Observations of ocean circulation have vastly increased during the last two decades, as a result of both international field programs and remote sensors on artificial earth satellites. Oceanographers are turning to inverse methods in order to combine these observations with numerical models of ocean circulation. Recent advances in high-performance computing enable rigorous implementation of inverse methods for massive data sets and complex models. Accordingly this book develops the fundamentals of inverse theory, emphasizing rigor over expedience.

In addition to interpolating the data and adding realism to the model solutions, inverse methods can yield estimates for unobserved flow variables, forcing fields and model parameters. Inverse formulations can resolve ill-posed modeling problems, provide assessment of oceanic observing systems and test models as formal scientific hypotheses.

Ocean models considered range from linear, finite-dimensional systems of equality and inequality constraints to the nonlinear primitive equations. Optimization methods include Kalman filters, descent algorithms and especially representer expansions. Exercises of varying difficulty rehearse technique, and supplement the central theoretical development.

This book is intended for environmental scientists and engineers interested in data analysis and quantitative forecasting, for advanced undergraduates in applied mathematics, and for graduate students in physical oceanography.

CAMBRIDGE MONOGRAPHS ON MECHANICS

INVERSE METHODS IN PHYSICAL OCEANOGRAPHY

Inverse Methods in Physical Oceanography

ANDREW F. BENNETT

College of Oceanography

Oregon State University

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011, 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

First published 1992
Reprinted with corrections 1999, 2003

Typeset in TeX

Library of Congress Cataloging-in-Publication Data

Bennett, Andrew F.

Inverse methods in physical oceanography / Andrew F. Bennett

p. cm. (Cambridge monographs on mechanics and applied mathematics)
Includes bibliographical references (p.) and index.

ISBN 0–521–38568–7 (hc)

1. Oceanography—Mathematics. 2. Inverse problems (Differential equations)
I. Title. II. Series.
GC10.4.M36B46 1992 92–5107
551.46'0015'1535—dc20 CIP

A catalog record for this book is available from the British Library

ISBN 0–521–38568–7 hardback

Transferred to digital printing 2003

CONTENTS

PREFACE

Inverse methods combine oceanic observations with theoretical models of ocean circulation. The methods lead to

- estimates of oceanic fields from sparse data, guided by physical laws;
- estimates of meteorological forcing fields;
- estimates of parameters in the physical laws;
- designs for oceanic observing systems;
- resolution of mathematically ill-posed modeling problems; and
- tests of scientific hypotheses.

The rapid development of inverse methods in physical oceanography over the last decade has been greatly influenced by the work of meteorologists and solid-earth geophysicists; the growing interest of oceanographers in inverse methods is largely in response to the arrival of great volumes of data collected by artificial earth satellites. Collected articles may be found in conference proceedings edited by Anderson & Willebrand (1989), and by Haidvogel & Robinson (1989). Even greater volumes of data are anticipated from planned, future missions. The emergence of inverse theory as a scientific tool has provided a major stimulus to numerical modeling of ocean circulation, which has moved beyond the infancy of thought-experiments but has only tentatively entered the maturity of operational forecasting long occupied by meteorologists. The extraordinary recent gains in computer performance will enable theoretical oceanographers, who have come to inverse theory relatively lately, to consider applying the most elegant and powerful inverse methods to their most complex models. (While this preface was being written, a computer manufacturer announced a 128–gigaflops, 32–gigabyte machine.) In this regard, oceanographers are even more fortunate in that their synoptic time scale is so much longer than the meteorologists' (months, instead of days), so powerful methods may also be brought to bear upon real-time forecasting problems.

The purpose of this book is, therefore, to introduce advanced graduate students, research scientists and that newly emerging professional, the ocean forecaster, to the *possibilities* for inverse theory, rather than the necessities which are the emphasis in Roger Daley's new and important book: *Atmospheric Data Analysis* (Daley, 1991). There the emphasis is almost entirely on objective analysis and the practical initialization problem in operational forecasting, equivalent to the first item in our list.

