribution as a model for bio-optical variability in the sea. *J. Geophys. Res.*, **100**(C7), 13237–13254.

Chassignet, E. P., H. E. Hurlburt, O. M. Smedstad, G. R. Halliwell, P. J. Hogan, A. J. Wallcraft, R. Baraille, and R. Bleck (2007), The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *J. Mar. Syst.*, **65**(1/4), 60–83.

Chassignet, E. P., and J. Verron (Eds.) (2006), *Ocean Weather Forecasting*. Springer-Verlag, Dordrecht, The Netherlands.

Christian, J. R., M. A. Verschell, R. Murtugudde, A. J. Busalacchi, and C. R. McClain (2002), Biogeochemical modelling of the tropical Pacific Ocean. I: Seasonality and interannual variability. *Deep-Sea Res. II*, **49**, 509–543.

Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, and E. Fiedler (2010), The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA). *Remote Sens. Environ.*, submitted to AATSR Special Issue.

Donlon, C. J., P. J. Minnett, C. Gentemann, T. J. Nightingale, I. J. Barton, B. Ward, and M. J. Murray (2002), Towards improved validation of satellite sea surface skin temperature measurements for climate research. *J. Climate*, **15**(4), 353–369.

Donlon, C. J., T. J. Nightingale, T. Sheasby, J. Turner, I. S. Robinson, and W. J. Emery (1999), Implications of the oceanic thermal skin temperature deviation at high wind speed. *Geophys. Res. Letters*, **26**(16), 2505–2508.

Donlon, C. J., I. S. Robinson, K. S. Casey, J. Vazquez, E. Armstrong, O. Arino, C. L. Gentemann, D. May, P. Le Borgne, J.-F. Piolle *et al.* (2007) The Global Ocean Data Assimilation Experiment (GODAE) High Resolution Sea Surface Temperature Pilot Project (GHRSST-PP). *Bull. Am. Meteorol. Soc.*, **88**(8), 1197–1213, doi: 10.1175/BAMS-88-8-1197.

Drange, H. (1996), A 3-dimensional isopycnic coordinate model of the seasonal cycling of carbon and nitrogen in the Atlantic Ocean. *Physics and Chemistry of the Earth*, **21**(5/6), 503–509.

Durand, D., L. H. Pettersson, O. M. Johannessen, E. Svendsen, H. Søiland, and M. Skogen (2002), Satellite observation and model prediction of toxic algae bloom. *Operational Oceanography: Implementation at the European and Regional Scales* (Elsevier Oceanography Series Vol. 66, pp. 505–515). Elsevier.

Espedal, H. A., and O. M. Johannessen (2000), Detection of oil spills near offshore installations using synthetic aperture radar (SAR). *Int. J. Remote Sensing*, **21**(11), 2141–2144.

Espedal, H. A., and T. Wahl (1999), Satellite SAR oil spill detection using wind history information. *Int. J. Remote Sensing*, **20**(1), 49–65.

Ewing, G. C. (Ed.) (1965), *Oceanography from Space* (report of a workshop in 1964, 469 pp.). Woods Hole Oceanographic Institution, Woods Hole, MA.

Fasham, M. J. R. (1993), Modelling the marine biota. In: M. Heimann (Ed.), *The Global Carbon Cycle* (pp. 457–504). Springer-Verlag, Berlin.

Fasham, M. J. R., H. W. Ducklow, and S. M. McKelvie (1990), A nitrogen-based model of phytoplankton dynamics in the oceanic mixed layer. *J. Mar. Res.*, **48**, 591–639.

Ferraro, G., B. Bulgarelli, S. Meyer-Roux, O. Muellenhoff, D. Tarchi, and K. Topouzelis (2008), The use of satellite imagery from archives to monitor oil spills in the Mediterranean Sea. In: V. Barale and M. Gade (Eds.), *Remote Sensing of the European Seas* (pp. 371–382). Springer-Verlag, Berlin.

Fiscella, B., A. Giancaspro, F. Nirchio, P. Pavese, and P. Trivero (2000), Oil spill detection using marine SAR images. *Int. J. Remote Sensing*, **21**(18), 3561–3566.

Friedrichs, M. A. M., R. R. Hood, and J. D. Wiggert (2006), Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data. *Deep-Sea Res. II*, **53**, 576–600.

GCOS-Secretariat (2009), *Guidelines for the Generation of Satellite-based Datasets and Products Meeting GCOS Requirements* (GCOS-128, 13 pp.). World Meteorological Organization, Geneva, Switzerland.

GCOS (2004), *Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC* (GCOS-92, 136 pp.). World Meteorological Organization, Geneva, Switzerland.

GCOS (2006), *Systematic Observation Requirements for Satellite-Based Products for Climate: Supplemental Details to the Satellite-based Component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC"* (GCOS-107, 90 pp.). World Meteorological Organization, Geneva, Switzerland.

Gregg, W. W. (2001), Tracking the SeaWiFS record with a coupled physical/biogeochemical/radiative model of the global oceans. *Deep-Sea Res. II*, **49**, 81–105.

Griffies, S. M. (2006), Some ocean model fundamentals. In: E. P. Chassignet and J. Verron (Eds.), *Ocean Weather Forecasting* (pp. 19–73). Springer-Verlag, Dordrecht, The Netherlands.

Gunson, J., A. Oschlies, and V. Garçon (1999), Sensitivity of ecosystem parameters to simulated satellite ocean color data using a coupled physical–biological model of the North Atlantic. *J. Mar. Res.*, **57**, 613–639.

Hemmings, J. C. P., R. M. Barciela, and M. J. Bell (2008), Ocean color data assimilation with material conservation for improving model estimates of air–sea CO_2 flux. *J. Marine Res.*, **66**, 87–126.

Hemmings, J. C. P., M. A. Srokosz, P. Challenor, and M. J. R. Fasham (2003), Assimilating satellite ocean color observations into oceanic ecosystem models. *Phil. Trans. Roy. Soc. Lond. A*, **361**(1802), 33–39.

Hemmings, J. C. P., M. A. Srokosz, P. Challenor, and M. J. R. Fasham (2004), Split-domain calibration of an ecosystem model using satellite ocean color data. *J. Marine Sys.*, **50**(3/4), 141–179.

IOCCG (1999), *Status and Plans for Satellite Ocean-Colour Missions: Considerations for Complementary Missions* (edited by J. A. Yoder, Report No. 2, 43 pp.). International Ocean-Colour Coordinating Group, Dartmouth, Canada.

IOCCG (2007), *Ocean-Colour Data Merging* (edited by W. W. Gregg, Report No. 6, 74 pp.). International Ocean-Colour Coordinating Group, Dartmouth, Canada.

IPCC (2007), *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (edited by R. K. Pachauri and A. Reisinger, 104 pp.). IPCC, Geneva, Switzerland.

Maritorena, S., and D. A. Siegel (2005), Consistent merging of satellite ocean color data sets using a bio-optical model. *Remote Sens. Environ.*, **94**, 429–440.

Merchant, C. J., D. T. Llewellyn-Jones, R. W. Saunders, N. Rayner, E. C. Kent, C. P. Old, D. Berry, A. R. Birks, T. Blackmore, G. K. Corlett *et al.* (2008) Deriving a sea surface temperature record suitable for climate change research from the along-track scanning radiometers. *Adv. Space Res.*, **41**, 1–11.

Natvik, L.-J., and G. Evensen (2003a), Assimilation of ocean colour data into a biochemical model of the North Atlantic, Part 1: Data assimilation experiments. *J. Mar. Syst.*, **40/41**, 127–153.

Natvik, L.-J., and G. Evensen (2003b), Assimilation of ocean colour data into a biochemical model of the North Atlantic, Part 2: Statistical analysis. *J. Mar. Syst.*, **40/41**, 155–169.

O'Carroll, A. G., J. R. Eyre, and R. W. Saunders (2008), Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations. *J. Atmos. Oceanic Technology*, **25**, 1197–1207.

Ohring, G., B. Wielicki, R. Spencer, W. J. Emery, and R. Datla (2005), Satellite instrument calibration for measuring global climate change: Report of a workshop. *Bull. Am. Meteorol. Soc.*, **86**(9), 1303–1313.

Oschlies, A., W. Koeve, and V. Garçon (2000), An eddy-permitting coupled physical–biological model of the North Atlantic, 2: Ecosystem dynamics and comparison with satellite and JGOFS local studies data. *Global Biogeochemical Cycles*, **14**, 499–523.

Palmer, J. R., and I. J. Totterdell (2001), Production and export in a global ocean ecosystem model. *Deep-Sea Res. I*, **48**(5), 1169–1198.

Robinson, I. S. (2004), *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Robinson, I. S., D. Antoine, M. Darecki, P. Gorringe, L. Pettersson, K. Ruddick, R. Santoleri, H. Siegel, P. Vincent, M. R. Wernand *et al.* (2008), *Remote Sensing of Shelf Sea Ecosystems: State of the Art and Perspectives* (edited by N. Connolly, Marine Board Position Paper No. 12., 60 pp.). European Science Foundation Marine Board, Ostend, Belgium.

Siddorn, J. R., J. I. Allen, J. C. Blackford, F. J. Gilbert, J. T. Holt, M. W. Holt, J. P. Osborne, R. Proctor, and D. K. Mills (2007), Modelling the hydrodynamics and ecosystem of the North-West European continental shelf for operational oceanography. *J. Mar. Syst.*, **65**(1/4), 417–429.

Solberg, A., and C. Brekke (2008), Oil spill detection in northern European waters: Approaches and algorithms. In: V. Barale and M. Gade (Eds.), *Remote Sensing of the European Seas* (pp. 359–370). Springer-Verlag, Berlin.

Solberg, A. H. S., G. Storvik, R. Solberg, and E. Volden (1999), Automatic detection of oil spills in ERS SAR images. *IEEE Trans. Geosc. Remote Sensing*, **37**(4), 1916–1924.

Trieschmann, O., T. Hunsänger, L. Tufte, and U. Barjenbruch (2003), Data assimilation of an airborne multiple remote sensor system and of satellite images for the North and Baltic sea. Paper presented at *Proc. SPIE 10th Int. Symposium on Remote Sensing* (pp. 51–60).

Wimmer, W., I. S. Robinson, and C. J. Donlon (2010), Long-term validation of AATSR SST data products using ship-borne radiometry in the Bay of Biscay and English Channel. *Remote Sens. Environ.*, accepted for publication (AATSR Special Issue).

15

Looking forward

This final chapter aims to answer the question posed in the first chapter about how important the use of EO satellites really is to the science of oceanography. Section 15.1 summarizes the main discoveries about the character, phenomena, and processes in the ocean which can be attributed to remote sensing. It draws the conclusion that satellite data have become an essential part of marine science, with a key part to play in developing new opportunities for oceanographic research and applications.

Section 15.2 identifies those satellite observations and data-processing systems that are now considered necessary elements of oceanographic science. It specifies the minimum core of measurements needed so that scientific agencies can deliver reliable ocean-monitoring, forecasting, and climate services for the general public good. It looks forward to the future plans announced by space agencies to provide the space and ground segment infrastructure that will meet the observational requirements for delivering such services. It also identifies opportunities for new research based on novel satellite missions.

Finally Section 15.3 summarizes the intellectual challenges that are presently holding back methods of satellite oceanography from realizing their full potential. It also points to new ideas on the horizon, some of which can be expected to open up exciting new areas of research within a few years.

