Earth System Modeling, Data Assimilation and Predictability

Second Edition

Since the publication of the first edition of this highly regarded textbook, the value of data assimilation has become widely recognized across the Earth sciences and beyond. Data assimilation methods are now being applied to many areas of prediction and forecasting, including extreme weather events, wildfires, infectious disease epidemics, and economic modeling. This second edition provides a broad introduction to applications across the Earth systems and coupled Earth–human systems, with an expanded range of topics covering the latest developments of variational, ensemble, and hybrid data assimilation methods. New toy models and intermediate-complexity atmospheric general circulation models provide hands-on engagement with key concepts in numerical weather prediction, data assimilation, and predictability. The inclusion of computational projects, exercises, lecture notes, teaching slides, and sample exams makes this textbook an indispensable and practical resource for advanced undergraduate and graduate students, researchers, and practitioners who work in weather forecasting and climate prediction.

Eugenia Kalnay completed her PhD at the Massachusetts Institute of Technology (MIT) under Jule Charney and became the first woman on the faculty in the Department of Meteorology. In 1979, she moved to NASA's Goddard Space Flight Center, where she developed the fourth-order global numerical model and led experiments in the new science called "data assimilation." In 1984, she became Head of NASA's Global Modeling and Simulation Branch. In 1987, she became Director of the National Oceanic and Atmospheric Administration's Environmental Modeling Center, where many improvements of models and data assimilation were developed for the National Weather Service forecasts. Her paper "The NCEP/NCAR 40-year reanalysis project" (Kalnay et al., 1996) is the most cited paper in geosciences. In 1997, Kalnay became Lowry Chair at the University of Oklahoma and in 1999 became Atmospheric and Ocean Sciences Department Chair and professor at the University of Maryland, where she was later elected a Distinguished University Professor.

Safa Mote is Assistant Professor of Computational and Applied Mathematics at Portland State University and Visiting Assistant Professor of Atmospheric and Oceanic Sciences at the University of Maryland who has worked on a wide range of challenging interdisciplinary problems. He has two PhD degrees in Physics and in Applied Mathematics and Statistics, and Scientific Computing (AMSC) from the University of Maryland. He designs mathematical models to propose and assess holistic policies that lead to sustainability in interconnected environmental, economic, climate, and health systems. He develops computational methods based on Dynamical Systems, Machine Learning, and Data Assimilation to forecast extreme weather and climate events, improve subseasonal to seasonal predictions, and create projections for the coupled energy–water–food nexus.

Cheng Da works on Coupled Data Assimilation as a postdoctoral research associate at the University of Maryland and the Global Modeling and Assimilation Office at NASA's Goddard Space Flight Center. Supported by the NASA Earth and Space Science Fellowship, he earned his PhD degree under the supervision of Professor Kalnay at the University of Maryland, focusing on the assimilation of precipitation and nonlocal observations in the ensemble data assimilation system and coupled data assimilation. Before this, he earned his bachelor's and Master's degrees in Meteorology at Florida State University, working on radiance assimilation from spaceborne sensors.

Earth System Modeling, Data Assimilation and Predictability

Atmosphere, Oceans, Land and Human Systems

Second Edition

EUGENIA KALNAY
University of Maryland, College Park, MD, US

SAFA MOTE
Portland State University, Portland, OR, US
University of Maryland, College Park, MD, US

CHENG DA
University of Maryland, College Park, MD, US
NASA Goddard Space Flight Center, Greenbelt, MD, US
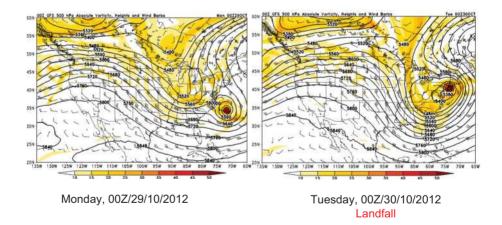
CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

**This book is dedicated to
the Grandmothers of Plaza de Mayo
for their tremendous courage and leadership in defense of human rights
and democracy.**

Front Cover Legend: Importance of ensemble forecasting and data assimilation for the forecast of the Hurricane Sandy landing in the New York metro area in 2012 (The actual trajectory of Sandy is shown with a thick black line with a black circle every 6 hours indicating when a new "assimilation of the observation of Sandy's location" took place). This figure, created by Clark Evans, Professor of Atmospheric Sciences at the University of Wisconsin–Milwaukee, (evans36@uwm.edu), shows the National Weather Service ensemble prediction of the devastating Sandy Hurricane landing.

Note that the ensemble trajectories of the NCEP 10-day ensemble forecasts of the Sandy hurricane trajectory were started on October 23, 12UTC (shown with blue lines at the first identification of Sandy), and 6 and 12 hours later (shown with green and red lines respectively). The majority of the earliest (in blue) ensemble forecasts miss the hurricane being "captured" by a strong atmospheric trough (see the inside cover of the book) and continue moving eastwards, driven by the Atlantic westerly winds, as the majority of the Atlantic hurricanes normally do. The green trajectories, that after 6 hours underwent one additional data assimilation, clearly turned west, indicating the influence of the trough that captured the hurricane. The red trajectories, started October 24 00UTC, from the next data assimilation just 12 hrs after the blue trajectories, have a majority that correctly turns west for the landfall.