It is assumed here that the reader has a basic knowledge of geophysical fluid dynamics and ocean modeling. The texts by Pedlosky (1987) and by Pickard & Emery (1990) are excellent references for oceanographic theory and practice; the latter also provides sketches of the major ocean current systems. It is also assumed that the reader is confident in linear algebra and advanced calculus. The material is based largely on lectures given to advanced graduate students in physical oceanography and to advanced undergraduates in applied mathematics. Some tools from functional analyses are exploited (embedding theorems, reproducing-kernel Hilbert spaces and *a priori* inequalities), but only in context and with no formality. In particular, the need for such tools is explained in simple terms. Some of these tools, while being powerful, can also be crude to the point of being misleading. The major analytical tool used throughout is the calculus of variations. An excellent and readily accessible discussion may be found in Courant & Hilbert (1953). A very brief introduction is offered at the end of the fifth chapter here.

This monograph does not review every contribution to the subject. About 50% of the contents derive from the work of the author and collaborators. It is hoped that the result will serve as a coherent presentation of the essential ideas. At this fundamental level, inverse methods in physical oceanography are not so different from those used in geophysics, as described in *Geophysical Data Analysis: Discrete Inverse Theory* (Menke, 1984) and in *Inverse Problem Theory* (Tarantola, 1987). A brief but useful theoretical review for oceanographers (Miller, 1987) may be noted. However, these outstanding texts do not provide for the requirements of oceanographers, that is, detailed analysis of the application of such methods to nonlinear, time-dependent, dynamical models of stratified ocean circulation;

detailed discussion of data requirements, including initial conditions and boundary conditions; a theory for the design of observing systems, and interpretations of simple data-assimilation techniques in terms of inverse methods. Close attention is given to many real inverse calculations drawn from the recent oceanographic literature.

The theoretical discussion is supported with exercises. These do not usually appear in monographs, so some justification is warranted. A number of the exercises merely invite the reader to derive results given in the text; had the full derivations been included, the size of this work would have increased enormously. Other exercises pursue lines of analysis which, while being of interest, were not considered essential. It is hoped that the exercises will encourage the habit of exploring inverse methods in simple settings using elementary analytical techniques, before applying the methods to complex ocean models using sophisticated numerical techniques.

To the extent that this monograph contains essentials, it should be useful for some time. The case studies will quickly become dated, as inverse calculations are made with increasingly more general models and more significant data sets.

The chapters may be summarized as follows:

1. Finite-dimensional inverse theory

Applications of finite-dimensional inverse theory to ocean modeling are discussed, with the objective of introducing some fundamental concepts in a relatively simple setting. These concepts are listed in the introduction to the chapter. Much of the theoretical material is well covered elsewhere, so the discussion of the concepts is brief and directed toward subsequent applications in more complex settings.

2. Smoothing of observations

Two fundamental techniques for estimating fields are discussed here. Both involve least-squares principles.

The first, the well-known "optimal interpolation" or "objective analysis," is based on the Gauss-Markov theorem. The essential concept is unbiased linear regression with minimal error variance. Measurements are expressed as linear functionals acting upon the oceanic field of interest. The necessary prior statistical information will be identified. Simplifying assumptions, such as statistical homogeneity or isotropy, will be discussed along with current practice in estimating the necessary statistics.

The second method, spline smoothing, is less well known to oceanographers but is indispensable later. The powerful constructs known as reproducing kernels and representers are introduced. The remarkable, virtual one-to-one relation between spline smoothing and Gauss-Markov smoothing is delineated.

As a refrain, Wunsch's finite-dimensional inverse methods, described in the preceding chapter using the simple language of linear algebra, are reinterpreted using both linear regression and the theory of representers.

This chapter is also brief, with preparation for subsequent chapters in mind rather than comprehensive presentation.

3. Data assimilation

Oceanographers have already met with some success in ocean forecasting when accurate initial conditions have been available. The quality of these conditions owes much to the smoothing of significant amounts of good initial data. These data may also be used to revise the covariance fields used in the smoothing; the resulting strategy (the Kalman filter) is the optimal "sequential" data-assimilation scheme, in the sense that the growth rate of the forecast error variance is the least. The Kalman filter is an intricate algorithm. It has proved to be so demanding of computing resources (time and memory) that it has, so far, been applied with any success only to simplified ocean models, such as linear models of the equatorial wave guide.

