

## Inverse Modeling of the Ocean and Atmosphere

*Inverse Modeling of the Ocean and Atmosphere* is a graduate-level textbook for students of oceanography and meteorology, and anyone interested in combining computer models and observations of the hydrosphere or solid earth. The scientific emphasis is on the formal testing of models, formulated as rigorous hypotheses about the errors in all the information: dynamics, initial conditions, boundary conditions and data. The products of successful inversions include four-dimensional multivariate analyses or maps of, for example, ocean circulation fields such as temperature, pressure and currents; analyses of residuals in the dynamics, inputs and data; error statistics for all the analyses, and assessments of the instrument arrays or observing systems.

A step-by-step development of maximally-efficient inversion algorithms, using ideal models, is complemented by computer codes and comprehensive details for realistic models. Variational tools and statistical concepts are concisely introduced, and applications to contemporary research models, together with elaborate observing systems, are examined in detail. The book offers a review of the various alternative approaches, and further advanced research topics are discussed.

Derived from the author's lecture notes, this book constitutes an ideal course companion for advanced undergraduates and graduate students, as well as being a valuable reference source for researchers and managers in the theoretical earth sciences, civil engineering, and applied mathematics. Tutors are also directed towards the author's ftp site where they may download complementary overheads for class teaching.

ANDREW BENNETT was awarded a Ph.D. in applied mathematics from Harvard University in 1971. He subsequently continued his research as a National Research Council Fellow at the University of Toronto, and as a Queen's Fellow in Marine Science at Monash University (Melbourne, Australia). Following eight years as a lecturer and senior lecturer in the Department of Mathematics at Monash University, he became a research scientist at the Institute of Ocean Sciences, Sidney, B.C., Canada. He has been a professor at the College of Oceanic and Atmospheric Sciences at Oregon State University since 1987, where his research interests include ocean data assimilation, turbulence theory, and regional modeling. Professor Bennett has won refereeing awards from the *Journal of Physical Oceanography* (1986) and the *Journal of Geophysical Research* (1995), and is also the author of *Inverse Methods in Physical Oceanography* (Cambridge University Press, 1992).

Cambridge University Press  
0521813735 - Inverse Modeling of the Ocean and Atmosphere  
Andrew F. Bennett  
Frontmatter  
[More information](#)

---

# Inverse Modeling of the Ocean and Atmosphere

Andrew F. Bennett  
College of Oceanic and Atmospheric Sciences  
Oregon State University



Cambridge University Press  
 0521813735 - Inverse Modeling of the Ocean and Atmosphere  
 Andrew F. Bennett  
 Frontmatter  
[More information](#)

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE  
 The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS  
 The Edinburgh Building, Cambridge CB2 2RU, UK  
 40 West 20th Street, New York, NY 10011-4211, USA  
 477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
 Ruiz de Alarcón 13, 28014 Madrid, Spain  
 Dock House, The Waterfront, Cape Town 8001, South Africa  
<http://www.cambridge.org>

© Andrew F. Bennett 2002

This book is in copyright. Subject to statutory exception  
 and to the provisions of relevant collective licensing agreements,  
 no reproduction of any part may take place without  
 the written permission of Cambridge University Press.

First published 2002

Printed in the United Kingdom at the University Press, Cambridge

*Typeface* Times 10.25/13.5 pt.    *System* L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub> [TB]

*A catalogue record for this book is available from the British Library*

*Library of Congress Cataloguing in Publication data*

Bennett, Andrew F., 1945–  
 Inverse modeling the ocean and atmosphere / Andrew F. Bennett.  
 p. cm.

Includes bibliographical references and index.

ISBN 0 521 81373 5

1. Oceanography – Mathematical models.    2. Meteorology – Mathematical models.  
 3. Inverse problems (Differential equations)    I. Title.

GC10.4.M36 B47    2002

551.46'001'5118 – dc21

2001052655

ISBN 0 521 81373 5 hardback

Cambridge University Press  
0521813735 - Inverse Modeling of the Ocean and Atmosphere  
Andrew F. Bennett  
Frontmatter  
[More information](#)

---

to Elaine, Luke and Antonia

## Contents

|  |          |
|--|----------|
| <i>Preface</i>   | xiii     |
| <i>Acknowledgements</i>  | xxi      |
| <b>Preamble</b>  | <b>1</b> |
| P.1 Linear regression in marine biology                                    | 1        |
| P.2 Data assimilation checklist  | 4        |
| <b>1 Variational assimilation</b>  | <b>7</b> |
| 1.1 Forward models   | 7        |
| 1.1.1 Well-posed problems  | 7        |
| 1.1.2 A “toy” example  | 8        |
| 1.1.3 Uniqueness of solutions  | 9        |
| 1.1.4 Explicit solutions: Green’s functions                                | 10       |
| 1.2 Inverse models   | 12       |
| 1.2.1 Overdetermined problems  | 12       |
| 1.2.2 Toy ocean data   | 12       |
| 1.2.3 Failure of the forward solution                                      | 13       |
| 1.2.4 Least-squares fitting: the penalty functional                        | 14       |
| 1.2.5 The calculus of variations:<br>the Euler–Lagrange equations          | 15       |
| 1.3 Solving the Euler–Lagrange equations using representers                | 18       |
| 1.3.1 Least-squares fitting by explicit solution<br>of extremal conditions | 18       |