15.1 ACHIEVEMENTS

15.1.1 Oceanographic discoveries from satellite data

Because most ocean phenomena revealed by satellite image data in the earlier chapters of this book have been explored in detail by a combination of

remote-sensing and *in situ* observations, their existence and their typical horizontal spatial structure is taken for granted by oceanographers. Yet it would be very difficult to characterize ocean features by their shape if this had not first been revealed by the synoptic spatial overview given by satellite images.

This is undoubtedly true about the mesoscale features described in Chapters 3 to 5. It was the earliest satellite images that revealed strong meanders in western boundary currents, producing eddies and rings associated with localized frontal zones. In regions where large ocean currents collide, as in the confluence of the Brazil and Malvinas Currents or the Agulhas retroflection zone, satellites have revealed very strong and often complex eddy structures. The same is true of the Antarctic Circumpolar Current, where mesoscale eddies tend to be more elongated. Moreover the ability to identify mesoscale eddies by either thermal, color. or sea surface height signatures allows us to see how ubiquitous they are across the ocean while also revealing their different characteristics in relation to the distribution of major wind-driven ocean currents. Our knowledge of the distribution of mesoscale eddy energy would lack detail and be globally much less comprehensive if the methods of satellite oceanography were not available.

The same is true of the process of upwelling, an important mechanism for raising nutrients to the surface ocean in limited geographical regions, whose primary production can support the ecosystems of large oceanic areas. It is easy to take for granted our basic knowledge about where primary upwelling zones are found, but our understanding is enriched by the graphical way in which the distribution of wind forcing from scatterometer maps can be matched to cool upwelling zones in thermal images and then to zones of primary production revealed on color images. The same is true of more sporadic or seasonal upwelling features, such as the wind jets produced in the eastern tropical Pacific off central America, and the monsoon changes evident in the Indian Ocean. Moreover, wherever phytoplankton blooms occur in open ocean or shelf seas, satellite images provoke questions about what controls their location, their timing, and why they vary from year to year. There can be no doubt that satellite image data has greatly enriched and informed our understanding of these mesoscale and submesoscale processes. Few field experiments are performed to study such processes in detail without reference to coincident satellite data that provide the wider context for interpreting the *in situ* and subsurface observations of conventional oceanography.

If we consider larger scale dynamical processes in the ocean, there is another class of phenomena that depend even more strongly on satellite observations, that of planetary waves (as discussed in Chapter 6). In the case of Rossby waves it is clear that without satellite data there would still be very little firm evidence for their existence, whereas they have now been clearly revealed in satellite observations of several different surface ocean variables, and their speed has been measured with sufficient confidence and precision to provide evidence that corrections to theoretical models that describe their propagation characteristics were needed. Satellite data also provided the first evidence of tropical instability waves and have allowed their behavior to be characterized. Moreover, scientific study of the basin-wide structure of the El Niño–La Niña phenomenon in the equatorial Pacific has been

greatly assisted by the view from satellites, used in conjunction with *in situ* observations.

As the other chapters confirm, there are many more features and phenomena that benefit from insights and perspectives gained from a satellite's viewpoint and unique sampling capability. Put together, these examples leave no doubt that, over the last three decades, satellite ocean data have been responsible for informing and underpinning much mesoscale and large-scale dynamical oceanographic research and its connection with ocean primary production.

15.1.2 Does ocean science need remote sensing?

The history of how the technical capabilities of satellite oceanography have developed is one in which the technology was initially pushed by the space agencies, responding to political encouragement to promote the development of high-technology industries. However, political and funding motivation has changed so that investment in Earth observation satellites now needs to be justified by what the application of remote-sensing data contributes to societal benefits, environmental science, and ultimately to managing the health of the planet and its ecosystems. It is therefore appropriate to ask whether we need to continue to provide satellites to study the ocean, or can we simply make do with the knowledge gained over the last 30 years?

The conclusion to be drawn from the examples in this book is that satellite data have not only made a significant contribution to the advancement of ocean science, but continue to do so. As Chapter 10 on air–sea fluxes and Chapter 13 on shelf and coastal seas have shown, scientific advances need a combination of *in situ* and satellite observations. If we were now to take away access to any satellite data of the ocean, the quality of oceanographic research in certain fields would be degraded. Oceanography without satellite data, relying entirely on buoys, drifters, and experiments from ships, would lack the wider spatial and temporal context which can only come from the perspective of satellite data. We have between 10 and 25 years (depending on the variable) of global monitoring of the ocean from space complementing a wide variety of *in situ* observations within the surface and deep ocean. These data have already given us baseline climatologies that can underpin future *in situ* oceanographic research, and it is important to emphasize that those 10 to 25 years also confirmed the degree of variability attached to those climatologies. This variability is associated with secular trends and low-frequency variability of oceans, plus random fluctuations of mesoscale ocean turbulence. Together with ever-changing forcing by the turbulent atmosphere, this means that there would be an unacceptably high level of uncertainty if oceanographers were in future to assume the background ocean state to be that of a previously measured climatology, instead of continuing to use satellites and *in situ* measurements to tell us what is actually happening to the ocean in near-real time. So from a purely scientific perspective there is a strong argument that continuity of routine ocean monitoring has become essential to support future experimental oceanographic science.

The case for ongoing ocean monitoring from space is robustly reinforced when the requirements of operational applications of ocean satellite data are added to the needs of scientific research. In fact the type of operational applications elaborated in Chapter 14 provide a more compelling argument for establishing routine ocean monitoring from satellites. This is needed to underpin ocean forecasting and now-casting services in support of all aspects of our interactions with the ocean, having a direct impact on many sectors of human endeavor including industry, transport, commerce, health and safety, leisure, resource management, marine environmental protection, and political governance. Now that the needs of these diverse user sectors are seen to strongly overlap within operational programs such as GMES, where they are combined as "downstream user requirements", the value and efficiency of routine ocean monitoring from satellites, funded for the public good, becomes self-evident. Another strong justification for using satellites to provide timely and contemporary knowledge of ocean properties comes from the meteorological community where monitoring the ocean is also recognized to be essential for improved weather forecasting. The third major operational justification for satellite ocean monitoring is to monitor the changing ocean in order to characterize the oceanic response to anthropogenic climate change and to inform the UNFCCC's international endeavor to reduce greenhouse gases.

It should be recognized that one of the great successes of satellite oceanography, of even more importance than the high-quality scientific research that it has produced, is the way it has developed and demonstrated public benefits that can be derived from applying our scientific knowledge. In Europe the Marine Core Service of the GMES initiative and in the U.S.A. the Coastwatch[1] program under NOAA are examples of the emergence of applied oceanography as a substantial new activity. To be effective such services require an integrated approach to measuring the ocean from a variety of *in situ* and remote-sensing platforms, coupled to the use of ocean-forecasting models. Without the spatially detailed, regularly repeated, global-coverage ocean monitoring made possible by the methods of satellite oceanography, ocean-forecasting systems would be denied the quantity of observations that give them empirical credibility. For this the oceanography of the 21st century requires a fleet of ocean-monitoring satellites if it is to serve the needs of a large human population on a small planet whose surface area is 70% ocean.

15.2 SECURING THE FUTURE FOR OCEAN REMOTE SENSING

15.2.1 Essential satellite oceanography

Having established that there is an ongoing need for routine satellite measurements of ocean variables for the foreseeable future, it is important to summarize the specific requirements, reviewing what has been discussed in the main chapters of this book.

[1] For further information see *http://coastwatch.noaa.gov/cwn/index.html*

Evidence of the experience from recent years, including oceanographic applications described in previous chapters, points to obvious standard variables that oceanographers need to be measured from satellites. These are sea surface height from which ocean surface currents can be derived, based on altimetry; sea surface temperature, based on infrared and microwave radiometry; ocean colour, the spectrum of water-leaving radiance or reflectance sampled in discrete bands across the visible waveband, from which chlorophyll-*a* concentration and the diffuse attenuation coefficient are derived; sea surface wind vectors from scatterometry and possibly from polarimetric microwave radiometry; sea ice parameters from imaging radars, passive microwave radiometry, thermal and optical measurements; sea state including wave height from altimetry and directional wave spectra from imaging radars; and sea surface roughness images from SAR, capable of being analyzed automatically to detect oil slicks and surface-dynamical features such as internal waves.

Table 15.1 lists these variables along with the desirable spatial and temporal resolution at which they are required for general operational applications, and an estimate of the provision of satellite/sensors needed to meet these minimum operational specifications. The implied products are global level 3 or level 4 analyzed datasets; the level 2 data products of individual sensors from which these datasets are constructed will need to be more finely resolved spatially. In addition to the ground segment provided by agencies for ocean data products from individual sensors, there needs to be a globally integrated processing infrastructure for producing these combined products in near-real time, along with error statistics and other ancillary data.

In specifying the minimum requirements for global ocean monitoring from satellites, it should be emphasized that the operational applications served by these satellite-derived data products are also dependent on a comparable set of *in situ* data. These include moored and drifting buoys, Argo floats and gliders, without which the capacity of ocean-forecasting systems to represent the interior dynamics and physics of the ocean would be seriously compromised. It is therefore essential that the requirements for satellite and *in situ* data are treated as a whole. They are complementary, their combination is synergetic, and it would be mistaken to treat them as alternatives and therefore in competition for the same funding. Moreover the quality of satellite data products is dependent on ongoing validation using appropriate sets of *in situ* measurements. It is important to recognize that monitoring the performance and uncertainties of each satellite instrument and its derived products, by maintaining a long-term comprehensive validation program throughout the lifetime of a sensor, is to be treated as an essential component of routine data delivery for operational applications.

15.2.2 Limitations of existing sensors and platforms

The data products identified in Table 15.1 represent minimum required specifications. It has already been noted in previous chapters that research and applications are being hindered by limitations of present sensor specifications.

Table 15.1. Ocean variables required to be monitored operationally from satellites.

Ocean variable	*Desirable specification (variable accuracy, spatial and temporal resolution) of analyzed products*	*Minimum complement of sensors needed to deliver the minimum specification*
Sea surface height	2 cm; 1/3° lat. × 1/3° long.; weekly	1 T/P class altimeter and 1 LEO altimeter
Sea surface temperature	0.1 K; 5 km; daily	1 ATSR-type dual-view IR sensor, 2 met-type LEO IR sensors, 1 GEO IR sensor, and 1 LEO MW sensor
Ocean color	5% reflectance, K, and 0.3 log(Chl); 5 km; weekly	2 multispectral VIS-NIR radiometers like MERIS or MODIS
Wind vector	1 m/s, 10° direction, 25 km, daily	2 LEO scatterometers
Sea state: SWH	10 cm; 10 km along-track	1 T/P class altimeter and 1 LEO altimeter
Sea state: directional wave spectra	10 cm in height and 15° in direction; samples at 200 km spacing; daily	4 SARs in LEO
Surface roughness	75 m; 2 days	4 SARs in LEO
Sea ice parameters	<5% concentration <1 km edge <1 K temperature Daily at <10 km resolution	Passive microwave imager (SSM/I or AMSR class), SAR, thermal and optical imagery, scatterometer, and SAR altimeter data

There are situations where there is a clear need for satellite measurements of a certain accuracy or resolution which cannot be provided because remote-sensing methodology is not yet adequate. The outstanding issues can be summarized as follows.