4. The spatial structure of the Kalman filter

An oft-quoted application of the Kalman filter involves forecasting the flight of a point projectile. Accuracy aside, there is no inherent significance in the cross-covariances of the six degrees of freedom which describe the flight. Ocean circulation, on the other hand, is conceived as the flow of a continuum possessing infinitely many degrees of freedom. For example, the spatial distributions of momentum and pressure may be represented as infinite Fourier series. A least-squares best fit to a model and to some data may be achieved with a series which, while being square-summable, might be divergent at isolated points such as observing sites. The corresponding fields of momentum and pressure would be "delta-correlated" in space, or have even worse behavior.

An analysis of the Kalman filter reveals the relationship of the physical realizability of the forecasts to the internal scales of the forecasting model, and to the assumed statistical scales of the model errors. The intuitive sequential assimilation technique known as "nudging" is shown to be a special and typically pathological case of the Kalman filter; the demonstration here uses hypothetical along-track altimetric data.

5. Generalized inverses of dynamical models

Most oceanic data assimilation is hindcasting (after an observing program), so the demand for a sequential assimilation scheme may be relaxed in favor of a best fit over the duration of the program. The fit may include the dynamics (simultaneously with the prior forcing estimate), the prior initial and boundary estimates, and the data. The approach is a natural extension of the inverse formulation in Chapter 1. To a mathematician, a conventional "forward" integration of a model, together with its initial conditions and boundary conditions, amounts to inverting the operator which they define. If the model is well posed then the inverse is non-singular. Specifying additional data would overdetermine the problem, rendering it ill-posed. The operator would then be singular, but would possess a generalized inverse in the least-squares sense. As Reid (1968) remarks, mathematicians must be possessed of a genetic urge to reinvent this least-squares inverse, seemingly without end. The first discussion by a meteorologist or an oceanographer appears to be Sasaki (1955).

The residual of the least-squares best fit, expressed here as the minimal value of a quadratic penalty functional, is the χ^2-variable with one degree of freedom per observation. The proof assumes that the prior estimates for the means and the covariances of the model errors, initial errors, and data errors are all correct. Thus two simple concepts have been extended and combined. The first is the concept of a model as a system of equations and side conditions. The second is the concept of model testing, by simple comparison with data. An inverse formulation combines these two into a null hypothesis concerning the first two moments of the model errors and data errors. Repeated inversions with different forcing estimates, initial values and data allow testing of the hypothesis, by comparing the theoretical and sample distributions of the minimal penalty functional.

For an inverse estimate of ocean circulation to be of greatest use, it should be supported by estimates of the error statistics for the best fit, that is, by estimates of the posterior error covariances. Several methods are available for constructing the generalized inverse and the posterior error statistics. Each method has its adherents. The case is made here for the use of the so-called representer fields. It is argued that they reveal the structure of the inverse, they lead to the best-conditioned algorithm, and they lead to the most efficient computation of posterior error covariances. Representer methods are widely used in solid-earth geophysics.

6. Antenna analysis

A benefit of constructing the inverse of a model and a data set is the opportunity to analyze the efficiency of the "antenna" which collected the data. Here, an antenna might be an array of current meters, or a ship-of-opportunity schedule, or an acoustic tomography web. If the model and antenna are linear, then the efficiency is independent of the actual data values, and may be analyzed prior to an observing program. That is, inverse methods lead to prior experimental design. The design will be sound if the error statistics for the model and the data have been correctly estimated. Several real examples of antennae are analyzed in detail.

7. Nonlinear quasi-geostrophic models

After listing five ways in which an inverse problem can be nonlinear, the chapter focuses on just one of the ways: through advection of the ocean circulation fields, by the circulation itself. The simplest example, based on a nonlinear quasi-geostrophic model in a doubly periodic domain, is examined in detail. The inverse is constructed using an iterative technique, which reduces the nonlinear inverse problem to a sequence of linear inverse problems. It can be proved that the sequence of inverse estimates is bounded, and therefore has limit points. Computations indicate the existence of actual limits, if the Reynolds number is sufficiently small. Posterior error statistics are derived.