|          |  |           |
|----------|--|-----------|
| 1.3.2    | The Euler–Lagrange equations are a two-point boundary value problem in time                              | 19        |
| 1.3.3    | Representer functions: the explicit solution and the reproducing kernel                                  | 19        |
| 1.4      | Some limiting choices of weights: “weak” and “strong” constraints  | 23        |
| 1.4.1    | Diagonal data weight matrices, for simplicity  | 23        |
| 1.4.2    | Perfect data   | 24        |
| 1.4.3    | Worthless data   | 24        |
| 1.4.4    | Rescaling the penalty functional   | 24        |
| 1.4.5    | Perfect dynamics: Lagrange multipliers for strong constraints  | 25        |
| 1.5      | Regularity of the inverse estimate   | 26        |
| 1.5.1    | Physical realizability   | 26        |
| 1.5.2    | Regularity of the Green’s functions and the adjoint representer functions                                | 26        |
| 1.5.3    | Nondiagonal weighting: kernel inverses of weights  | 27        |
| 1.5.4    | The inverse weights smooth the residuals   | 29        |
| <b>2</b> | <b>Interpretation</b>  | <b>31</b> |
| 2.1      | Geometrical interpretation   | 32        |
| 2.1.1    | Alternatives to the calculus of variations   | 32        |
| 2.1.2    | Inner products   | 32        |
| 2.1.3    | Linear functionals and their representer; unobservables  | 33        |
| 2.1.4    | Geometric minimization with representer  | 34        |
| 2.1.5    | Equivalence of variational and geometric minimization: the data space                                    | 35        |
| 2.2      | Statistical interpretation: the relationship to “optimal interpolation”                                  | 37        |
| 2.2.1    | Random errors  | 37        |
| 2.2.2    | Null hypotheses  | 37        |
| 2.2.3    | The reproducing kernel is a covariance   | 38        |
| 2.2.4    | “Optimal Interpolation”, or best linear unbiased estimation; equivalence of generalized inversion and OI | 38        |
| 2.3      | The reduced penalty functional   | 41        |
| 2.3.1    | Inversion as hypothesis testing  | 41        |
| 2.3.2    | Explicit expression for the reduced penalty functional   | 42        |
| 2.3.3    | Statistics of the reduced penalty: $\chi^2$ testing  | 43        |
| 2.4      | General measurement  | 45        |
| 2.4.1    | Point measurements   | 45        |
| 2.4.2    | Measurement functionals  | 45        |
| 2.4.3    | Representer for linear measurement functionals   | 46        |
| 2.5      | Array modes  | 48        |
| 2.5.1    | Stable combinations of representer   | 48        |

|  |           |
|--|-----------|
| 2.5.2 Spectral decomposition, rotated representers                               | 48        |
| 2.5.3 Statistical stability, clipping the spectrum                               | 49        |
| 2.6 Smoothing norms, covariances and convolutions                                | 51        |
| 2.6.1 Interpolation theory   | 51        |
| 2.6.2 Least-squares smoothing of data; penalties for roughness                   | 52        |
| 2.6.3 Equivalent covariances   | 53        |
| 2.6.4 Embedding theorems   | 55        |
| 2.6.5 Combining hypotheses: harmonic means of covariances                        | 56        |
| <b>3 Implementation</b>  | <b>58</b> |
| 3.1 Accelerating the representer calculation                                     | 59        |
| 3.1.1 So many representers . . .   | 59        |
| 3.1.2 Open-loop maneuvering: a time chart  | 59        |
| 3.1.3 Task decomposition in parallel   | 61        |
| 3.1.4 Indirect representer algorithm; an iterative time chart                    | 61        |
| 3.1.5 Preconditioners  | 62        |
| 3.1.6 Fast convolutions  | 64        |
| 3.2 Posterior errors   | 66        |
| 3.2.1 Strategy   | 66        |
| 3.2.2 Restatement of the “toy” inverse problem                                   | 67        |
| 3.2.3 Representers and posterior covariances                                     | 69        |
| 3.2.4 Sample estimation  | 71        |
| 3.2.5 Memory-efficient sampling algorithm  | 73        |
| 3.3 Nonlinear and nonsmooth estimation   | 74        |
| 3.3.1 Double, double, toil and trouble   | 74        |
| 3.3.2 Nonlinear, smooth dynamics; least-squares                                  | 75        |
| 3.3.3 Iteration schemes  | 76        |
| 3.3.4 Real dynamics: pitfalls of iterating                                       | 78        |
| 3.3.5 Dynamical linearization is not statistical linearization                   | 81        |
| 3.3.6 Linear, smooth dynamics; non-least-squares                                 | 82        |
| 3.3.7 Nonsmooth dynamics, smooth estimator                                       | 83        |
| 3.3.8 Nonsmooth estimator  | 84        |
| <b>4 The varieties of linear and nonlinear estimation</b>                        | <b>86</b> |
| 4.1 State space searches   | 87        |
| 4.1.1 Gradients  | 87        |
| 4.1.2 Discrete penalty functional for a finite difference<br>model; the gradient | 87        |
| 4.1.3 The gradient from the adjoint operator                                     | 89        |
| 4.1.4 “The” variational adjoint method for strong dynamics                       | 90        |
| 4.1.5 Preconditioning: the Hessian   | 92        |
| 4.1.6 Continuous adjoints or discrete adjoints?                                  | 93        |