The measurement of sea surface slope using altimetry for detecting ocean currents is limited in several ways (see Section 3.4.2):

- If SSHA could be measured more reliably in shelf seas and closer to the coast, there would be considerable benefits to shelf sea dynamics and the study of coastal upwelling (see Sections 5.1 and 13.3)
- The large spacing between adjacent altimeter tracks severely restricts the capacity to map synoptic mesoscale details of ocean surface currents and to quantify eddy kinetic energy (see Section 3.4.4). This could be remedied partly by the use of swath altimeters which map SSHA in detail over a swath of up to 200 km, and partly by using a co-ordinated constellation of altimeters.
- Measuring absolute currents from the retrieval of ADT as well as SSHA requires

improved independent knowledge of the geoid. The use of GRACE data has already helped (see Section 2.4.5), but further progress is expected based on the analysis of data from the recently launched GOCE mission.

The sampling frequency for updating measurements of surface winds and waves is limited by the small number of sensors in orbit. A more complete view of the variability of sea state and a better sampling of extremes requires a much higher sampling rate, implying the need for constellations of satellites.

The specification for ocean color measurements is based on requirements for monitoring the open ocean. The scope of applications of ocean color data would be greatly enhanced by attending to the major outstanding issues:

- The problem of retrieving reliable ecosystem information from ocean color sensors in Case 2 optical conditions has not yet been satisfactorily solved (see Sections 13.2.6 and 14.3).
- Studies of coastal and estuarine ecosystems demand fine-resolution ocean color remote sensing down to scales of a few meters (see Section 7.5). There is a growing need to provide a global capability for mapping local regions, which may need a rethink of how to produce the required infrastructure of platforms, sensors, and data management.
- Even where present sensors are effective in monitoring the open ocean, cloud cover remains a serious problem. The combination of few sensors and random cloud cover seriously degrades actual sampling frequency. It is worth exploring ways to maximize opportunities to view the ocean when it is cloud-free, using a constellation of sensors or experimenting with sensors on geostationary platforms.

The methods for monitoring SST are satisfactorily evolving through the GHRSST activity which has facilitated the synergetic use of different sensor types, although it remains a limitation that satellite thermal microwave radiometry has such a coarse spatial resolution and needs to be better co-ordinated in terms of a sustained space segment. It is clear from Sections 13.2.4 and 14.4.2 that there will be benefits for shelf sea oceanography if microwave radiometers could deliver better resolution and reliably measure close to the coast.

The methodological limitations outlined above represent a set of technology and system challenges for satellite oceanographers to solve. They have a high priority because there are applications of ocean satellite data already being held back until these problems are remedied. However, an even higher priority is to make sure that the well-established baseline of core ocean measurements is secure for the future (as discussed in the next section).

15.2.3 Future sensors, platforms, and systems for observing the ocean

It is important to review the ocean-observing sensors that are expected to be launched in the next decade, since without a reliable ongoing program of satellites

the health of research in satellite oceanography would be under serious threat. As things stand at the end of 2009, the prognosis for ocean-viewing satellites for the next decade looks to be generally secure, with an adequate provision of sensors presently in operation to meet the baseline needs expressed in Table 15.1. These are the sensors listed in the tables in Chapter 2. Equally important is that some space agencies are already committed to preserving continuity. Several satellite systems are currently under construction with the aim of putting replacement sensors in orbit in time to allow an overlap period for intercalibration before current sensors fail. Nonetheless there is some risk that by the middle of the decade there will be less redundancy than at present when several similar sensors from different agencies are measuring the same variable. A certain amount of redundancy is essential to underpin a robust operational service of ocean data delivery from satellites. Looking forward 5 to 10 years it appears that, without such redundancy, any unexpected premature failure of a sensor could cause a temporary loss of data supply.

There is a bewildering array of EO satellites in orbit just now, with over 100 sensors in operation at the present time, serving a wide variety of global environmental measurement needs. Around one-third of these are capable of contributing useful information for ocean science, although only a handful have ocean monitoring as their primary task. The Committee on Earth Observing Satellites[2] (CEOS) is responsible for co-ordinating the programs of all the different space agencies. CEOS has produced the *Earth Observation Handbook* which is now maintained online.[3] The information it contains on sensors needs to be critically examined because it is not always easy for the reader to distinguish between state-of-the-art sensors geared up to deliver near real–time data products of high quality, and sensors with an apparently similar function that are in fact part of an agency's development and experience-building program but their data are not freely published. However, there is a useful section in the *CEOS EO Handbook* which provides a catalog of satellite missions, distinguishing operational missions from the rest.

Meeting the requirements of Table 15.1

Operational meteorology missions include an ongoing commitment to wide-swath infrared sensors in LEO, to be provided through the Eumetsat MetOp satellite series and the planned NPOESS program of the U.S.A. There is also a commitment to GEO IR sensors spaced around the world. Within Europe, the GMES Sentinel series provides dedicated operational continuity of many ocean-relevant instruments including ocean color, temperature, altimetry, and SAR. These LEO and GEO sensors will partially meet the baseline need for SST, providing Classes 1 and 3 (in Table 14.1) towards datasets mentioned in Table 15.1. The dual-view sensor needed for quality control (Class 4 in Table 14.1) will be provided by a succession of new sea and land surface temperature radiometers (SLSTRs) on ESA's Sentinel-3 satellite series. The need for a microwave sensor (Class 2 in Table 14.1) will be met by the

[2] The CEOS website is at *http://www.ceos.org/*

[3] The *Earth Observation Handbook* is to be found at *http://www.eohandbook*

AMSR-2 instrument on Japan's Global Change Observation Mission (GCOM) to be launched in 2012. However, unlike other thermal sensors this does not yet appear to have a backup.

For ongoing altimetry there is a commitment to continue the TOPEX/Poseidon class altimeter in the Jason satellite series, and the RA-2 on Envisat will be replaced by the SRAL altimeter on Sentinel-3. In addition, the SIRAL altimeter launched in 2010 on ESA's CryoSat mission to monitor polar ice thickness should also be available to fill the gap before Sentinel-3 is launched, should the RA-2 fail prematurely. The French agency, CNES, also intends to launch a Ka-band altimeter (AltiKa) on the Indian satellite Saral by the end of 2010. These additional sensors should help to ensure that up to four platforms deliver altimeter data, which is better for operational ocean-forecasting models than using only two, although the picture for the latter part of the next decade is less favorable for additional altimeters.

The sea state requirements for measuring SWH into the future can be met by the planned altimeters mentioned above. The need for wind vector measurements should be met by the ASCAT sensor on the MetOp series of European meteorological sensors, although there seem to be no firm plans for other operational scatterometers in the next few years, which implies some risk. The requirements for sea surface roughness measurements from SAR, needed for retrieving directional wave spectra as well as oil spill detection, are likely to be met by plans for a growing number of SARs through the next decade, including a new C-band SAR on Sentinel-1, as well as PALSAR, TerraSAR-X, and Radarsat mentioned in Section 14.5.

Finally, considering ocean color requirements, despite the importance of improving coastal ecosystem measurements in Case 2 conditions, it seems that continuity from MODIS will be lost when it is replaced by the VIIRS instruments on the NPOESS platforms. This has only the basic wavebands needed for open-ocean chlorophyll retrievals. However, ESA will sustain the MERIS heritage with the Ocean and Land Colour Instrument (OLCI) on Sentinel-3, extending its functionality with additional wavebands. Between them, OLCI and VIIRS will just about meet the baseline need for ocean color monitoring stated in Table 15.1.

Operationally targeted satellite series

It is noteworthy that the continuity needed for ongoing ocean monitoring will be guaranteed by a combination of operational meteorological sensors (MetOp, NPOESS, and the GEO platforms), the Jason program which is now largely underpinned by a federation of operational users co-ordinated by EUMETSAT, and the ESA Sentinel program. This confirms that the driving motivation is to support operational applications of satellite ocean data. This is clearly the case for the Sentinel series of satellites, developed by ESA to deliver the satellite measurements needed by the European GMES initiative. Ultimately the GMES Space infrastructure is expected to be funded by the European Union on behalf of all European users of the environmental data products it produces (as discussed in Section 14.2.1). The scope of the Sentinel program, with five different satellite types, is far wider than ocean applications alone. However, the purpose of Sentinel-3 (Aguirre *et al.*, 2007) is

to ensure that the data requirements of the GMES Marine Core Service are adequately met. For this reason it will carry a radar altimeter (SRAL), an ocean color sensor (OLCI), and a surface temperature sensor (SLSTR). Note that SLSTR was based on the heritage of the dual-viewing AATSR sensor on Envisat, in response to the request by GHRSST that this type of sensor was needed to provide high quality of accuracy and stability rather than coverage.

While Sentinel-3 was designed to fulfill the main ocean requirements its sensors also meet the requirements for coarse-resolution land mapping. In a comparable way, Sentinel-1, which carries an X-band SAR mainly for land mapping and interferometry, contributes to the ocean-monitoring requirement for sea surface roughness, and Sentinel-2, which will carry a fine-resolution, land-mapping visible waveband sensor, has the potential to contribute to monitoring of coastal and estuarine waters.

Explorer missions

ESA has also produced several satellites in its Earth Explorer series which will be able to contribute to ocean monitoring, although with no commitment to sustained data delivery. The Gravity and Ocean Circulation Explorer (GOCE) mission was launched in 2009 in order to measure the gravity field above the Earth and hence deduce the ocean geoid to a higher resolution than hitherto (see section 11.7.3 of *MTOFS*—Robinson, 2004). It is expected to lead to more reliable estimates of the ADT and hence the retrieval of absolute geostrophic surface currents of the sea.

As mentioned above, Cryosat was launched in early 2010, carrying the SIRAL altimeter which will be used to map aspects of polar ice sheets. It also has the capacity to operate in a mode that can retrieve SSH where this does not compromise its primary cryospheric mission. Since the Sentinel-3 altimeter is expected to be very similar, there may be the opportunity to use data from SIRAL to develop novel processing schemes for ocean echo waveforms using the synthetic aperture capabilities of SIRAL.

Finally, the Soil Moisture and Ocean Salinity (SMOS) mission was launched in late 2009 and this L-band passive microwave radiometer has just started to deliver its first data. It remains to be seen how finely it will be able to resolve sea surface salinity, SSS, but since it is the first sensor to attempt such measurements from space its results will be of great interest to the wider oceanographic community as well as to satellite oceanography specialists. It will be followed by the comparable NASA Aquarius mission.

The combination of these three ESA Earth Explorer missions, with the prospect of a decade of operationally reliable ocean measurements to come from the Sentinel series, added to the range of existing ocean sensors, holds out the promise that satellite oceanography research will be as exciting in the second decade of the new millennium as it was in the first.

15.3 CHALLENGES FOR SATELLITE OCEANOGRAPHERS

In the end it is factors such as a burning curiosity to understand how the world works and the creative drive to solve theoretical or experimental problems that propels scientists towards success. Therefore it seems appropriate to conclude this book with a reminder of some of the outstanding challenges and opportunities that remain for creative scientific initiative in the field of satellite oceanography. These are some of the problems that are likely to attract and motivate the next generation of scientists in this field. They are simply listed under a number of categories. A simple phrase or a few sentences should be enough to serve as a reminder of what is discussed in more detail elsewhere in the book.

