MODELING METHODS FOR MARINE SCIENCE

This textbook on modeling, data analysis, and numerical techniques for marine science has been developed from a course taught by the authors for many years at the Woods Hole Oceanographic Institution.

The first part of the book covers statistics: singular value decomposition, error propagation, least squares regression, principal component analysis, time series analysis, and objective interpolation. The second part deals with modeling techniques: finite differences, stability analysis, and optimization. The third part describes case studies of actual ocean models of ever-increasing dimensionality and complexity, starting with zero-dimensional models and finishing with three-dimensional general circulation models. Throughout the book hands-on computational examples are introduced using the MATLAB programming language and the principles of scientific visualization are emphasized.

_Modeling Methods for Marine Science_ is a textbook for advanced students of oceanography on courses in data analysis and numerical modeling. It is also an invaluable resource as a reference text for a broad range of scientists undertaking modeling in chemical, biological, geological, and physical oceanography.

**DAVID M. GLOVER** is a Senior Research Specialist in the Department of Marine Chemistry and Geochemistry at Woods Hole Oceanographic Institution. He is the author or co-author of 67 published articles, book chapters and abstracts. Dr. Glover’s research uses satellite data, model results, and shipboard data to elucidate the mechanisms and processes by which the oceans play a major role in the maintenance of the global climate.

**WILLIAM J. JENKINS** is a Senior Scientist in the Department of Marine Chemistry and Geochemistry at Woods Hole Oceanographic Institution. He has published 84 peer-reviewed journal and book articles. Dr. Jenkins is the Director of the National Ocean Sciences Accelerator Mass Spectrometry Facility (NOSAMS). In 1983 he received the Rosenstiel Award in Oceanographic Science from the University of Miami and in 1997 he received the Henry Bryant Bigelow Award in Oceanography from the Woods Hole Oceanographic Institution. Dr. Jenkins’ interests include studying tracers as applied to oceanic physical, chemical, biological, and geological processes; air–sea and ice–water exchange of gases; ocean biological productivity and its controls; radiogenic and primordial noble gas isotopes in the sea, atmosphere, lakes, ground waters, sediments and rocks; climatic changes in the ocean and its effects on biogeochemical systems; and radiocarbon and the global carbon cycle in the past 60,000 years.

**SCOTT C. DONEY** is a Senior Scientist in the Department of Marine Chemistry and Geochemistry at Woods Hole Oceanographic Institution. He has authored or
co-authored more than 160 peer-reviewed journal and book articles. He was awarded the James B. Macelwane Medal from the American Geophysical Union in 2000 and an Aldo Leopold Leadership Program Fellowship in 2004. He has traveled extensively, lending his expertise to a number of national and international science programs, most recently as inaugural chair of the Ocean Carbon and Biogeochemistry (OCB) Program. He has also testified before both the US House of Representatives and the US Senate. His research interests include marine biogeochemistry and ecosystem dynamics, ocean acidification, the global carbon cycle, climate change, and the intersection of science and policy.
MODELING METHODS
FOR MARINE SCIENCE

DAVID M. GLOVER
WILLIAM J. JENKINS

and

SCOTT C. DONEY
Woods Hole Oceanographic Institution
This book is dedicated to
Tina, Susan, and Andrea.
They have endured our preoccupations with
loving support, rare good humor,
and infinite patience.
## Contents

**Preface**

<table>
<thead>
<tr>
<th>Resources, MATLAB primer, and introduction to linear algebra</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Resources</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Nomenclature</td>
<td>3</td>
</tr>
<tr>
<td>1.3 A MATLAB primer</td>
<td>3</td>
</tr>
<tr>
<td>1.4 Basic linear algebra</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Problems</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement theory, probability distributions, error propagation and analysis</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Measurement theory</td>
<td>14</td>
</tr>
<tr>
<td>2.2 The normal distribution</td>
<td>18</td>
</tr>
<tr>
<td>2.3 Doing the unspeakable: throwing out data points?</td>
<td>27</td>
</tr>
<tr>
<td>2.4 Error propagation</td>
<td>29</td>
</tr>
<tr>
<td>2.5 Statistical tests and the hypothesis</td>
<td>31</td>
</tr>
<tr>
<td>2.6 Other distributions</td>
<td>33</td>
</tr>
<tr>
<td>2.7 The central limit theorem</td>
<td>37</td>
</tr>
<tr>
<td>2.8 Covariance and correlation</td>
<td>39</td>
</tr>
<tr>
<td>2.9 Basic non-parametric tests</td>
<td>42</td>
</tr>
<tr>
<td>2.10 Problems</td>
<td>45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Least squares and regression techniques, goodness of fit and tests, and nonlinear least squares techniques</th>
<th>49</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Statistical basis for regression</td>
<td>49</td>
</tr>
<tr>
<td>3.2 Least squares fitting a straight line</td>
<td>52</td>
</tr>
<tr>
<td>3.3 General linear least squares technique</td>
<td>62</td>
</tr>
<tr>
<td>3.4 Nonlinear least squares techniques</td>
<td>68</td>
</tr>
<tr>
<td>3.5 Problems</td>
<td>74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Principal component and factor analysis</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Conceptual foundations</td>
<td>76</td>
</tr>
<tr>
<td>4.2 Splitting and lumping</td>
<td>81</td>
</tr>
</tbody>
</table>
# Contents