x Contents

|          |   |            |
|----------|---|------------|
| 4.2      | The sweep algorithm, sequential estimation and the Kalman filter                        | 94         |
| 4.2.1    | More trickery from control theory   | 94         |
| 4.2.2    | The sweep algorithm yields the Kalman filter  | 95         |
| 4.3      | The Kalman filter: statistical theory   | 98         |
| 4.3.1    | Linear regression   | 98         |
| 4.3.2    | Random errors: first and second moments   | 98         |
| 4.3.3    | Best linear unbiased estimate: before data arrive                                       | 99         |
| 4.3.4    | Best linear unbiased estimate: after data have arrived                                  | 100        |
| 4.3.5    | Strange asymptotics   | 102        |
| 4.3.6    | “Colored noise”: the augmented Kalman filter  | 104        |
| 4.3.7    | Economies   | 104        |
| 4.4      | Maximum likelihood, Bayesian estimation, importance sampling<br>and simulated annealing | 105        |
| 4.4.1    | NonGaussian variability   | 105        |
| 4.4.2    | Maximum likelihood  | 105        |
| 4.4.3    | Bayesian estimation   | 111        |
| 4.4.4    | Importance sampling   | 112        |
| 4.4.5    | Substituting algorithms   | 113        |
| 4.4.6    | Multivariate importance sampling  | 114        |
| 4.4.7    | Simulated annealing   | 115        |
| <b>5</b> | <b>The ocean and the atmosphere</b>   | <b>117</b> |
| 5.1      | The Primitive Equations and the quasigeostrophic equations                              | 118        |
| 5.1.1    | Geophysical fluid dynamics is nonlinear   | 118        |
| 5.1.2    | Isobaric coordinates  | 118        |
| 5.1.3    | Hydrostatic balance, conservation of mass   | 119        |
| 5.1.4    | Pressure gradients  | 120        |
| 5.1.5    | Conservation of momentum  | 121        |
| 5.1.6    | Conservation of scalars   | 121        |
| 5.1.7    | Quasigeostrophy   | 122        |
| 5.2      | Ocean tides   | 124        |
| 5.2.1    | Altimetry   | 124        |
| 5.2.2    | Lunar tides   | 124        |
| 5.2.3    | Laplace Tidal Equations   | 125        |
| 5.2.4    | Tidal data  | 127        |
| 5.2.5    | The state vector, the cross-over measurement functional<br>and the penalty functional   | 130        |
| 5.2.6    | Choosing weights: scale analysis of dynamical errors                                    | 131        |
| 5.2.7    | The formalities of minimization   | 134        |
| 5.2.8    | Constituent dependencies  | 135        |
| 5.2.9    | Global tidal estimates  | 135        |



|   | Contents | xi         |
|---|----------|------------|
| 5.3 Tropical cyclones (1). Quasigeostrophy; track predictions   |          | 138        |
| 5.3.1 Generalized inversion of a quasigeostrophic model   |          | 138        |
| 5.3.2 Weak vorticity equation, a penalty functional, the Euler–Lagrange equations   |          | 138        |
| 5.3.3 Iteration schemes; linear Euler–Lagrange equations  |          | 141        |
| 5.3.4 What we can learn from formulating a quasigeostrophic inverse problem   |          | 142        |
| 5.3.5 Geopotential and velocity as streamfunction data: errors of interpretation  |          | 142        |
| 5.3.6 Errors in quasigeostrophic dynamics: divergent flow   |          | 143        |
| 5.3.7 Errors in quasigeostrophic dynamics: subgridscale flow, second randomization  |          | 144        |
| 5.3.8 Implementation; flow charts   |          | 146        |
| 5.3.9 Track prediction  |          | 148        |
| 5.4 Tropical cyclones (2). Primitive Equations, intensity prediction, array assessment  |          | 150        |
| 5.5 ENSO: testing intermediate coupled models   |          | 156        |
| 5.6 Sampler of oceanic and atmospheric data assimilation  |          | 165        |
| 5.6.1 3DVAR for NWP and ocean climate models  |          | 165        |
| 5.6.2 4DVAR for NWP and ocean climate models  |          | 167        |
| 5.6.3 Correlated errors   |          | 168        |
| 5.6.4 Parameter estimation  |          | 168        |
| 5.6.5 Monte Carlo smoothing and filtering   |          | 169        |
| 5.6.6 Documentation   |          | 170        |
| <b>6 Ill-posed forecasting problems</b>   |          | <b>172</b> |
| 6.1 The theory of Oliger and Sundström  |          | 172        |
| 6.2 Open boundary conditions for the linear shallow-water equations   |          | 173        |
| 6.3 Advection: subcritical and supercritical, inflow and outflow  |          | 175        |
| 6.4 The linearized Primitive Equations in isopycnal coordinates: expansion into internal modes; ill-posed forward models with open boundaries |          | 177        |
| 6.5 Resolving the ill-posedness by generalized inversion  |          | 180        |
| 6.6 State space optimization  |          | 182        |
| 6.7 Well-posedness in comoving domains  |          | 183        |
| <b>References</b>   |          | <b>186</b> |
| <b>Appendix A Computing exercises</b>   |          | <b>195</b> |
| A.1 Forward model   |          | 195        |
| A.1.1 Preamble  |          | 195        |