Newly emerging remote-sensing techniques or methods urgently needing improvement

- New coastal altimetry methods: improvement of altimeter signal processing and corrections to improve the accuracy of altimeter data in shallow seas and close to the coast, opening up important fields of research in dynamics of shelf seas and coastal currents (see Section 13.3).
- Swath altimetry—exploring and developing techniques for off-nadir altimetry providing high-resolution, two-dimensional mapping of SSHA.
- Higher resolution passive microwave imagery for better SST, salinity, sea ice parameters, and surface wind over the ocean.
- Use of novel techniques to retrieve ocean currents using SAR and SAR interferometer systems.
- Better and more complete specification of uncertainty estimates for all satellite ocean measurements.
- The use of global navigation satellite system reflectometry (GNSS-R). "Signals of opportunity" from GPS are reflected from the sea and received by a detector in low-Earth orbit. Comparison between the direct and reflected signal has the potential to carry information about surface height and sea state at the points of reflection. If the technique is feasible there is scope for a constellation of small cheap satellites to monitor ocean properties at quite high sampling densities (see, e.g., Clarizia *et al.*, 2009).
- Exploiting the new gravity data from GOCE to improve mapping of absolute dynamic topography from satellite altimeter data.
- Developing systems to exploit altimetry records containing evidence of tsunamis, in conjunction with conventional methods of tsunami warnings and prediction (see Section 11.5.4).
- Evaluating and improving the retrieval of sea surface salinity from SMOS data.
- Improving the prediction of diurnal thermocline structure for on-the-fly conversion of skin SST or subskin SST to foundation SST (see Section 14.4.2).
- Improving the retrieval of seawater properties from ocean color remote-sensing data in optically complex (Case 2) waters.

Making the most of new satellite data-merging tools and operational oceanography systems

- Explore novel ways of assimilating satellite ocean data into ocean-forecasting models that respect the character of the data.
- Better use of complementary satellite measurements in synergy.
- Developing novel techniques for constraining ecosystem models with satellite ocean color data, using the best knowledge of inherent optical properties.
- With the production of new merged satellite data products (e.g., new, high-resolution, analyzed SST products resulting from GHRSST), there is a need to appraise their content, to assess how closely they represent the real ocean and whether different optimal interpolation schemes degrade or enhance genuine features.
- Develop and analyze new climatologies based on variability signals from high-resolution, analyzed satellite data products. Does this offer insights into climate change in the ocean?
- Use of new, high-resolution maps of SST and surface winds from satellite data to evaluate high-resolution coupling between the ocean and atmosphere.
- Interactive use of satellite-derived wave data (from altimetry and SAR) to evaluate and improve the performance of wave-forecasting models.

New oceanographic questions and problem solving that exploits the latest satellite data products

- Study the stability of ocean fronts using new surface velocity retrievals derived by SAR Doppler centroid methods.
- Exploit the latest satellite and *in situ* datasets to improve accuracy and spatial resolution of air–sea fluxes and flux climatologies.
- Use high-resolution satellite data to explore the impact of space-time resolution on the accuracy of globally integrated fluxes.
- Investigate the possibility of Rossby wave interactions with western boundary flows, leading to potential modulation of ocean circulation.
- Connecting satellite-derived climate data records with historical *in situ* measurements to provide the best climate datasets for use in climate research.

These are just a few examples of the type of research questions that arise as methods of satellite oceanography improve and deliver higher quality or better resolved ocean data. My hope is that by the time young scientists reach this stage of the book they will already have formed ideas and questions of their own that may be quite different from these. If so I wish them success in following them up, and look forward one day to reading the outcome when they are published. Then perhaps someone else will come along and compile the results from another generation of satellite oceanographers into another book like this!

15.4 REFERENCES

Aguirre, M., B. Berruti, J.-L. Bezy, M. Drinkwater, F. Heliere, U. Klein, C. Mavrocordatos, P. Silvestrin, B. Greco, and J. Benveniste (2007), Sentinel-3: The ocean and medium-resolution land mission for GMES operational services. *ESA Bulletin*, **131**(August 2007), 24–29.

Clarizia, M. P., C. P. Gommenginger, S. Gleason, M. A. Srokosz, C. Galdi, and M. Di Bisceglie (2009), Analysis of GNSS-R delay-Doppler maps from the UK-DMC satellite over the ocean. *Geophys. Res. Letters*, **36**(L02608), doi: 10.1029/2008GL036292.

Robinson, I. S. (2004), *Measuring the Ocean from Space: The Principles and Methods of Satellite Oceanography* (669 pp.). Springer/Praxis, Heidelberg, Germany/Chichester, U.K.

Index

AATSR 58, 364, 577, 581
absorption, optical 33
achievements of ocean remote sensing 607–610
action balance equation 463
Aden, Gulf of 99–100, 101–102
ADEOS-1&2 56, 415, 547
advection, role in shelf sea temperature distribution 491, 492, 497, 498, 500, 502
aerial photography 533
aerosols 32
Agulhas Current 70–71, 87–96, 118–119, 124, 126–127, 135, 137
Agulhas retroflection 71, 85, 86, 87
Agulhas rings 85, 96, 135
air pressure 362
air–sea fluxes 359–386
  CO_2 563, 590
  gas 362, 372–377
  heat 362–363, 378–381, 590
  measuring from space 370–382
  momentum 294, 362
  parameterizations 360, 362–363, 601, 618
  theoretical principles 361–362
  satellite data relevant for 363–370
air–sea interaction 223–224, 231, 233, 294, 340, 347–350, 393, 421, 425
air temperature 362, 363, 369, 378, 381
air temperature patterns of ENSO 393–394
airborne remote sensing 530, 532
AIRS 370
Alaska, Gulf of 85
Alboran Sea 251
albedo 427
algal bloom monitoring in shelf seas 519–523
along-slope current 492, 493, 494
ALOS 587
AltiKa 615
altimeter applications to:
  air–sea fluxes 231, 365–366
  El Niño 407–410
  fronts 124–126
  internal waves 459, 479
  large-scale dynamic processes 202–203, 210–211
  mean sea level 437–441
  mesoscale eddies 80–89
  monsoon 422–425
  shelf seas 318–319, 327–328, 485, 488, 491, 524–527
  storm surges 441–442
  surface currents 412–415
  surface waves 309–312, 317, 323–326
  tsunamis 442–444
  upwelling 165–166, 176
  wind over the ocean 336–337, 348–349
altimeter
  constellation 327,442, 443

altimeter (*cont.*)
  pulse-limited footprint 298
  pulse reflected from rough sea 298–299
  retracking 527–528
  significant wave height measurement 298–300
  estimate of wave period 300
altimeter-derived wave climatologies 323–324
altimetry 46–50, 60, 86–87, 435
  improved processing for coastal seas 527–528
  problems in coastal seas 523–524
Amazon River plume 178–179, 249, 252
AMI 61 334, 411–412
AMSR-E 45, 59, 338, 368, 577, 599
AMSR-2 615
AMSU-A 369
analyzed datasets *see* data products, analysed
analysis winds 340, 341, 342, 343
anchovy 269, 271
Andaman Sea 458, 469
anomalies 59, 197–200, 249–250, 269–270, 280, 405, 406, 408, 416–417
Antarctic Circumpolar Current (ACC) 70–71, 72–73, 87, 89, 126, 143, 148–153, 186–187, 233, 243
Antarctic circumpolar waves 233–234
Antarctic coastal waters 267, 283
Antarctic sea ice cover 427–430, 432, 433
anticyclonic eddy 76–78, 92, 102, 103
Apel, John 462–463
applied satellite oceanography 539–543, 547–9
Aqua satellite 45, 56, 370, 547
Aquarius satellite 45, 602 616
aquaculture *see* marine aquaculture
Arabian Sea 80–81, 84, 102, 422–426
Arctic ice cover 433–434, 599
Argentine Basin 83, 90
Argo floats 553
artificial neural network (ANN) 369, 515, 518
ASAR 61, 304, 312, 313–316, 319, 322 427, 466, 468, 469, 584–585, 586
ASCAT 81, 334, 336, 615
assimilation
  altimeter data 442, 553
  ocean color data 555, 558–564, 618
  sequential 551–554, 562–564, 570, 618
  SST data 553
  wave models 321
ATBD 60
Atlantic meridional overturning circulation 427
Atlantic Ocean 204, 216, 220, 226, 227, 229–231, 232, 248
Atlantic Ocean, north 243, 249, 250, 252, 264
Atlantic Ocean, northeast 243, 245, 251, 494
Atlantic Ocean, northwest 241, 243, 269, 270
Atlantic Ocean, nouth 247, 304
Atlantic Ocean, subtropical 252
atlas of internal wave features 462
atmospheric
  boundary layer 95, 106, 133, 342, 359, 363, 365, 366, 369, 370, 378, 380
  boundary layer stability 365, 366, 385
  correction
    altimetry 46
    generic 19, 20
    ocean color 32, 566–568
    infrared 37–39
  internal waves 454–455
  pressure 47
  remote sensing 359
  transmission 28–29
  windows 29
attenuation coefficient, diffuse optical (K_d) 34, 564, 565–568
ATSR 36, 38–39, 41, 58, 407
AVHRR 36, 38, 41, 58, 71, 364, 399, 404, 405, 424, 493, 495–497, 499, 504, 577
AVNIR 275

β plane 208
background error 553–554, 558
bacteria 556
Baltic Sea 441, 486, 521–522
Barents Sea 99, 183
baroclinic flow 74–75, 208, 210–211
barotropic flow 74–75, 442
bathymetry measurement with SAR 515–516

Bass Strait 487
beach management 520, 528, 530, 532–533
BEAM 60, 66
beams of internal wave energy 455
benchmark in the sky 435–437
Bengal, Bay of 85, 422
Benguela Current 70, 71, 162–165, 168–170
Benguela ecosystem 253
Bering Sea 143, 146, 431, 434, 486
BEST 60, 66
Bilko 60, 66
biogeochemical models 240, 549 ,556–557, 559 (*see also* ecosystem models)
biogeochemical cycling 253
biogeochemical fluxes 253, 282
biogeographical province 153–154 252, 253, 259, 260–261, 264
Biscay, Bay of 469, 471–478, 494, 498
black body 36
blooms, phytoplankton 174, 178, 182–3, 184–187, 488, 489, 494, 514, 519, 520, 521
bloom timing 187–188
bottom-mounted pressure gauges 443
Bragg scattering 51
Bragg-resonant waves 301, 334, 460, 463, 464, 467, 584
Brazil Current 83, 89, 90, 91, 146
Brazil–Malvinas confluence zone 120, 146, 153–155, 253
"bright pixel" atmospheric correction adjustment 509
Bristol Channel 531
Brunt–Väisälä frequency 455
bulk domains at air–sea interface 361
butterfish 271
buoyancy frequency 455