8. Open-ocean modeling: quasi-geostrophy

Satellite-borne observing systems are becoming increasingly accurate, and computers become ever faster, yet realistic models of the world ocean circulation are still beyond our reach. Subsurface data are only available in limited regions, while mesoscale ocean circulation (which meteorologists would describe as synoptic-scale circulation) cannot be resolved by present global numerical models running on readily accessible computers. Thus open-ocean modeling remains an important activity. However, boundary data are invariably unreliable or fragmentary. To make matters worse, seemingly well-chosen boundary conditions, even for quasi-geostrophic models, lead to ill-posed problems. We must rely on inverse methods for correcting boundary errors, by exploiting good data collected in the interior of the open-ocean region. Inverse methods also offer some hope for resolving ill-posedness.

Tests of inverse methods must distinguish between shortcomings of the methods and inadequacies of the test data. Really good oceanographic data sets are unavailable, so oceanographers must turn to meteorologists who have data that are incomparably better in both quality and quantity. A regional quasi-geostrophic model for tropical cyclone steering flows is therefore formulated as an inverse problem, and tested with real data. A discussion of the transient-tracer problem follows naturally from the material in this chapter.

9. Primitive-equation models

Open boundary conditions for primitive-equation models are notoriously ill-posed. Some leading meteorologists have declared the problem to be insoluble (Robert & Yakimiw, 1986; Yakimiw & Robert, 1990); they conclude that boundaries for limited-area numerical weather prediction models must be pushed so far out that they have no influence upon the region of interest. This strategy is not available to oceanographers; the size of the model region is restricted by the availability of initial data, especially below the surface. Recourse must be made to inverse formulations, which either resolve conflicts arising from too many boundary conditions or else remove uncertainties arising from too few boundary conditions. Of course, inverse formulations using good data would also lead to more realistic hindcasts, but ill-posedness is the motivating problem in this chapter. The ideal algorithm for solving inverse problems is characterized.

10. Outstanding problems

The β-spiral and finite-dimensional inverse calculations of general circulation described in Chapter 1 are all based on linear dynamical models. Inversion of steady, nonlinear circulation models remains as a major technical challenge. Even if the dynamics of the model are linear, the dependence of the solution upon parameters such as eddy diffusivities is highly nonlinear. Experience in engineering control theory suggests that parameter estimation can be highly unstable.

Theories of error statistics for parameter estimates are linearizations of doubtful utility.

Simple variational inverse methods for ocean models are inadequate if non-smooth circulations are admissible, or if the dynamics are non-smooth owing to branching processes, for example. The processes may represent unresolved scales of motion such as convective mixing. Regardless of the choice of inverse method, real oceanic inverse problems will require massive computing resources, which will only be available from parallel computers. Inverse methods must be adapted to these new architectures.

Finally, and probably the most challenging problem of all, theoretical oceanographers must persuade experimental oceanographers to collect statistics of dynamical errors, so that least-squares inverse calculations may be correctly weighted.

Acknowledgement

Figures 1.2.1 and 1.2.2 are reprinted with the permission of Professor F. Schott and Pergamon Press, 1977. Figures 1.3.2 and 1.3.3 are reprinted with the permission of the American Association for the Advancement of Science, 1977. Figures 1.5.1, 1.5.2, 1.5.3, 1.5.4, 1.6.1 and 1.6.2 are reprinted with the permission of the American Geophysical Union, 1978. Figures 1.7.1, 1.7.2, 1.7.3, 1.8.1, 1.8.2, 1.8.3 and 1.8.4 are reprinted with the permission of the American Geophysical Union, 1987. Figures 2.2.2 and 2.2.3 are reprinted with the permission of the American Meteorological Society, 1976. Figure 2.2.4 is reprinted with the permission of the American Meteorological Society, 1986. Figures 3.2.1 and 3.2.2 are reprinted with the permission of the American Meteorological Society, 1987. Figures 3.2.3 and 3.2.4 are reprinted with the permission of the American Meteorological Society, 1988. Figures 3.5.1, 3.5.2, 3.5.3 and 3.5.4 are reprinted with the permission of the American Meteorological Society, 1989. Figure 6.3.5 is reprinted with the permission of the American Meteorological Society, 1983. Figures 6.3.6, 6.3.7, 6.3.8 and 6.3.9 are reprinted with the permission of the American Geophysical Union, 1990. Figures 6.3.10 and 6.3.11 are reprinted with the permission of Pergamon Press PLC, 1985. Figure 6.3.12 is reprinted with the permission of the American Meteorological Society, 1984. Figure 6.3.13 is reprinted with the permission of Dr. P. McIntosh, 1984. Figure 6.3.14 is reprinted with the permission of the American Meteorological Society, 1987. Figures 6.3.15, 6.3.16, 6.3.17 and 6.3.18 are reprinted with the permission of the American Meteorological Society, 1990. Figures 7.6.1, 7.6.2, 7.6.3, 7.6.4, 7.6.5, 7.6.6, 7.6.7, 7.6.8 and 7.6.9 are reprinted with the permission of the American Meteorological Society, 1992. Figures 7.6.10 and 7.6.11 were prepared by Mr. Edward Zaron, who also provided thoughtful comments. Figures 8.5.2, 8.5.3 and 8.6.1 are reprinted with the permission of *Tellus*, 1988. Figure 9.1.1 is reprinted with the permission of Dr. T. Whitworth and Elsevier Science Publishers, 1990. Figure 10.1.1 is reprinted with the permission of the American Meteorological Society, 1986. Figures 10.2.1, 10.2.2, 10.2.3 and 10.2.4 are reprinted with the permission of the American Geophysical Union, 1991. Figure 10.3.1 is reprinted with the permission of Dr. M. Ghil and Springer-Verlag GmbH & Co. KG, 1987. Figure 10.3.2 is reprinted with the permission of the American Geophysical Union, 1989. Figure 10.4.2 is reprinted with the permission of the American Meteorological Society, 1992.