xii Contents

|   |  |            |
|---|--|------------|
| A.1.2   | Model  | 196        |
| A.1.3   | Initial conditions                                   | 196        |
| A.1.4   | Boundary conditions                                  | 196        |
| A.1.5   | Model forcings                                       | 197        |
| A.1.6   | Model parameters                                     | 197        |
| A.1.7   | Numerical model                                      | 197        |
| A.1.8   | Numerical model parameters                           | 199        |
| A.1.9   | Numerical code and output                            | 199        |
| A.1.10  | To generate a plot                                   | 199        |
| A.2   | Variational data assimilation                        | 199        |
| A.2.1   | Preamble   | 199        |
| A.2.2   | Penalty functional                                   | 200        |
| A.2.3   | Euler–Lagrange equations                             | 200        |
| A.2.4   | Representer solution                                 | 200        |
| A.2.5   | Solutions for A.2.2–A.2.4                            | 201        |
| A.3   | Discrete formulation                                 | 205        |
| A.3.1   | Preamble   | 205        |
| A.3.2   | Penalty functional                                   | 205        |
| A.3.3   | Weighted residuals                                   | 206        |
| A.4   | Representer calculation                              | 209        |
| A.4.1   | Preamble   | 209        |
| A.4.2   | Representer vector                                   | 209        |
| A.4.3   | Representer matrix                                   | 209        |
| A.4.4   | Extremum   | 209        |
| A.4.5   | Weights  | 209        |
| A.5   | More representer calculations                        | 210        |
| A.5.1   | Pseudocode, preconditioned conjugate gradient solver | 210        |
| A.5.2   | Convolutions for covarying errors                    | 211        |
| <br><b>Appendix B Euler–Lagrange equations for a numerical weather prediction model</b> |  | <b>212</b> |
| B.1   | Symbols  | 213        |
| B.2   | Primitive Equations and penalty functional           | 214        |
| B.3   | Euler–Lagrange equations                             | 214        |
| B.4   | Linearized Primitive Equations                       | 216        |
| B.5   | Linearized Euler–Lagrange equations                  | 217        |
| B.6   | Representer equations                                | 219        |
| B.7   | Representer adjoint equations                        | 220        |
| <br><i>Author index</i>   |  | 221        |
| <i>Subject index</i>  |  | 225        |

## Preface

Inverse modeling has many applications in oceanography and meteorology. Charts or “analyses” of temperature, pressure, currents, winds and the like are needed for operations and research. The analyses should be based on all our knowledge of the ocean or atmosphere, including both timely observations and the general principles of geophysical fluid dynamics. Analyses may be needed for flow fields that have not been observed, but which are dynamically coupled to observed fields. The data must therefore contribute not only to the analyses of observed fields, but also to the inference of corrections to the dynamical inhomogeneities which determine the coupled fields. These inhomogeneities or inputs are: the forcing, initial values and boundary values, all of which are themselves the products of imperfect interpolations. In addition to input errors, the dynamics will inevitably contain errors owing to misrepresentations of phenomena that cannot be resolved computationally; the data are therefore also required to improve the dynamics by adjusting the empirical coefficients in the parameterizations of the unresolved phenomena. Conversely, the model dynamics must have some credibility, and should be allowed to influence assessments of the effectiveness of observing systems. Finally, and perhaps most compelling of all, geophysical fluid dynamical models need to be formulated and tested as formal scientific hypotheses, so that the development of increasingly realistic models may proceed in an orderly and objective fashion. All of these needs can be met by inverse modeling. The purpose of this book is to introduce recent developments in inverse modeling to oceanographers and meteorologists, and to anyone else who needs to combine data and dynamics.

What, then, is inverse modeling and why is it so named? A conventional modeler formulates and manipulates a set of mathematical elements. For an ocean model, the set includes at least the following:

- (i) a domain in four-dimensional space, representing an ocean region and a time interval of interest;
- (ii) a system of inhomogeneous partial differential equations expressing the phenomenological dynamics of the circulation (there will be inhomogeneities owing to internal fields which force the dynamics beneath the ocean surface, and the equations will include empirical parameters representing unresolved phenomena);
- (iii) initial conditions for the equations, representing the ocean circulation or state at some time; and
- (iv) boundary conditions which may be inhomogeneous owing either to forcing of the ocean at the ocean surface, or to fields of flow and thermodynamic conditions imposed at lateral open boundaries.

There will be subtle yet profoundly important differences for, say, atmospheric models; consider the character of boundary conditions, for example. However, ocean models will be invoked henceforth as the default choice for concise discussion.