C-band radar 50, 61, 312–313, 334, 366, 464, 466–467, 585, 615
Cadiz, Gulf of 470
Caicos Bank 77, 278
calibration
  data product algorithms 22
  sensors (generic) 18–19, 444
California Current 146, 171, 173, 251, 261
Canary Coast (northwest Africa) 170–171
Canary Islands 180
Cape Cod (Massachusetts) 454, 458, 465, 466, 479
capillary waves 293, 583
carbon-based primary production model 261–262, 266–267
carbon dioxide CO_2 360, 362, 369, 370, 372, 373, 376, 377
Cardigan Bay Front 504, 505
Caribbean Sea 251, 277
Case 2 challenge in shelf seas 517–519
Celtic Sea 251, 267, 468, 486, 493–495, 497–501, 504–506, 508, 513
Celtic Sea Front 504–506, 508, 513
CEOS 446, 447, 614
CDOM 127, 517, 518
challenges facing satellite oceanography 485–486, 517, 523–524, 530, 617–618
Channel Islands, California 316
Charnock constant 366–368, 378
Chattonella spp. 561
Chemical energy equivalent of fixed carbon (J_C) 259
Chlorophyll-*a* 33, 34, 126–127, 560
Chlorophyll
  anomalies 249–250, 416–417
  concentration (C) 256, 257, 258, 259, 260
  global distribution 246–250
  patchiness 184, 242, 251, 560
Chlorophyll-specific light absorption coefficient (a_c^*) 256, 257
CHRIS 273, 275
circulation, ocean 143, 233, 548, 551, 562, 587, 590, 599, 602, 618
classification of sea floor type 273
CleanSeaNet 586–588
climate 589
  change 223–224, 249, 360, 385, 540, 588, 610, 618
  forecasting 223, 227
  monitoring 588–601, 610
  role of the ocean 223–224, 589–591
climate variables (*see also* essential climate variables)
  ocean color 597, 600–601
  ocean salinity 597, 602
  sea ice 597–599
  sea level 597, 599

climate variables (*cont.*)
sea state 322–326, 597, 601–602
SST 597, 599–600
climatologies 59, 190, 199–201, 247–249, 323–326, 385, 597–601, 618
cloud cover, problems of 96, 103, 123–124, 407, 491, 494, 507, 513–514, 520, 549, 555, 574, 584, 601, 613
cloud detection 19, 39–40
CMIS 338
CMOD4 wind–backscatter model 336
CMOD5 wind–backscatter model 336
COARE algorithm 367, 368, 380, 385
coastal altimetry 523–528, 612, 617
demand for 524
potential applications of 525–526
coastal applications of remote sensing 532–534
coastal ecosystems 277, 279, 282, 533, 613
coastal remote sensing 528–534
coastlines 435
coccolithophore lookalikes 255
coccolithophores 99, 101, 129, 254, 255, 494, 514
colored dissolved organic material (CDOM) 34
Columbia River 178
Compact Airborne Spectral Imager (CASI) 277
conical scan 38–39
consolidated data 65
Congo River 252
convection in atmosphere over warm sea 344, 347, 395, 406, 411
convergent/divergent surface currents 103–104, 106–107, 117–118, 131, 180, 460
coral bleaching 279–282
Coral Reef Watch 280, 281
coral reefs 272, 277, 279–282
Coral Sea 248
corals 279
Coriolis satellite 56, 338
Coriolis force 73–74, 116–117
Coriolis parameter 49, 75–76, 196, 455
crest-tracking of Rossby waves 222
Crozet Island 186–187
CROZEX 186–187
Cryosat 599, 615
cryosphere 590
currents *see* ocean surface currents
current jets 149–150, 153
current shear 227, 231
cusp formations on images 228–229
cyclonic eddy 76–78, 92, 103
CZCS 57

Dalton number 363, 378, 379, 380
data cube 201–202, 217–218
data formats 60, 579–581
data merging 197–200, 445, 570–582, 601
data processing, general 7–8, 16–28, 197–200, 547–550, 569–582, 611
data products
absolute dynamic topography (ADT) 573
aerosol optical depth 579–580
altimetry 49, 547, 549, 551, 553, 573, 611–612
analyzed 572, 576, 581
ancillary data fields 579–580
anomalies 59, 197–200, 249–250, 403, 444
chlorophyll 549
climatologies 59, 246–249, 403, 444
composite 26–27, 41–42, 45 197–200, 246, 248, 444, 70–572
confidence values 27, 60, 554, 579–580
error statistics 22, 444, 554, 560, 570–572, 579–580, 582
generic 55, 57–60, 65–66, 547–549, 551
geophysical derivation 21–22
GHRSST L2P 365, 579–582
levels 0 to 4 17, 18, 26–27, 57, 65, 242, 547, 572
matchup dataset 22, 560–561
merged 445, 570–582
metadata 60
meteorological 45
navigation 23–26, 60
near-real time 65, 545
ocean color 34–35, 241–251, 549, 551, 555–568, 573–575, 611–612
sea ice 45, 428–431, 611–612
sea surface height anomaly (sea level anomaly) 49, 81, 82, 84, 573
significant wave height 49, 309–311, 573, 611–612

SST 41–42, 547, 549, 551, 553, 570, 575–582, 611–612
surface currents 412–415
validation 22–23, 65, 444, 547
versions 65
waves and spectra 313–316
wind speed 45, 49, 341–343, 543, 579–580, 611–612
data selection criteria 57–60, 62–63, 65, 570–572
deep chlorophyll maximum (DCM) 475–477
degree heating weeks 280, 281, 282
delay-Doppler altimeter 528
depth profile of chlorophyll 475–477
detritus 556–557
diagnostic dataset, high resolution (HR–DDS) 582
dielectric constant 43, 337
diffuse attenuation coefficient for light (K) 474–475
diffusivity, of a gas 373
dimethylsulfide (DMS) 183, 253, 254, 253
direction, internal waves interpreted from images 463
disaster monitoring 446–447, 531–532
Disaster Monitoring Constellation 531
dispersion relation, ocean surface waves 295
dissolved organic nitrogen (DON) 556
diurnal variability 364, 571, 578–9
diurnal warming 40, 41, 364–365, 578
DMSP 55, 338, 428
Doppler centroid analysis 111, 134–135, 137, 618
Doppler frequency shift 53, 135
Dover Strait 441, 500, 511
drag coefficient 362, 378
drawdown of CO_2 377
DUACS 49, 84, 445, 573
dual-frequency altimetry 366, 376
dual view 39
dynamic topography *see* topography, ocean dynamic

early ocean images 70
Earth Explorer missions 616
Earth system science 541, 589
East Atlantic Pattern (EAP) 326
East China Sea 145, 147, 486
ECMWF analysis winds 340
ecosystem models *see* marine ecosystem models
ecosystems observed in shelf seas 516–523
eddies, ocean 15, 45, 69–110, 176, 553
eddy detection
ocean color 73, 79, 98–102
radar images 79–80, 103–111
sea level anomaly 78, 80–90
surface temperature 71, 73, 78, 79, 91–97
eddy kinetic energy (EKE) 89–91, 92
Ekman transport 159–162, 174–175, 180, 395
El Niño 227, 232, 251, 252, 269, 393–421, 590
forecasting 393, 402–403. 410, 414, 420–421
observed from satellites 402–404, 419–421
signatures
altimetry 404–409
chlorophyll 415–417
rainfall over the sea 418–419
sea surface temperature 404–407
wind fields 409–415
electromagnetic spectrum
general 28–30
microwaves, use of 29–30
radar wavebands 30
ellipsoid, reference 46, 47
Emiliania huxleyi 101, 254, 255
emission, thermal 36–37
emissivity 36, 37, 43, 371, 428
English Channel 302, 315, 319, 486, 499, 500, 505
ensemble Kallman filter (EnKF) 562, 563
ENSO 232–233, 392, 393–402, 425, 590
ENSO indicators 400, 401, 406, 410, 411
Envisat 56, 307, 308, 309, 370, 547
equator 247
equatorial Kelvin waves 225–227
error
background 552–554, 558
observation 554
errors, measurement 560
errors, sampling 560
ERS–SAR 61, 427, 515, 583–587, 615, 617
ERSEM 557

ERSST 399
ERS-1&2 55, 307, 308, 309, 336, 386, 411–412, 584
essential climate variables (ECV) 588, 589, 591–596
essential ocean observations 611–612
estuarine remote sensing 528–534
Eumetsat 339, 546. 551, 615
EuroGOOS 544–545
European Marine Safety Agency (EMSA) 586–587
European Space Agency (ESA) 36, 41, 45, 66, 546, 551, 574, 615
eutrophication 255, 561
eutrophic water 264–265
extreme wave height statistics 325

f-ratio (new to total production) 258, 261
fast Fourier transform (FFT) 213
filament structures in images 78, 99, 105, 125–126, 149–151
file formats 60, 66
film modulation of surface roughness 464–467
filter, impulse response 228–230
filters, spatial low pass 211–213, 215–216, 229–231, 232–233
fisheries 267–271, 393, 395–397, 401, 415, 558
fisheries management 268, 270
fisheries policy 267
Flamborough Front 504, 505, 513
flooding, coastal 435–443, 533
flux *see* air–sea flux
flux climatologies 385, 386
footprint 9, 10
foundation SST 40, 41, 579
Fourier analysis 213, 301
Fourier transform 301
friction velocity 365, 366
front climatology 139, 142–146
front detection
  automated 138–142
  ocean color 126–129
  radar images 128–135
  sea level anomaly 124–126
  surface temperature 118–124
front-following algorithm 140–142
frontal meandering 92, 117
frontal variability 146–151
fronts 70, 104, 115–155
fronts, dynamic structure 116
Fundamental Climate Data Record (FCDR) 594–598, 601
future, for ocean remote sensing 282, 326–328, 382–386, 610–616

Galapagos Islands 181, 182
gas flux equation 362
gas solubility 362, 365, 372, 373
gas transfer velocity 362, 372, 373, 374, 375, 376
GCOM-W1 satellite 601, 615
GCOS 588, 592–601
GEO (geostationary orbit) *see* orbit, geostationary
geographical domain for coastal altimetry 526
geoid 46, 47, 50, 86
geolocation 21, 23–24
geophysical data record (GDR) 49, 309, 310, 573
Georges Bank 261
Geosat 55, 307, 308, 309
Geosat altimeter 60
Geosat Follow-on 56, 307, 308, 309, 573
Geosat F–O altimeter 60, 573
GEOSS 391 (footnote), 447
geostrophic flow 49, 73–75, 81, 116–117, 125
GHRSST 42, 45, 365, 407, 569, 575–583, 599–600, 613, 618
  data specification (GDS) 576, 579–580
  -L2P *see* data products, GHRSST L2P
  long-term storage and reanalysis facility (LTSRF) 407, 600
GLI 57, 267
Global Climate Record (GCR) 588
Global Ocean Data Assimilation Experiment (GODAE) 550–551, 575–6
global ocean monitoring 246–250, 540–541
Global Ocean Observing System (GOOS) 544–545
Global Precipitation Climatology Project (GPCP) 418, 420
global primary production 264–265
global thermohaline ocean circulation 427