4.3 Optimum multiparameter (OMP) analysis 86  
4.4 Principal component analysis (PCA) 95  
4.5 Factor analysis 104  
4.6 Empirical orthogonal functions (EOFs) 109  
4.7 Problems 118  

5 Sequence analysis I: uniform series, cross- and autocorrelation, and Fourier transforms 120  
5.1 Goals and examples of sequence analysis 120  
5.2 The ground rules: stationary processes, etc. 122  
5.3 Analysis in time and space 123  
5.4 Cross-covariance and cross-correlation 131  
5.5 Convolution and implications for signal theory 132  
5.6 Fourier synthesis and the Fourier transform 135  
5.7 Problems 139  

6 Sequence analysis II: optimal filtering and spectral analysis 141  
6.1 Optimal (and other) filtering 141  
6.2 The fast Fourier transform (FFT) 144  
6.3 Power spectral analysis 145  
6.4 Nyquist limits and data windowing 157  
6.5 Non-uniform time series 162  
6.6 Wavelet analysis 166  
6.7 Problems 169  

7 Gridding, objective mapping, and kriging 171  
7.1 Contouring and gridding concepts 171  
7.2 Structure functions 179  
7.3 Optimal estimation 183  
7.4 Kriging examples with real data 186  
7.5 Problems 192  

8 Integration of ODEs and 0D (box) models 194  
8.1 ODE categorization 194  
8.2 Examples of population or box models (0D) 197  
8.3 Analytical solutions 200  
8.4 Numerical integration techniques 203  
8.5 A numerical example 210  
8.6 Other methods 218  
8.7 Problems 222  