From a mathematical perspective, the partial differential operators, initial operators and boundary operators can all be seen as acting in combination upon the solution for the ocean circulation, and producing the inhomogeneities or inputs which are, again, the subsurface forcing, initial values, surface forcing and any values at open boundaries. The combined operator is nonsingular if there exists a unique and analytically satisfactory – say, continuously differentiable – solution for each set of analytically satisfactory inputs. If the operator is nonsingular, then there is a well-defined and unique inverse operator. From the mathematical perspective, the solution is the action of the inverse operator on the inputs. Computing this action is, in the infinite wisdom of convention, “forward ocean modeling”.

Characterizing our knowledge of ocean circulation as solutions of well-posed, mixed initial-value boundary-value problems does not correspond to our real experience of the ocean. Ship surveys, moored instruments, buoys drifting freely on the ocean surface or floating freely below the surface, and earth satellites orbiting above cannot observe continuous fields throughout an ocean region, even for one instant. Yet these data, after control for quality, are in general far more reliable than either the parameterizations of turbulence in the dynamics or the crudely interpolated forcing fields, initial values and boundary values. The quality-controlled data belong to any rational concept of a model, and the set of mathematical elements that defines a model is readily extended to include them. Specifically, functionals corresponding to methods of measurement, and numbers corresponding to measurements of quantities in the real ocean (such as velocity components, temperature, density and the like), may be added to the set. Each functional maps a circulation field into a single number. For example, monthly-mean sea level at a coastal station defines a kernel or integrand which selects sea level from the many circulation variables, which has a rectangular time window of one month and which is sharply peaked at the coastal station. Note that the new mathematical elements

include both additional operators (the measurement functionals), and additional input (the data).

It is always assumed, but almost never proved, that the operator for the original model is nonsingular. It can always be assumed that applying the measurement functionals to the unique solution of the original forced, initial-boundary value problem does not produce numbers equal to the real data. The extended operator can therefore have no inverse, and so must be singular. It seems natural, even a compulsion (Reid, 1968), to determine the ocean circulation as some uniquely-defined best-fit to the extended inputs (forcing, initial, boundary and observed). The singular extended operator then has a generalized inverse operator, and the best-fit ocean circulation is the action of the generalized inverse on the extended inputs. This book outlines the theoretical and practical computation of the action, for best-fits in the sense of weighted least-squares. The practical computations will be only numerical approximations, so the theme of the book should therefore be expressed as “inverting numerical models and observations of the ocean and atmosphere in a generalized sense”. Abandoning precision for brevity, the theme is “inverse modeling the ocean and atmosphere”.

How, then, does inverse modeling meet the needs of oceanographers and meteorologists? The best-fit circulation is clearly an analysis, an optimal dynamical interpolation in fact, of the observations. All the fields coupled by the dynamics are analyzed, even if only some of them are observed. The least-squares fit to all the information, observational and dynamical, yields residuals in the equations of motion as well as in the data, and these residuals may be interpreted as inferred corrections to the dynamics or to the inputs. There are emerging techniques that can in principle distinguish between additive errors in dynamics and internal forcing, but these techniques are so new and unproven that it would be premature, even by the standards of this infant discipline, to include them here. Empirical parameters may also be tuned to improve the analysis. (The tuning game, sometimes described as a “fiddler’s paradise” [Ljung and Söderström, 1987], is outlined here.) The conditioning or sensitivity of the fit to the inputs, as revealed during the construction of the generalized inverse, quantifies the effectiveness of the observing system. The natural choices for the weights in the best fit are inverses of the covariances of the errors in all the operators and inputs. These covariances must be stipulated by the inverse modeler. They accordingly constitute, along with stipulated means, a formal hypothesis about the errors in the model and observations. The minimized value of the fitting criterion or penalty functional yields a significance test of that hypothesis. For linear least-squares, the minimal value is the  $\chi^2$  variable with as many degrees of freedom as there are data, provided the hypothesized means and covariances are correct. A failed significance test does discredit the analyzed circulation and also any concomitant assessment of the observing system, but does not end the investigation: detailed examination of the residuals in the equations, initial conditions, boundary conditions and data can identify defects in the model or in the observing system. Thus model development can proceed in an orderly and objective fashion. This is not to deny the crucial roles of astute and inspired insight in oceanic and

atmospheric model development as in all of science; it is rather to advocate a minimal level of organization especially when inspiration is failing us, as seems to be the case at present.

In spite of an explicit emphasis here on time dependence, the spirit of this approach is close to geophysical inverse theory (see for example Parker, 1994), specifically the estimation of permanent strata in the solid earth using seismic data. The retrieval of instantaneous vertical profiles of atmospheric temperature and moisture using multi-channel microwave soundings from satellites (see for example Rodgers, 2000) bears a striking resemblance to the seismic problem, and is indeed both named and practised as inverse theory. Yet the context here – time-dependent oceanic and atmospheric circulation – is so different that to call it inverse theory seems almost misleading.