To my editors, Peter-John Leone and Alan Harvey, and to my production manager, Louise Calabro Gruendel, all at Cambridge University Press: sincere thanks for all their help.

I am grateful to Carl Wunsch for suggesting that I write this book, and for his continuing encouragement. My own interest in inverse methods dates from his celebrated Woods Hole seminar in 1977, on the circulation of the Western North Atlantic. Much of my subsequent education has been stimulated by, I hope, endless arguments with my collaborators Peter McIntosh, Paul Budgell and Robert Miller. Thoughtful comments by Paul Budgell, Michele Rienecker, Leonard Walstad and Carl Wunsch on a draft manuscript were very helpful and greatly appreciated. Special thanks are owed to Bill McMechan for his careful typing, infinite patience and cheerful visits during which heaps of scrawl were exchanged for immaculate text. David Reinart did all the art work.

At various stages of this labor I have been fortunate to have the support of the National Science Foundation (OCE 8800004), the Office of Naval Research (N 00014–90–J–1142) and the National Aeronautics and Space Administration (NAGW–1305, NAGW–2715).

Finally, I wish to thank my wife Elaine for her unfailing support.

Preface to the Second Printing

The potential for inverse theory is now becoming realized: global ocean tides (Egbert, Bennett & Foreman, 1994; Egbert & Bennett, 1996; Egbert, 1998); global numerical weather prediction (Bennett, Chua & Leslie, 1996, 1997), and the tropical Pacific ocean-atmosphere (Bennett et al., 1998). These applications involve models having as many as 10^9 computational degrees of freedom in space and time, and incorporate as many as 10^5 observations. Much of this progress owes to the parallel computing power mentioned in Chapter 10, but an equally significant development has been the indirect use of representers (Egbert et al., 1994). The representer coefficients $\hat{\mathbf{b}}$ may be found without explicit construction of the representers themselves. That is, (5.4.8) may be solved iteratively, given an algorithm that calculates $\mathbf{Pb}$ for any $\mathbf{b}$. As pointed out by Gary Egbert, one backward and one forward integration suffice. The coefficients $\hat{\mathbf{b}}$ having been found, the coupling of the Euler-Lagrange equations (5.3.2) – (5.3.5) is broken, thanks to (5.4.22). A similar approach has been advocated by Amodei (1995). In essence, the penalty functional is being minimized not by a search in the state or control space, but in the space of observable degrees of freedom or 'data space'.

It would be impossible to review here the literature of the 1990's on data assimilation in meteorology and oceanography. Survey articles include Courtier et al. (1993) and Anderson, Sheinbaum & Haines (1996); for collections of papers see Malanotte-Rizzoli (1996) and Ghil et al. (1997). Two major new texts are: Parker (1994) and Wunsch (1996).

Many thanks are owed to my secretary Florence Beyer, to the College artist David Reinert, and to Alan Harvey of Cambridge University Press for assembling the second printing.