9 A model building tutorial 223  
9.1 Motivation and philosophy 224  
9.2 Scales 226  
9.3 A first example: the Lotka–Volterra model 232
## Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.4</td>
<td>A second example: exploring our two-box phosphate model</td>
<td>238</td>
</tr>
<tr>
<td>9.5</td>
<td>A third example: multi-box nutrient model of the world ocean</td>
<td>245</td>
</tr>
<tr>
<td>9.6</td>
<td>Problems</td>
<td>249</td>
</tr>
<tr>
<td>10</td>
<td>Model analysis and optimization</td>
<td>250</td>
</tr>
<tr>
<td>10.1</td>
<td>Basic concepts</td>
<td>251</td>
</tr>
<tr>
<td>10.2</td>
<td>Methods using only the cost function</td>
<td>258</td>
</tr>
<tr>
<td>10.3</td>
<td>Methods adding the cost function gradient</td>
<td>264</td>
</tr>
<tr>
<td>10.4</td>
<td>Stochastic algorithms</td>
<td>271</td>
</tr>
<tr>
<td>10.5</td>
<td>An ecosystem optimization example</td>
<td>274</td>
</tr>
<tr>
<td>10.6</td>
<td>Problems</td>
<td>278</td>
</tr>
<tr>
<td>11</td>
<td>Advection–diffusion equations and turbulence</td>
<td>280</td>
</tr>
<tr>
<td>11.1</td>
<td>Rationale</td>
<td>280</td>
</tr>
<tr>
<td>11.2</td>
<td>The basic equation</td>
<td>281</td>
</tr>
<tr>
<td>11.3</td>
<td>Reynolds decomposition</td>
<td>282</td>
</tr>
<tr>
<td>11.4</td>
<td>Stirring, straining, and mixing</td>
<td>287</td>
</tr>
<tr>
<td>11.5</td>
<td>The importance of being non</td>
<td>287</td>
</tr>
<tr>
<td>11.6</td>
<td>The numbers game</td>
<td>289</td>
</tr>
<tr>
<td>11.7</td>
<td>Vertical turbulent diffusion</td>
<td>293</td>
</tr>
<tr>
<td>11.8</td>
<td>Horizontal turbulent diffusion</td>
<td>294</td>
</tr>
<tr>
<td>11.9</td>
<td>The effects of varying turbulent diffusivity</td>
<td>296</td>
</tr>
<tr>
<td>11.10</td>
<td>Isopycnal coordinate systems</td>
<td>297</td>
</tr>
<tr>
<td>12</td>
<td>Finite difference techniques</td>
<td>299</td>
</tr>
<tr>
<td>12.1</td>
<td>Basic principles</td>
<td>299</td>
</tr>
<tr>
<td>12.2</td>
<td>The forward time, centered space (FTCS) algorithm</td>
<td>300</td>
</tr>
<tr>
<td>12.3</td>
<td>An example: tritium and $^3$He in a pipe</td>
<td>304</td>
</tr>
<tr>
<td>12.4</td>
<td>Stability analysis of finite difference schemes</td>
<td>309</td>
</tr>
<tr>
<td>12.5</td>
<td>Upwind differencing schemes</td>
<td>316</td>
</tr>
<tr>
<td>12.6</td>
<td>Additional concerns, and generalities</td>
<td>321</td>
</tr>
<tr>
<td>12.7</td>
<td>Extension to more than one dimension</td>
<td>323</td>
</tr>
<tr>
<td>12.8</td>
<td>Implicit algorithms</td>
<td>329</td>
</tr>
<tr>
<td>12.9</td>
<td>Problems</td>
<td>330</td>
</tr>
<tr>
<td>13</td>
<td>Open ocean 1D advection–diffusion models</td>
<td>332</td>
</tr>
<tr>
<td>13.1</td>
<td>Rationale</td>
<td>332</td>
</tr>
<tr>
<td>13.2</td>
<td>The general setting and equations</td>
<td>333</td>
</tr>
<tr>
<td>13.3</td>
<td>Stable conservative tracers: solving for $K/w$</td>
<td>334</td>
</tr>
<tr>
<td>13.4</td>
<td>Stable non-conservative tracers: solving for $J/w$</td>
<td>338</td>
</tr>
<tr>
<td>13.5</td>
<td>Radioactive non-conservative tracers: solving for $w$</td>
<td>340</td>
</tr>
<tr>
<td>13.6</td>
<td>Denouement: computing the other numbers</td>
<td>343</td>
</tr>
<tr>
<td>13.7</td>
<td>Problems</td>
<td>344</td>
</tr>
</tbody>
</table>
### Contents

14 One-dimensional models in sedimentary systems 346
   14.1 General theory 346
   14.2 Physical and biological diagenetic processes 350
   14.3 Chemical diagenetic processes 353
   14.4 A modeling example: CH$_4$ at the FOAM site 356
   14.5 Problems 363

15 Upper ocean 1D seasonal models 365
   15.1 Scope, background, and purpose 365
   15.2 The physical model framework 369
   15.3 Atmospheric forcing 371
   15.4 The physical model’s internal workings 376
   15.5 Implementing the physical model 380
   15.6 Adding gases to the model 387
   15.7 Implementing the gas model 393
   15.8 Biological oxygen production in the model 402
   15.9 Problems 407

16 Two-dimensional gyre models 409
   16.1 Onward to the next dimension 409
   16.2 The two-dimensional advection–diffusion equation 412
   16.3 Gridding and numerical considerations 419
   16.4 Numerical diagnostics 426
   16.5 Transient tracer invasion into a gyre 433
   16.6 Doubling up for a better gyre model 442
   16.7 Estimating oxygen utilization rates 449
   16.8 Non-uniform grids 451
   16.9 Problems 452

17 Three-dimensional general circulation models (GCMs) 453
   17.1 Dynamics, governing equations, and approximations 454
   17.2 Model grids and numerics 463
   17.3 Surface boundary conditions 469
   17.4 Sub-grid-scale parameterizations 474
   17.5 Diagnostics and analyzing GCM output 479

18 Inverse methods and assimilation techniques 489
   18.1 Generalized inverse theory 489
   18.2 Solving under-determined systems 494
   18.3 Ocean hydrographic inversions 499
   18.4 Data assimilation methods 506