Inverse modeling is but one formulation of the vaguely defined activity known as “data assimilation”. The most widely practised form of oceanic or atmospheric data assimilation involves interpolating fields at one time, for subsequent use as initial data in a model integration which may even be a genuine forecast. Once nature has caught up with the forecast, the latter serves as a first-guess or “background” field for the next synoptic analysis. As might be imagined, this cycle of synoptic analysis and forecasting is a major enterprise at operational centers, and is very extensively developed for meteorological applications. Characterization of operational systems for observing the weather, in particular studying the statistics of observational errors, has been and remains the subject of vast investigation. Comprehensive references may be found at appropriate places in the following chapters, but that description of operational detail will not be repeated here. Nor will the intricate, “diagnostically-constrained” multivariate forms of synoptic interpolation be discussed in detail. Geostrophy, for example, is an approximate diagnostic constraint on synoptic fields of velocity and pressure. The emphasis instead will be on elaborating the new data assimilation schemes that could be consistently described as nonsynoptic, “prognostically-constrained” interpolation. The unapproximated law of conservation of momentum, for example, is a prognostic constraint. Again, the nature of this latter activity is so different in technique and broader in scope, in comparison with the conventional cycle of synoptic analysis and forecasting, that to call these new schemes “data assimilation” seems to be misleading yet again. The name “inverse modeling” is chosen, for better or worse.

What else has been left out here? Monte Carlo methods are immensely appealing in any application, and data assimilation is no exception. Sample estimates of means and covariances of circulation fields may be generated from repeated forward integrations of a model driven by suitably constructed pseudo-random inputs. The sample moments of the solutions are then used for conventional synoptic interpolation. The calculus of variations is not required. These assimilation methods, especially “ensemble Kalman filtering”, are highly competitive with variational inverse methods in terms of development effort and computational efficiency, but are even more immature and so are mentioned only briefly. The very basics of statistical simulation and Monte Carlo methods in general are outlined in these chapters.

The reader should not be discouraged by the technical definition of inverse modeling given in previous paragraphs. The calculus of several variables, a rudimentary knowledge of partial differential equations and the same numerical analysis used to solve the forward model are enough mathematics for the computation of generalized inverses. Abstraction is restricted to the one place in this book where an elegant expression of generality is of real benefit. The Hilbert Space analysis sketched in Chapter 2 exposes the geometrical structure of the generalized inverse, and explains the efficiency of the concrete algorithms developed in Chapter 1. The geometrical interpretation is a straightforward adaptation of the theory of Laplacian spline interpolation. A beautiful treatment of L-splines may be found in an applied meteorology journal (Wahba and Wendelberger, 1980), to the eternal credit of the authors, reviewers and editors. Any temptation to make use of the Hilbert Space machinery for abstract definitions of adjoint operators is easily resisted, as such abstraction offers no real insight into the problem of interest. The adjoint operators arise naturally when the elementary calculus of variations is used to derive the classic Euler–Lagrange conditions for the weighted, least-squares best-fit. Unlike the Hilbert Space definition of an adjoint operator, the variational calculus need not be preceded by a linearization of the dynamics and measurement functionals. This flexibility leads to critically important alternatives for iterative solution techniques that are linear.

It is essential to distinguish the formulative and interpretive aspects of inverse modeling from its mathematical aspects. Least-squares may be used to estimate any quantity, but it is the estimator of maximum likelihood for Gaussian or normal random variables. Such variability can reasonably be expected in the ocean and atmosphere, on the synoptic scale and larger, away from transient and semi-permanent fronts, and in variables not subject to phase changes. Least-squares is especially attractive from a mathematical perspective, since it leads to linear conditions for the best fit when the constraints are also linear. The linearity of the extremal conditions permits powerful analyses which yield efficient solution methods. There are many least-squares algorithms, such as optimal interpolation, Kalman filtering, fixed-interval smoothing, and representers. The relationships between these statistical, control-theoretic and geometrical approaches are explained in this book. Aside from unifying the mathematics, recognizing the mathematical relationships facilitates the identification of scientific assumptions.

For example, if the data were collected in much less time than the natural scales of evolution of the dynamics and the internal forcing, then there would be little to gain by assuming that there are errors in the dynamics or internal forcing. It would suffice to admit errors only in the initial conditions, surface forcings, open boundary values and data. This assumption massively reduces the finite dimension of the “state space” for the numerical model, by eliminating those variable fields or “controls” defined both throughout the ocean region and throughout the time interval of interest. Boundary values, initial values and empirical parameters would be retained as controls. The reduced state or “control” space may be sufficiently small that a conventional

gradient search for a minimum in the state space is feasible. The condition of the fit in state space determines the efficiency of the search. It is, however, becoming increasingly necessary to consider time intervals during which the dynamical errors are bound to become significant. That is, the initial conditions would be ineffective as controls for guiding the model solution towards the later data. Using distributed controls, that is, admitting errors in the dynamics throughout space and time, leads to huge numbers of computational degrees of freedom. (There are in general far fewer statistically independent degrees of freedom, but these are not readily identified. Indeed, the methods developed here serve to identify them.) Hence there could be no prospect of a well-conditioned search in the control space or in the equivalent state space. The power of the methods described in these chapters is that they identify a huge subspace of controls (known as the null space) having exactly no influence on guiding the solution towards the data. The methods restrict the search for optimal controls to those lying entirely in the comparatively tiny, orthogonal complement of the null space (known as the data subspace). Again, as in the choice of a least-squares estimator, there is an interplay between scientific formulation and mathematical technique. The two should nonetheless always be carefully distinguished.