GlobColour 574, 601
GLONASS 21
GMES 541, 542, 551, 586, 610, 615–616
GNSS-R 328, 617
GOCE 50, 125, 573, 599, 613, 616, 617
GOES 55, 577
GOSIC 592
GPS 21, 617
GRACE 50, 86, 125, 573, 613
Grand Banks 241–242
gravity 46, 125, 617
Great Whirl eddy 80–81, 83–84, 102
Greenland Sea 427, 431
Group on Earth Observations (GEO) 391 (footnote), 446, 447
group velocity, ocean surface waves 295
Guinea, Gulf of 252
Gulf Stream 69, 70, 83, 86, 89–94, 98, 118, 126, 128, 135–136, 138, 146–148, 223

habitat mapping, shallow seabed 272, 273
haddock 269, 270
hammerhead patterns in images 93, 99, 107
harmful algal blooms (HAB) 272, 519, 520, 555, 558, 561–562
Hawaii 180, 252
HDF format files 60, 66, 242, 445
high-resolution visible sensors
  AVNIR 275
  CHRIS 273, 275
  DMC 531
  ETM 274
  HRG 274
  HRV 272, 274, 530
  IKONOS 275, 277
  LISS 274
  MSS 274
  Quickbird 275, 531
  Thematic Mapper 272, 274, 530
  WorldView-1 275, 531
  WorldView-2 275
HNLC (high-nutrient low-chlorophyll) 184, 186, 262
HOAPS 379, 381, 383, 384
hotspots, coral bleaching 280–281
Hovmöller plots 200–202, 212–219, 226–227, 229–230, 409–410, 414, 420
HRV 272, 274, 530
human impact from ocean observations 391–447
Humboldt upwelling (Peru and Chile) 171–172
humidity *see* specific humidity
hurricanes 344–349
  effect of SST 347–349
  impact on marine biology 350
  forecasting
    impact of satellite winds 343–347
    impact of altimetry 348–349
hydrodynamic modulation of surface roughness 103–104, 110–111, 301, 463–465
hydrogen sulfide 255
hydrological cycle 602

ice edge phytoplankton blooms 181–184
IGDR 310, 311
IKONOS 275, 277
image-processing software 64, 66
images
  analysis 27
  averaging 27
  classification 272–278
  color contoured 119–120
  composite 26–27, 41–42
  contrast stretching 119–120
  creating from point samples 9–11
  filtering 120–123, 405
  geometric distortion 11
  resampling 21, 23
  satellite co-ordinates 22, 57
  smoothing 27
images of data derived from
  AMSR-E 97, 228, 338
  ASAR 136, 137, 314–316, 319, 469, 585
  ATSR or AATSR 94, 120, 121, 123, 149, 205, 217, 227
  AVHRR 71, 104, 139, 144, 145, 147, 281, 345, 349, 404, 405, 424, 493, 496, 497, 499, 504
  ERS scatterometer 346, 412
  Jason altimeter 82, 311, 337, 444
  Landsat Thematic Mapper 254
  MERIS 514, 521

images of data derived from (*cont.*)
MODIS (color) 73, 98, 99, 127, 164, 168–170, 172, 173, 176, 177, 179, 184, 495, 510, 512, 513
MODIS (SST) 93, 119, 163, 167, 169, 170, 172, 173, 177, 512
multimission altimetry 84–85, 87, 90–92, 187, 189, 217, 349, 408, 411, 425, 438, 440
OCTS 416
SAR 104, 105, 107, 108, 130, 132, 133, 302, 354, 458, 466, 470, 473, 515
SeaWiFS 100, 102, 129, 153, 182, 185, 188–189, 206, 241–246, 248, 250, 254, 263, 270, 416–417, 426, 473, 498, 508, 522
SeaWInds (QuikScat) 165, 176, 177, 335, 353, 375, 423
SPOT 278
SSM/I 429, 430, 431, 432, 434
TerraSAR-X 465
TMI 230
TOPEX 150, 151, 203, 206, 324, 326, 409, 419, 440
Windsat 339
imaging radars 53 (*see also* SAR)
imaging spectrometer 30
Indian Ocean 72–73, 85, 102, 203, 216, 231, 241, 247, 267, 422–426
incidence angle (radar) 51, 52
inertial frequency 455
infrared radiometer applications to
air–sea fluxes 364–365
El Niño 404–407
fronts 118–123
internal waves 459
large-scale dynamic processes 204–205
mesoscale eddies 91–95
ocean biology 245, 258, 261, 267, 269, 271, 272, 279, 280, 281
shelf seas 491–494, 496, 497–507
upwelling 162–165, 175–177
infrared radiometry 35–42, 58, 576–577, 583
infrared windows 28–29, 36
inherent optical properties (IOP) 253, 518, 568
instantaneous field of view (IFOV) 7, 8, 9, 10
in situ observations 397–399, 403, 540, 544, 546–547, 549–550, 553, 590, 611
in situ sensors on elephant seals 427
interdisciplinary marine science 240
interfacial waves 460, 461
internal
solitary waves 458, 459, 461, 471
soliton of depression 461, 462, 463, 468, 469
soliton of elevation 468. 459, 470
solitons 461, 462
tides 456, 467, 468, 471–474, 475–479
wave breaking 492
wave packets 459, 462, 468, 469
wave thermal signatures 459
waves
analytical description 453–456
detecting propagation characteristics 462–463
effects detected by ocean color 471–479
importance in oceanography 456–457
new insight from remote sensing 479–480
propagation speed estimation 468
SAR signatures 457–470
signature polarity inversion 468–470
solitary wave packets observed with SAR 459–462
surface roughness modulation 463–467
typical scales in shelf seas 489
International Charter on Space and Major Disasters 446, 531
Internet sources of satellite data 58–60, 62–63, 65–66, 569
intertropical convergence zone (ITCZ) 231
IPCC 588, 589, 592, 594
Irish Sea 486, 500–508, 513
Irish Sea (western) Front 504, 505, 508, 513
Irminger Basin 252
iron enrichment 184–5
iron limitation 184–187
islands 247–248
island wakes 179–181
Islay Front 504, 513
isobars 74

Japan, fisheries 267
Japan Sea 267

Jason-1&2 46, 48, 56 82, 307, 308, 309, 310, 311, 407, 573
JCOMM 446
JGOFS 252

Ka-band radar 30, 61, 300, 334, 615
Kelvin waves 225, 232
  equatorial (EKW) 224, 225–227, 229, 410, 420
kinematic flow properties 88–89
Korteweg de Vries (K-dV) equation 461, 469
Ku-band radar 30, 365
Kuroshio Current 83, 89, 105, 131, 143, 145, 223

L-band radar 30, 45, 466, 587, 602, 616
La Niña 393–421
Labrador Sea 85, 183, 270
Landsat 55, 101, 272, 274, 277, 528, 530
Laplacian image filter 121
latent heat of vaporization 363, 378
latent heat flux 362, 363, 378, 380, 381, 382, 383
laser altimeter 533
Legeckis waves 227
lengthscales in shelf and coastal seas 487, 489, 490, 511, 520, 529, 530
LEO *see* orbit, near polar
LIDAR 267, 334, 533
light penetration in the ocean 564–565
linear phase speed (interfacial waves) 461–462
loggerhead turtles 269
log-normal distribution of chlorophyll 249
Lombok Strait 131–132
Longitude–time plots *see* Hovmöller plots
long-wave propagation 74–75

macroalgae 272, 277
Madden–Julian oscillation 231–232
Malvinas (or Falkland) Current 83, 89, 91, 128–129, 146, 153
management 268–270
mangroves 272, 277
map projections 23–26
marginal ice zone 145, 181–184, 318
marine
  aquaculture 271–272, 558, 561
    Core Services (MCS) 545–6, 551, 574, 586, 610, 616
marine ecosystems 187, 240, 246, 249, 252, 277, 279, 282, 456
  ecosystem models 253, 522–523, 555–569, 618
  environmental management 540, 547, 558, 610
  insurance 317
maximum cross correlation (MCC) 188
mean available wind power potential 352
mean sea level (MSL) 437–441
mean square slope (of surface waves) 366, 376
meddies 79
Mediterranean Oscillation Index (MOI) 326
Mediterranean outflow 79
Mediterranean Sea 85, 88, 92, 94, 107, 110, 251,
Medspiration Project 576, 581
melting ice 182–3
meridional component of wave speed 216–219
MERIS 35, 57, 244, 250, 253, 269, 417, 479, 514, 521, 574, 584, 601, 615
MERSEA Project 546
mesoscale, the ocean 15, 45, 69, 72–76,
mesoscale variability 50, 69–190, 553
mesotrophic waters 264–265
Meteosat 55
Metop 56, 334, 615
methane CH_4 370
methods of satellite oceanography
  diversity of techniques 28–53
  general 7–28
Mexico, Gulf of 178, 251
microsatellites 327
microwave radiometer applications to
  air–sea fluxes 364, 365, 368, 369
  ENSO 418–419
  large-scale dynamic processes 204–205, 227–230
  mesoscale eddies 96–97
  ocean biology 249, 280
  sea ice 427
  shelf seas 491
  wind over the ocean 337–340

microwave radiometry 42–45, 59, 96–97, 427, 576–577, 617
mid-latitude depression 441
Mississippi River 178
MISST Project 576, 581
mixing, vertical 457, 471, 487, 489, 490, 493, 494, 498, 502
modes of planetary waves 208–209, 211, 216, 221, 223, 226
MODIS 57, 58, 244, 269, 417, 479, 494–495, 510, 512, 513, 574, 584, 601, 615
modulation transfer function 303, 304, 313
molecular diffusion 360–361
Monin–Obukhov theory 378
monsoon 231, 392, 421–426
monsoon causes 421
monsoon index 425
movie sequence of images 196, 200
Mozambique Channel 479
MPI method, wave spectral inversion 303, 304, 313, 314
multi-sensor histogram edge detection (MSHED) 138–139
multispectral radiometer 30
MyOcean 546, 555, 573

nadir view 10, 46, 336
NASA 45, 576
NASA Ocean Color website 241–250
National Oceanographic and Atmospheric Administration (NOAA) 36, 70, 576, 610
navigation through ice 429, 431
NCEP analysis wind fields 340, 341, 368, 369, 377
near infrared *see* visible and near IR
near-real time data 65, 365, 445, 545, 575, 581
NEODAAS 493, 496, 497, 498, 499, 508
NetCDF 60, 66, 445, 579–581
net heat exchange 362
net longwave radiation 362, 370, 371
net shortwave radiation 362, 370, 371
neutral stability ABL 365, 366, 378
new production (P_{new}) 258, 261
New Zealand 85
Newfoundland Basin 85, 183, 241
Nimbus-7 55
Niño-3/4 region 398, 410
Nitrogen 556–557
NOAA meteorological satellite series 55
NOAA Photo Library 70
nonlinear wave theory 461
Norte event 175–177
Northeast Passage 434, 435
Northwest Passage 434
North Atlantic Oscillation (NAO) 325,–326, 439–441
North Brazil Current (NBC) 179
North Equatorial Countercurrent (NECC) 179
North Sea 486, 498, 500, 501, 505, 511, 514
Norwegian coast 561–562
Norwegian Sea 427
Nova Scotia 241, 261, 270, 486
NPOESS 338, 339, 614, 615
NPP 339
NPZD model 556–557, 564
NSCAT 61, 386
NSIDC 428, 430, 431, 433, 434
nuisance bloom 519, 521
numerical ocean prediction (NOP) 544, 547–554, 570–572, 590, 609–610, 618
numerical weather prediction *see* weather forecasting
nutrient supply 247, 251, 257, 258, 261, 269
nutrient transport 456

objective analysis 385, 552–554, 572
ocean biology 239–283, 555
ocean climate 540, 588–591, 618
ocean color
  algorithms 34
  anomalies 249–250, 416–417
  Case 1 and Case 2 waters 34, 517, 565–568
  data products 34, 35
  climate applications 597, 600–601
  limitations in Case 2 waters 247, 252, 517–519, 565–568, 613
  remote sensing 30–35
  sensors 57
ocean color applications to
  El Niño 415–417
  fronts 126–128
  internal waves 458, 471–479