19 Scientific visualization 516
   19.1 Why scientific visualization? 516
   19.2 Data storage, manipulation, and access 518
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.3</td>
<td>The perception of scientific data</td>
<td>521</td>
</tr>
<tr>
<td>19.4</td>
<td>Using MATLAB to present scientific data</td>
<td>526</td>
</tr>
<tr>
<td>19.5</td>
<td>Some non-MATLAB visualization tools</td>
<td>530</td>
</tr>
<tr>
<td>19.6</td>
<td>Advice on presentation graphics</td>
<td>532</td>
</tr>
<tr>
<td>Appendix A</td>
<td>Hints and tricks</td>
<td>537</td>
</tr>
<tr>
<td>A.1</td>
<td>Getting started with MATLAB</td>
<td>537</td>
</tr>
<tr>
<td>A.2</td>
<td>Good working practices</td>
<td>539</td>
</tr>
<tr>
<td>A.3</td>
<td>Doing it faster</td>
<td>541</td>
</tr>
<tr>
<td>A.4</td>
<td>Choose your algorithms wisely</td>
<td>543</td>
</tr>
<tr>
<td>A.5</td>
<td>Automating tasks</td>
<td>544</td>
</tr>
<tr>
<td>A.6</td>
<td>Graphical tricks</td>
<td>545</td>
</tr>
<tr>
<td>A.7</td>
<td>Plotting oceanographic sections</td>
<td>547</td>
</tr>
<tr>
<td>A.8</td>
<td>Reading and writing data</td>
<td>549</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>552</td>
</tr>
<tr>
<td></td>
<td>Index</td>
<td>564</td>
</tr>
</tbody>
</table>
Preface

If you are a student of science in the twenty-first century, but are not using computers, then you are probably not doing science. A little harsh, perhaps, and tendentious, undoubtedly. But this bugle-call over-simplification gets to the very heart of the reason that we wrote this book. Over the years we noticed, with increasing alarm, very gifted students entering our graduate program in marine chemistry and geochemistry with very little understanding of the applied mathematics and numerical modeling they would be required to know over the course of their careers. So this book, like many before it, started as a course – in this case, a course in modeling, data analysis, and numerical techniques for geochemistry that we teach every other year in Woods Hole. As the course popularity and web pages grew, we realized our efforts should be set down in a more formal fashion.

We wrote this book first and foremost with the graduate and advanced undergraduate student in mind. In particular, we have aimed the material at the student still in the stages of formulating their Ph.D. or B.Sc. thesis. We feel that the student armed with the knowledge of what will be required of them when they synthesize their data and write their thesis will do a much better job at collecting the data in the first place. Nevertheless, we have found that many students beyond these first years find this book useful as a reference. Additionally, many of our colleagues in the ocean sciences, broadly defined (chemical, biological, geological, and yes, even physical), find this book a useful resource for analyzing or modeling data.

Readers will find this book to be self-contained inasmuch as we introduce all of the concepts encountered in the book, including bringing the reader up to speed on ocean science and physics. Consequently, prerequisites for this book are few. However, exposure to linear algebra, statistics, and calculus sometime in the reader’s past will be helpful, but not absolutely required. Additionally, this book uses MATLAB™ as its computational engine and some programming in MATLAB™ is required; for that reason exposure to programming concepts will be helpful as well. We have chosen MATLAB™ (rather than some other mathematics and statistics package) because we find it subsumes arcane details (e.g. data formats) without concealing the process of analysis. There are a number of very useful MATLAB™ m-files in this book (some written by us, some donated), which we have made available at http://www.cambridge.org/glover, the web page the publisher maintains for this book. These m-files are working, practical examples
Preface

(i.e. code that runs), and each chapter contains detailed problems sets that include computer based assignments and solutions. A fair amount of MATLAB\textsuperscript{TM} instruction occurs throughout the book and in Appendix A, which we call Hints and Tricks, so familiarity with MATLAB\textsuperscript{TM} will be helpful but not required as well.\footnote{MATLAB is a registered trademark of The MathWorks, Inc., of Natick, MA 01760, USA. In order to avoid the appearance of crowded redundancy we are dropping the TM from the name, but when we write MATLAB we are referring to the trademarked product.}