As a final example of the distinction and interplay between scientific formulation and mathematical manipulation, consider errors in models of small-scale flows. As already implied, these errors are likely to be highly intermittent or nonGaussian. Thus, inversions of observations collected in mixing fronts and jets, or in free convection, or during phase changes, will require estimators other than least squares. Only brute-force minimization techniques, such as simulated annealing or Monte Carlo methods in general, appear to be available for most estimators. On the other hand, multi-processor computer architecture may favor brute-force inversion. These brute-force techniques will be mentioned here, but only briefly, since by their nature regrettably little is known about them.

The content of this book closely follows an upper-level graduate course for physical oceanography students at Oregon State University. Their preparation includes

- graduate courses in fluid dynamics, geophysical fluid dynamics and ocean circulation theory;
- a graduate course in numerical modeling of ocean circulation;
- a graduate course in time series analysis including Gauss–Markov smoothing or “objective analysis”;
- graduate courses in ordinary and partial differential equations, computational linear algebra and numerical methods in general;
- FORTRAN and basic UNIX skills;
- or an equivalent preparation in atmospheric science.

The curriculum does not require great depth or fresh familiarity with all of the above material. The following would suffice.



1. Some minimal exposure to hydrodynamics, preferably in a rotating reference frame, including approximations such as hydrostatic balance, the shallow-water equations and geostrophic balance. The well-known texts by Batchelor (1973), Pedlosky (1987), Gill (1982), Holton (1992) and Kundu (1990) may be consulted. Graduate students in physics or mechanical or civil engineering would have no problem with the curriculum, although some jargon may cause them to glance at a text in oceanography or meteorology.
2. The knowledge that oceanic and atmospheric circulation models are expressed as partial differential equations (pdes) that may be numerically integrated, most simply using finite differences. The text by Haltiner and Williams (1980) on numerical weather prediction is very useful.
3. Access to Stakgold's classic (1979) text on boundary value problems. The theoretical notions most useful here are (i) odes and pdes can only have well-behaved solutions if precisely the right number of initial and boundary value conditions are provided and (ii) the solution of such well-posed problems for linear odes and pdes can be expressed using a Green's function or influence function. As for computational linear algebra and numerical methods in general, the synopses in Press *et al.* (1986) are very useful.
4. Comfort with the very basics of probability and statistics, including random variables, means, covariances and minimum-variance estimation. Again, the synopses in Press *et al.* (1986) make a good first reading.
5. As much FORTRAN as can be learned in a weekend.

The content of the Preamble, and of each of the six chapters and the two appendices, is outlined on their first pages. The Preamble attempts to communicate the nature of variational ocean data assimilation, or any other assimilation methodology, through a commonplace application of basic scientific method to marine biology. The example might seem out of context, and indeed it is, but that underscores the universality and long history of the approach advocated here. Its arrival in the context of oceanic and atmospheric circulation has of course been delayed by the fantastic mathematical and computational complexity of circulation models. The Preamble includes a "data assimilation checklist", which the student or researcher is encouraged to consult regularly. Chapter 1 is the irreducible introduction to variational assimilation with dynamical models; a "toy" model consisting of a single linear wave equation with one space dimension serves as an illustration. Chapter 2 complements the control-theoretic development of Chapter 1 with geometrical and statistical interpretations; analytical considerations essential to the physical realism of the inverse solutions are introduced. Chapter 3 addresses efficient construction of the inverse and its error statistics, and introduces iterative techniques for coping with nonlinearity. Chapter 4 surveys alternative algorithms for linear least-squares assimilation, and for assimilation with nonlinear or nonsmooth models or with nonlinear measurement functionals. Difficulties to be expected with nonlinear techniques are outlined – proven remedies are still

lacking. Chapter 5 reviews large-scale geophysical fluid dynamics, discusses several real oceanic and atmospheric inverse models in detail, and concludes with notes on a selection of contemporary efforts, both research and operational. Chapter 6 applies inverse methods to forward models based on singular operators.

The material in this book can be presented in thirty one-hour lectures. An overhead projector is a great help: minimal-text, math-only, large-font overhead transparencies allow the audience to listen, rather than transcribe incorrectly. The overheads are available as TEX source files via an anonymous ftp site (<ftp.oce.orst.edu>, <dist/bennett/class/overheads>). Students should be able to begin the computing exercises in Appendix A after studying the first four sections of Chapter 1. The inverse tidal model of §5.2 in Chapter 5 is accessible after studying Chapters 1 and 2. The nonlinear inverse models of tropical cyclones and ENSO in §5.3–5.5, and the accelerated algorithms used in their construction, require a study of Chapter 3. The complete variational equations for the tropical cyclone inversion may be found in Appendix B. A first reading of Chapter 4 is assumed in §5.6, the survey of contemporary applications of advanced assimilation with oceanic and atmospheric data.