large-scale dynamic processes 206, 211
mesoscale eddies 98–102
monsoon 425–426
ocean biology 239–283, 471–479
shelf seas 491, 494, 495, 507–514, 516–523
suspended sediments 508–515
upwelling 162–165, 168, 180–182
ocean dynamics, large-scale 195–234
ocean dynamics, mesoscale 50, 69–190, 553
ocean forecasting 540, 543–554, 618
ocean forecasting models 421 (*see also* numerical ocean prediction or NOP)
Ocean Niño Index (ONI) 398–401, 410, 411
ocean-observing system (OOS) 544–545
ocean state determination 540
ocean surface currents 50, 86, 410–415
ocean surface waves
frequency spectrum 296, 314
mean period 297, 300, 320
peak period 297, 300, 320
wave
age 297, 304, 367–368, 601
buoys 296, 297, 300, 304, 307, 320, 322, 323, 325
climate 322–325, 601
damping 367, 464, 583
energy 295, 318
height *see* significant wave height
measurement 294–316
modulation 53, 463–467
period 295
prediction models 319–321, 618
propagation and characteristics 294–306
skewness 297, 300, 324, 327
spectra 295–297, 300, 303–305, 312–314, 318, 321, 322, 327, 464
spectrometry 304–307
speed 295
theory 294–297
wavelength 295
wavenumber 295, 296, 303, 305, 314
spectrum 296. 314
zero-crossing period 297, 300
Ocean Vector Winds Science Team (NASA) 335
oceanic internal waves *see* internal waves
oceanographic use of wind data 341–343
octet (8-day average) of data 243, 244, 245, 248, 249, 263
OCTS 57, 267, 269, 415–417, 574
offshore industry 294, 297, 317, 322, 324, 325
offshore wind power generation 350–353
oil slicks 53, 583
oil spill monitoring 583–588
Okhotsk, Sea of 143, 146, 183
Okubo–Weiss parameter 88–89
OLCI 601, 615, 616
oligotrophic water 252, 257, 264, 265
operational data monitoring 445, 543, 544, 570–572, 610
operational oceanography 270, 445, 541, 543–546, 547–551, 558, 569–582, 610
orbit
geostationary 12
near-polar 12, 13
period 12
repeat cycle 46
repeat track 47
Sun-synchronous 13, 47, 571
tracking 47
Orinoco River 252
OSCAR (ocean surface current analysis) 412–415
OSI–SAF 428, 432
outgassing of CO_2 377
oxygen 360, 362

Pacific–Antarctic Ridge 252
Pacific Ocean 206, 216, 226–229, 231–233, 241, 247
Pacific Ocean, equatorial 247, 393–420
Pacific Ocean fronts 143–144
Pacific Ocean, north 243, 252, 267, 269
PALSAR 587, 615
Panama, Gulf of 143, 175–176
Papagayo, Gulf of 143, 175–176
parameter estimation in models 321, 322, 559
partial pressure 362, 369, 370, 372, 376
particle transport 457
particulate organic carbon (POC) 261
particulate organic nitrogen (PON) 556
Patagonian Shelf 128–129, 155 243, 486

patchiness of phytoplankton 471
Pathfinder SST data 71, 143, 404, 405, 422, 424
phase speed, ocean surface waves 295, 297
phase speed, linear internal waves 460
phenology, marine 249, 269
photic zone 261, 456, 509
photon flux density (E_{PAR}) 256, 257, 258, 259, 260
photosynthesis 247, 256, 257, 259, 262, 264, 279
  cross-section for 259–260
photosynthetic rate parameters 257, 260
photosynthetically available radiation (PAR) 35, 262–264, 371, 564–565, 566–568 (*see also* photon flux density)
photosynthetically stored radiation rate (PSR) 256, 259
physical forcing of primary production 241, 245, 247, 249, 252–253, 282–283
phytoplankton 240–254, 456, 556–557
  biomass 241, 246, 256, 260, 261, 264, 269, 270, 282
  blooms *see* blooms, phytoplankton
  functional groups 253, 557
  spatial distribution 241–245, 246–250, 251–253
  seasonal variability 241, 243–245, 248–249
  transport 457
picture file formats 66
Pinatubo, Mt. 39
pixels
  averaging 26
  cloudy or cloud-free 19
  location 23
  nearest neighbor substitution 24
  resampling 21, 24
Planck function 36
planetary waves 195–234
Plymouth Front 504, 505, 513
polar orbit *see* orbits, near-polar
polar remote sensing 426
polarimetric radiometer 339
polarity of internal wave signatures 463, 468–470
Poseidon altimeter 46, 60, 81, 82, 366, 386, 573
pressure field 73–74
Prestige oil tanker disaster 584–585
primary production 128, 152–155, 255–267, 456, 555
  empirical models 258
  enhanced by internal waves 477–479
  estimates from remote sensing 258–262, 264–267
  process-based models 258–261
primary production rate (P) 256–259
  column integrated (P_{tot}) 257, 258, 259, 261
propagating features 200–202, 212–219
public good 391, 407, 446–447, 542–543, 546, 550

quantum yield (φ_μ) 256, 257
Quickbird 275, 531
Quikscat 56, 334, 375, 386, 410, 423

RA 60, 573
RA-2 60, 310, 573
radar backscatter cross-section 51, 52, 103–106, 299–300, 366, 464
radar ocean wave spectrometer 304–307
radars *see* sensors, active microwave
Radarsat 55, 313, 427, 586, 615
Radon transform 213–216
  three-dimensional 217–219
radiance, water-leaving 33
rain over the ocean 231–232, 418–419, 421–422
rainfall patterns of ENSO 393–397, 419–420
rank-ordered wave packets 462, 468, 469, 470
Rayleigh wind distribution 375
rays of internal wave energy 455
reduced gravity 208
reefs 247–248, 279–282
reflectance
  optical (R) 34, 568
  remote sensing (R_{RS}) 568
relaxation time (surface waves) 464
remote-sensing information flow 8
research 171–174, 268–270
  applied 541, 542
  motivation 541–543
  pure 541

resolution
  spatial 15
  temporal 15
Reynolds number 72
Reynolds SST analysis 386
ring, in surface current field 89
river plumes 145, 177–179, 252
Roberts image filter 121
ROFI (region of freshwater influence) 487, 489, 505, 511
roll vortex 95, 106–107, 133, 33
Ross Ice Shelf 184
Rossby radius of deformation 72–76, 80, 82
Rossby rototiller effect 222
Rossby wave,
  importance in ocean science 223–224, 410
  speed 209–210, 212–219
  theory 208–210
  theory enhanced 221–222
Rossby waves 196, 206–211, 220–223, 229, 410
  chlorophyll signature 206, 222
Rossby waves and eddies ambiguity 222
roughness, sea surface 43, 51–53, 299, 366
roughness height 366
roughness signatures 53, 103–111, 457–470
runoff, rivers or land 487, 488, 520

S-band radar 30
safety at sea 268, 294, 317, 540
salinity, ocean 43, 45, 597, 602, 616, 617
salinity fronts 145
sampling
  capabilities of satellite sensors 9–16, 608–609
  geostationary sensor 14, 570–571
  interval 14, 47, 587
  non-scanning sensor 15
  polar-orbiting sensor 14–15, 570–571
  space-time tradeoff 15, 570–571
  spatially averaged 11, 570–571
SAR applications to
  coasts and estuaries 530, 532–533
  fronts 129–133
  internal waves 459–470
  mesoscale eddies 103–111
  sea ice 427
  shelf seas 491, 494, 515–516
  surface waves 300–304, 312–316, 318–319
  upwelling 176, 180
  wind over the ocean 336
SAR (synthetic aperture radar) 53, 61
  automated image analysis 479
  complex image 301, 303, 313, 315
  geometry relative to surface waves 302
  image intensity modulation 460, 461, 463, 470
  image spectrum 301, 303, 304, 315
  image wave modulation transfer function 301
  interferometry 617
  inverse wave spectral transform 303–304
  wave mode imagettes 312–315
satellites used for ocean observing
  ADEOS-1&2 56, 415, 547
  ALOS 587
  Aqua 45, 56, 370, 547
  Aquarius 45, 602 616
  Coriolis 56, 338
  Cryosat 599, 615
  Disaster Monitoring Constellation 531
  DMSP 55, 338, 428
  Envisat 56, 307, 308, 309, 370, 547
  ERS-1&2 55, 307, 308, 309, 336, 386, 411–412, 584
  GCOM-W1 601, 615
  Geosat 55, 307, 308, 309
  Geosat Follow-on 56, 307, 308, 309, 573
  geostationary 12, 14, 41
  GOCE 50, 125, 573, 599, 613, 616, 617
  GOES 55, 577
  GRACE 50, 86, 125, 573, 613
  Jason-1&2 46, 48, 56 82, 307, 308, 309, 310, 311, 407, 573
  Landsat 55, 101, 272, 274, 277, 528, 530
  Meteosat 55
  Metop 56, 334, 615
  Nimbus-7 55
  NOAA 55
  NPOESS 338, 339, 614, 615
  NPP 339
  ocean-viewing 54–56
  orbits 11–13
  Quikscat 56, 334, 375, 386, 410, 423
  Radarsat 55, 313, 427, 586, 615
  Seasat 55

satellites used for ocean observing (*cont.*)
Seastar 56
Sentinel-1 615, 616
Sentinel-3 317, 582, 601, 614, 615–6
SMOS 45, 602, 616, 617
SPOT 55, 272, 274, 277, 530
Terra 56, 547
TerraSAR-X 465, 587, 615
TIROS-N 55
TOPEX/Poseidon 46, 48, 55, 81, 307, 308, 309, 407–410, 573, 601
TRMM 45, 56, 338, 418
unique perspective from 9–16
scale mismatch between observation and phenomenon 529–532
scanning techniques 10, 11, 38–39
scattering, optical 33
scatterometer applications
air–sea fluxes 364
El Niño 343, 410–412
monsoon 422–423
ocean biology 343
shelf seas 491, 527
surface currents 412–415
upwelling 162–165, 174–177, 343
wind over the ocean 334–336
scatterometers 51–52, 61
SCIAMACHY 370
Scully-Power, Paul 109
sea bed reflection 33, 272, 273, 276, 277,
sea grass 272, 276, 277
sea ice 183–4, 392, 426–435, 597, 598–599
concentration 428–432, 434
composite maps 429, 430, 432
data quality constraints 429–430
detection 427–428
edge 428–431, 433, 434
extent 392, 428–435
index 428, 430–435
sea level 437–441, 597, 599
anomaly *see* sea surface height anomaly
rise 437–440, 590
sea state 293–328, 597, 601–602 (*see also* waves, ocean surface)
sea surface height *see* topography, ocean surface
sea surface height anomaly (SSHA) 48, 49, 403, 407–410, 425, 612
sea surface mean square slope 299
sea surface roughness *see* roughness, sea surface
sea surface temperature *see* SST
Seasat 55
seasonal ice zone (SIZ) 183
seasonal variation of shelf seas SST 500–502
Seastar 56
SeaWiFS 35, 57, 240–250, 267, 269, 415–417, 426 498, 508, 522, 559, 574, 584
SeaWinds 61, 334, 410, 423
secondary circulation 117, 120, 126
sensible heat flux 362, 363, 369, 378, 379, 380, 381, 382, 384
sensors for ocean observation
AATSR 58, 364, 577, 581
active microwave (generic radars) 30, 46–53
AIRS 370
AltiKa 615
AMI 61 334, 411–412
AMSR-E 45, 59, 338, 368, 577, 599
AMSR-2 615
AMSU-A 369
ASAR 61, 304, 312, 313–316, 319, 322 427, 466, 468, 469, 584–585, 586
ASCAT 81, 334, 336, 615
ATSR 36, 38–39, 41, 58, 407
AVHRR 36, 38, 41, 58, 71, 364, 399, 404, 405, 424, 493, 495–497, 499, 504, 577
CMIS 338
ERS–SAR 61, 427, 515, 583–587, 615, 617 generic types 28–53
Geosat altimeter 60
Geosat F-O altimeter 60, 573
GLI 57, 267
high-resolution visible *see* high-resolution visible sensors
infrared radiometers (generic) 29, 35–42
MERIS 35, 57, 244, 250, 253, 269, 417, 479, 514, 521, 574, 584, 601, 615
MODIS 57, 58, 244, 269, 417, 479, 494–495, 510, 512, 513, 574, 584, 601, 615
NSCAT 61, 386
OCTS 57, 267, 269, 415–417, 574
OLCI 601, 615, 616
PALSAR 587, 615