We teach our course in a one-semester blitz divided into three parts. And, yes, taking the course is a little like drinking from a fire hose, but we feel that there is something beneficial about the Zen-like concentration required. The first part of the book deals with the mathematical machinery of data analysis that generally goes under the heading of statistics, although strictly speaking some of it is not really statistics (e.g. principal component analysis). The second part deals with the techniques of modeling that we choose to cover in this book: finite differences, stability analysis, and optimization. The third part of this book deals with case studies of actual, published models, of ever increasing dimensionality and complexity, starting with zero-dimensional models and finishing with three-dimensional general circulation models. Our goal is to instill a good conceptual grasp of the basic tools underlying the model examples. We like to say the book is correct, but not mathematically rigorous. Throughout the book the general principles and goals of scientific visualization are emphasized through technique and tools. A final chapter on scientific visualization reviews and cements these principles.

This book makes a very nice basis for a one- or two-semester course in data analysis and numerical modeling. It begins with data analysis techniques that are not only very useful in interpreting actual data, but also come up again and again in analyzing model output (computa). This first “third” of the book could also be used in a one-semester data analysis course. It begins with an introduction to both MATLAB and singular value decomposition via a review of some basic linear algebra. Next the book covers measurement theory, probability distributions, and error propagation. From here the book covers least squares regression (both linear and nonlinear) and goodness of fit ($\chi^2$). The next analysis technique is principal component analysis which begins with covariance, correlation, and ANOVA and ends with factor analysis. No data analysis course would be complete without a treatment of sequence data starting with auto- and cross-correlation, proceeding through Fourier series and transforms, and optimal filtering, and finishing up with, of course, the FFT. We finish the data analysis third of the book with a chapter on gridding and contouring techniques from simple nearest neighbor methods to objective interpolation (kriging).

The middle third of this book is the transitional segment of any course that attempts to bring together data analysis techniques and numerical modeling. However, this portion of the book can also be used as part of a more traditional course on numerical modeling. We begin with integration of ordinary differential equations and introduce some simple but useful zero-dimensional models. At this point we pause for a chapter and present a tutorial on model building, practical things one needs to consider no matter how simple or complex the problem. We then demonstrate how the parameters in such models can be optimized.
with respect to actual data. When the problems become too complex to be expressed as ordinary differential equations we use partial differential equations, and a discussion and practical introduction to the advection–diffusion equation and turbulence is presented. Next the concept of finite differences is developed to solve these complex problems. In the final chapter of this middle section we cover the important topics of von Neumann stability analysis (Fourier resurfaces), conservation, and numerical diffusion.

We find that, at this point, the students are primed and ready to tackle some “real” models. However, this final third of the book could be used to augment a modeling survey course, although, to get the most out of such a survey the students would need to be well versed in numerical modeling techniques. The book takes the reader through a series of models beginning with simple one-dimensional models of the ocean that rely heavily on lessons learned in the earlier, nonlinear regression section. There are also one-dimensional models of the upper sediment and a very thorough exposition of a one-dimensional, seasonal model of the upper ocean water column. This last section of the book transitions to two-dimensional gyre models and culminates with a chapter on three-dimensional, general circulation models. Up to this point all of these models have been “forward” models. The final third of the book wraps up with “inverse” models. Here we introduce the concepts of inversion and data assimilation and return to the lessons learned from the singular value decomposition chapter at the beginning of the book. It is followed by three-dimensional inversions involving two-dimensional slices of the ocean.

There are certainly many “mathematical methods” books on the market. But this book is the only one we know of that attempts to synthesize the techniques used for analyzing data with those used in designing, executing, and evaluating models. So, where on one’s bookshelf does this volume fit in? It goes in that gap that exists between your copies of Stumm and Morgan’s *Aquatic Chemistry* and Broecker and Peng’s *Tracers in the Sea* on one side, and Pedlosky’s *Geophysical Fluid Dynamics* and Wunsch’s *The Ocean Circulation Inverse Problem* on the other. Although we are, and our examples reflect this, oceanographers, we feel that scientists in other fields will find our explanations and discussions of these techniques useful. For while the density, pressure, and nature of the problem being analyzed and/or modeled may be vastly different from the ones commonly encountered in the pages of this book, the mathematics remain the same.

Over the years we have had a great deal of help (particularly from our students, who take great pride and pleasure in finding mistakes in our notes) in pulling together the information found between the covers of the book you hold in your hand. We thank each and every one of our students, friends, and colleagues who have contributed to the betterment of this work. However, at the end of the day, we take full responsibility for the accuracy of our work, and deficiencies therein are our responsibility.