The research monograph by Bennett (1992) contains almost all of the theoretical development found here, but none of the guidelines for implementation and few case studies with real data or real arrays. Certain advanced theoretical considerations, such as Kalman filter pathology in the equilibrium limit and continuous families of representers for excess boundary data, are only briefly mentioned here if at all, but may be found in the earlier monograph. There has been a rapid growth in the literature of nonsynoptic data assimilation during the last decade. A full literature survey would be impractical and of doubtful value as so much work has been highly application-specific. Shorter but very useful survey articles include Courtier *et al.* (1993); Anderson, Sheinbaum and Haines (1996) and Fukumori (2001); for collections of expository papers and applications see Malanotte-Rizzoli (1996), Ghil *et al.* (1997) and Kasibhatla *et al.* (2000). The last-mentioned is noteworthy for its interdisciplinary range, and also for a set of exercises on various assimilation techniques. The major text by Wunsch (1996) principally develops in great detail the time-independent inverse theory for steady ocean circulation, using a finite-dimensional formulation which certainly complements the analytical development here and which may be the more accessible for being finite-dimensional. On the other hand the essential mathematical condition of the inverse problem is established at the analytical or continuum level, and the “look and feel” of geophysical fluid dynamics is retained by an analytical formulation.

Inverse modeling suffers not so much from the lack of good data, credible models and adequate computing resources, as from a lack of experience. This book is intended to be of assistance to the generation of investigators who, it is hoped, will acquire that experience.

*Monterey, June 2001*

## Acknowledgements

Many of my collaborators and colleagues over the last two decades have influenced this book, most recently Boon Chua, Gary Egbert and Robert Miller. The exercises in Appendix A were devised and constructed by Boon Chua and Hans-Emmanuel Ngodock. The book is based upon lecture notes for two summer schools on inverse methods and data assimilation, held at Oregon State University in 1997 and 1999. Numerous suggestions from the summer school participants led to significant improvement of the notes. Successive versions were all patiently typed by Florence Beyer, and the book manuscript was created from the notes by William McMechan.

Permission to reproduce copyrighted figures was received from: Inter-Research (Fig. P.1.1), Elsevier (Fig. 5.2.7), the American Geophysical Union (Fig. 5.2.8), Nature (Fig. 5.3.5), Springer-Verlag (Figs. 5.4.1, 5.4.2, 5.4.3, 5.4.5), and the American Meteorological Society (Figs. 5.5.1, 5.5.2, 5.5.3, 5.5.4, 5.5.5, 5.5.6, 5.5.7, 5.5.8). Dudley Chelton kindly provided Fig. 5.2.5. All the other figures were drafted by David Reinert. Thanks for the cover illustrations are owed to Michael McPhaden and Linda Stratton of the NOAA Pacific Marine Environmental Laboratory. The support of the National Science Foundation (OCE-9520956) is also acknowledged.

It has been a pleasure to work with Cambridge University Press: my editors Alan Harvey and Matt Lloyd; also Nicola Stern, Susan Francis, Jo Clegg, and Frances Nex.

The manuscript was completed at the Office of Naval Research Science Unit in the Fleet Numerical Meteorology and Oceanography Center, under the auspices of the Visiting Scientist Program of the University Corporation for Atmospheric Research. I am grateful to Oregon State University for an extended leave of absence, to Manuel

xxii Acknowledgements

Fiadeiro at ONR and Michael Clancy at FNMOC, to Meg Austin and her team at the UCAR VSP Office for their friendly efficiency, and to Captain Joseph Swaykos, USN and his entire staff, for their hospitality during an educational and productive stay at “Fleet Numerical”.

Internet sites mentioned in this book:

[bioloc.coas.oregonstate.edu/pictures/gallery2/index.html](http://bioloc.coas.oregonstate.edu/pictures/gallery2/index.html)  
[ftp.oce.orst.edu, cd/dist/chua/IOM/IOSU](ftp://ftp.oce.orst.edu/cd/dist/chua/IOM/IOSU), 216  
[ftp.oce.orst.edu, cd/dist/bennett/class](ftp://ftp.oce.orst.edu/cd/dist/bennett/class), 146, 195  
[ftp.oce.orst.edu, cd/dist/bennett/class/overheads](ftp://ftp.oce.orst.edu/cd/dist/bennett/class/overheads), xx  
[www.gfdl.gov/~smg/MOM/MOM.html](http://www.gfdl.gov/~smg/MOM/MOM.html), 167  
[www.polar.gsfc.nasa.gov](http://www.polar.gsfc.nasa.gov) 166  
[www.fnmoc.navy.mil](http://www.fnmoc.navy.mil), 166  
[www.pac.dfo-mpo.gc.ca/sci/osap/projects/plankton/zoollab\\_e.htm#Copepod](http://www.pac.dfo-mpo.gc.ca/sci/osap/projects/plankton/zoollab_e.htm#Copepod)  
[www.oce.orst.edu/po/research/tide/global.html](http://www.oce.orst.edu/po/research/tide/global.html), 136  
[www.pmel.noaa.gov/toga-tao/home.html](http://www.pmel.noaa.gov/toga-tao/home.html), 156  
<http://diadem.nersc.no/project>, 169  
[www.units.it/~mabiolab/set\\_previous.htm](http://www.units.it/~mabiolab/set_previous.htm), 1

Note: [oce.orst.edu](http://oce.orst.edu) changes to [coas.oregonstate.edu](http://coas.oregonstate.edu) in 2002.

The publisher has used its best endeavors to ensure that the URLs for external websites referred to in this book are correct and active at the time of going to press. However, the publisher has no responsibility for the websites and can make no guarantee that a site will remain live or that the content is or will remain appropriate.