passive microwave (generic) 29, 42–45
Poseidon altimeter 46, 60, 81, 82, 366, 386, 573
RA 60, 573
RA-2 60, 310, 573
SCIAMACHY 370
SeaWiFS 35, 57, 240–250, 267, 269, 415–417, 426 498, 508, 522, 559, 574, 584
SeaWinds 61, 334, 410, 423 SEVIRI 58, 577
SIRAL 528, 599, 615, 616
SLSTR 582, 614, 616
SRAL 615
SSM/I 45, 59, 338, 368, 369, 418, 428, 599
SSM/T-2 369
TMI 45, 59, 338, 368, 418
TOPEX 46,48, 60, 386, 407, 409, 410, 418, 419, 440, 573
VIIRS 601, 615
visible and near-IR (generic) 29, 30–32, 35, 57
Windsat 59, 338, 339, 368
Sentinel-1 satellite series 615, 616
Sentinel-3 satellite series 317, 582, 601, 614, 615–6
SEVIRI 58, 577
shear flow, evidence in images 72, 78–81, 85, 88, 89, 93, 99, 131
shelf break front 145, 492
shelf edge phenomena 492–496
shelf sea
dynamical phenomena 497–508
ecosystems 516–523
global locations 487,
lengthscales *see* lengthscales
remote sensing 486–522
satellite sensors 491
suspended sediments 489, 491, 495, 508–516
tidal fronts *see* tidal mixing fronts
timescales *see* timescales
shipping 297, 317, 322, 324, 344
ship-routing 317
shoreline threat 328, 435, 436, 441–443, 530, 532, 533
Siberian coast 431
side-looking airborne radar (SLAR) 267, 583
sigma-zero, σ_0 *see* radar backscatter cross-section 464
significant wave height (SWH) 49, 50–51, 295, 297, 299, 300, 307, 309, 314, 316, 317, 320–326, 367–368
singular value decomposition 227
SIRAL 528, 599, 615, 616
skin SST 40, 41, 364, 371, 578
skin, thermal 40, 364, 578
sky light 32
sky radiance 37
slick-modulated signatures on SAR images 106–108
SLSTR 582, 614, 616
SMOS 45, 602, 616, 617
Sobel image filter 121
societal benefits of applied oceanography 540–543
soliton 185, 227
South Africa 70, 71
Southern Gyre eddy 80–81, 83–84
Southern Ocean 182–3, 252, 427
Southern Oscillation 393, 401
Index (SOI) 400–401, 418, 420
Southland Front 148
specific humidity 363, 368, 378, 381
specific inherent optical properties (SIOP) 518, 568
SPOT 55, 272, 274, 277, 530
SPRA method for wave spectra 304, 313, 314
Spring bloom 242, 243, 245, 249, 251, 252, 269
spring–neap tidal cycle 490, 494, 505
SRAL 615
SSALTO 573
SSM/I 45, 59, 338, 368, 369, 418, 428, 599
SSM/T-2 369
SST (sea surface temperature)
anomaly 204, 205, 211, 230, 233, 397–400, 403, 405–407
definitions 40–41
general applications 591, 597, 599–600, 618
microwave radiometry 42–43
multi-channel (MCSST) 42
Pathfinder 42, 404, 405, 422, 424
thermal infrared radiometry 35–42

SST data applications to
  air–sea fluxes 364–365, 371, 372, 373
  El Niño 387–399, 402–407
  fronts 118–123
  large-scale dynamic processes 204–205, 211
  mesoscale eddies 91–95
  monsoon 423–424
  ocean biology 245, 249, 258, 261, 267, 269, 271, 272, 279, 280, 281
  upwelling 162–165, 175–177
  wind over the ocean 347–349
Stanton number 363, 378, 380
state variables 551, 552, 554, 556–559, 562–568, 570
Stefan–Boltzmann constant 371
Stokes drift 135
storm surges 344, 441–442
strain, surface current field 88
stratification parameter for tidal shelf seas 503, 505, 506
sublayer, molecular diffusion 361, 364, 457
subskin SST 40 ,41, 364, 578
Sulu Sea 459
sun glint 458, 471
sun glitter 32, 109–110
supervised classification 273
surface
  films 106–109, 366, 367, 376, 385, 457, 464–467
  slicks 457, 454, 459, 464–467
  solar irradiance (SSI) 35
  waves 293–327 (*see also* ocean surface waves)
surfactants 106–107, 367, 376, 464, 466
suspended
  particulate material (SPM) 34, 514
  sediment map structures 511–513
  sediments monitored in shelf seas 508–516
swath altimetry 442, 612
swath width 11, 14, 53
swell 293, 295, 300, 301, 302, 304, 316, 318, 319, 321, 322, 366, 367, 532–533
swordfish 271
synergy, exploiting different remote-sensing methods 102–103, 104, 149–152, 154–155, 184–187, 189 240, 419–421, 445
synthetic aperture radar *see* SAR

TAO array 397, 398, 399, 402
Taylor–Proudman theorem 489
Tehuantepec, Gulf of 143, 175–177
temperature
  brightness 36–38, 44
  sea surface 575–582 (*see also* SST)
Terra satellite 56, 547
TerraSAR-satellite 465, 587, 615
Thematic Mapper 272, 274, 530
thermal signatures of shelf sea dynamics 497–508
thermal skin layer 364, 371, 372
thermocline
  diurnal 41, 95, 124, 180, 364, 578–579, 617
  permanent 74–75, 207–208, 395, 410, 471
  seasonal 459, 493, 500–503
tidal analysis 436–437
tidal corrections, altimetry 523, 527
tidal currents 487–492, 498, 502–505, 508, 511–513, 515, 516
tidal mixing fronts 145, 489, 502–507, 513
tidal mixing front, cross-section 503, 507
tide gauges 436, 437, 439, 440, 441, 442
tides 47, 435–437
tilt modulation 301
timescales in shelf and coastal seas 487, 520, 529, 530
TIROS-N 55
TMI 45, 59, 338, 368, 418
TOGA 50, 402
top-of-atmosphere (TOA) radiance 595, 601
TOPEX altimeter 46,48, 60, 386, 407, 409, 410, 418, 419, 440, 573
TOPEX/Poseidon mission 46, 48, 55, 81, 307, 308, 309, 407–410, 573, 601
topography
  absolute dynamic (ADT) 50, 86–87, 125–126, 150–152, 186–187, 612, 616
  mean dynamic (MDT) 50, 126
  ocean dynamic 47, 48
  ocean surface 46, 80–89
  ocean surface time mean (MSS) 48
Torres Strait 441
toxic phytoplankton 520, 521

Tracker algorithms for altimeter 523,
transfer coefficient 362
transfer velocity 362, 367, 372–377
trend, sea ice 430, 432, 433,434, 435
trend, mean sea level 437–441
TRITON mooring 397, 398, 399, 402
TRMM 45, 56, 338, 418
tropical cyclones 344–349, 441–442
tropical cyclone heat potential 349–350
tropical instability waves 227–231
  importance of 231
tropical seas, shallow habitats 272–279
tsunami model 443
tsunamis 442–444, 531, 617
tuna, albacore 269, 271
tuna, bluefin 271
turbulence in geophysical flows 72, 73, 76, 93, 100, 104, 361
turbulent flow, evidence in images 72, 89, 91–97, 98–103, 104, 106, 109–111
turbulent flux 361, 362, 378–382
turtles 269–270
typhoon 344
Tyrrhenian Sea 107–108

ultra-violet (UV) detection 583
UNESCO 66
UNFCCC 588, 592, 600, 610
unstable ABL 366, 378, 380
upwelling 117–118, 153, 229
  causes and consequences 159–162
  coastal 70, 71, 160
  detection by, satellites in general 162–167
    altimetry 165–166
    ocean color 164, 168
    SST 162–163, 167, 169–170, 172–173, 258, 269
    vector winds 165
  equatorial 161, 395, 396
  index 160, 174
  regions of the world 167–171
  research 171–174
Ushant front 504, 505, 507, 513

validation of models 320, 321, 559–561
validation *see* data products, validation
variability
  climatic 161, 174, 190
  decadal 322, 325, 426
  diurnal 364, 571, 578–9
  interannual 145, 148, 186, 187, 230, 234, 252, 269, 322, 325, 326, 370, 392, 415, 422, 424, 426, 432, 433, 437, 439
  long-term 282, 325, 399
  mesoscale 72–78, 115–118, 159–189
  seasonal 144, 148, 154, 161, 174, 196, 198, 246, 251, 282, 322, 403, 405, 415, 421, 437, 500–502
variability scales in shelf seas 488–492
vector winds 338, 339, 340, 346, 352
velocity bunching 301
vertical mixing *see* mixing, vertical
vertical structure of shelf seas 489
Vertically Generalized Production Model 265–267
VIIRS 601, 615
visible and near IR
  atmospheric window 28, 29
  remote sensing *see* ocean color remote sensing
  sensors *see* sensors, visible and near-IR
visualization of internal wave beams 455, 456
volcanic dust 39, 407
vorticity, planetary 207
vorticity, surface current field 88
vortices in image fields 228–229

WAM (wave model) 315, 320, 532
Warm Pool, equatorial Pacific 395, 397, 402, 406, 412, 418
water depth effect on sea bed reflection 273, 276
water quality monitoring 516–523, 555, 558
water vapor 38, 360, 362, 363, 368–369, 370, 378, 381
wave
  dispersion 295
  power extraction 325
  properties *see* ocean surface waves
  refraction 316, 318
  shadows 316, 318
waves on beaches 294, 295, 302, 314, 318, 327, 532–533
WCRP 446, 588

weather forecasting (NWP) 231, 320, 540, 544, 547, 576, 602
weather response to ENSO 393–397,
Weibull distribution 352
westerly wind bursts (equatorial Pacific) 409–411, 420
wind
  data, applications in oceanography 341–343
  -driven offshore dynamical features 52, 174–177
  farms 350–353
  friction velocity (u^*) 297
  measurement 333–354
    altimetry 336–337
    microwave polarimetry 338–339
    microwave radiometry 337–338
    SAR 336
    scatterometry 334–336
  power density distribution 352
  power potential 351
  shadow 180 341, 353
  speed 45, 49, 52
  stress 360, 362, 364, 365, 366, 386
Windsat 59, 338, 339, 368
World Weather Watch (WWW) 544
WorldView-1 275, 531
WorldView-2 275
WOCE 50

X-band radar 30, 334, 465, 587, 615, 616

Yellow Sea 143, 486

Z_{90} (depth viewed by visible remote sensing) 474–475, 509, 514
Zapiola Rise 154–155
zonal mean speed of planetary waves 216, 218
zooplankton 249, 556–557