Why did Sandy turn west around 00Z/29/10/2012?
It was captured by a deep trough!

Monday, 00Z/29/10/2012   Tuesday, 00Z/30/10/2012
                                          Landfall

500 hPa NCEP analysis of absolute vorticity, winds, and heights

Contents

Foreword to the First Edition

During the 50 years of numerical weather prediction, the number of textbooks dealing with the subject has been very small, the latest being the 1980 book by Haltiner and Williams. As you will soon realize, the intervening years have seen impressive developments and success. Eugenia Kalnay has contributed significantly to this expansion, and the meteorological community is fortunate that she has applied her knowledge and insight to writing this book.

Eugenia was born in Argentina, where she had exceptionally good teachers. She had planned to study physics but was introduced to meteorology by a stroke of fate; her mother simply entered her in a competition for a scholarship from the Argentine National Weather Service! But a military coup took place in Argentina in 1966 when Eugenia was a student, and the College of Sciences was invaded by military forces. Rolando Garcia, then Dean of the College of Sciences, was able to obtain for her an assistantship with Jule Charney at the Massachusetts Institute of Technology. She was the first female doctoral candidate in the department and an outstanding student. In 1971, under Charney's supervision, she finished an excellent thesis on the circulation of Venus. She recalls that an important lesson she learned from Charney at that time was that if her numerical results did not agree with accepted theory, it might be because the theory was wrong.

What has she written in this book? She covers many aspects of numerical weather prediction and related areas in considerable detail, on which her own experience enables her to write with relish and authority. The first chapter is an overview that introduces all the major concepts discussed later in the book. Chapter 2 is a presentation of the standard equations used in atmospheric modeling, with a concise but complete discussion of filtering approximations. Chapter 3 is a roadmap to numerical methods providing students without a background in the subject with all the tools needed to develop a new model. Chapter 4 is an introduction to the parameterization of subgrid-scale physical processes, with references to specialized textbooks and papers. I found her explanations in Chapter 5 of data assimilation methods and in Chapter 6 on predictability and ensemble forecasting to be not only inclusive but thorough and well presented, with good attention to historical developments. These chapters, however, contain many definitions and equations. (I take this wealth as a healthy sign of the technical maturity of the subject.) This complexity may be daunting for many readers, but this has obviously been recognized by Eugenia. In response, she has devised many simple graphical sketches that illustrate the important relations and definitions.

An added bonus is the description in an appendix of the use of *Model Output Statistics* by the National Weather Service, its successes, and the rigid constraints that it imposes on the forecast model. She also includes in the appendices a simple adaptive regression scheme based on Kalman filtering and an introduction to the generation of linear tangent and adjoint model codes.

Before leaving the National Centers for Environmental Prediction (NCEP) in 1998 as Director of the Environmental Modeling Center, Eugenia directed the *Reanalysis Project*, with Robert Kistler as Technical Manager. This work used a 1995 state-of-the-art analysis and forecast system to reanalyze and reforecast meteorological events from past years. The results for November 1950 were astonishing. On November 24 of that year, an intense snowstorm developed over the Appalachians that had not been operationally predicted even 24 hours in advance. This striking event formed a test situation for the emerging art of numerical weather prediction in the years immediately following the first computations in 1950 on the ENIAC computer discussed in Chapter 1. In 1953, employing his baroclinic model, and with considerable "tuning," Jule Charney finally succeeded in making a 24-hour forecast starting on November 23, 1950, of a cyclonic development, which, however, was still located some 400 kilometers northeast of the actual location of the storm. This "prediction" played a major role in justifying the creation of the Joint Numerical Weather Prediction Unit in 1955 (Chapter 1). By contrast, in the Reanalysis Project, this event was forecast extremely well, in both intensity and location – as much as three days in advance. (Earlier than this the associated vorticity center at 500 mb had been located over the Pacific Ocean, even though at that time there was no satellite data!) This is a remarkable demonstration of the achievements of the numerical weather prediction community in the past decades, achievements that include many by our author.

After leaving NCEP in 1998, Eugenia was appointed Lowry Chair in the School of Meteorology at the University of Oklahoma, where she started writing her book. She returned to Maryland in 1999 to chair the Department of Meteorology, where she continues to do research on a range of topics, including applications of chaos to ensemble forecasting and data assimilation. We look forward to future contributions by Professor Kalnay.

Norman Phillips

Preface to the Second Edition

It has been two decades since Cambridge University Press published the first edition of my book on data assimilation in 2003. Its title was *Atmospheric Modeling, Data Assimilation and Predictability*, but it was always referred to as "the Data Assimilation book." The book was very well received by numerical weather prediction graduate students and researchers, with ~4,500 citations at this time. In the section that follows, "Reviews and Comments on the First Edition," we have included four book reviews as well as a sample of other reviews and endorsements of the first edition.

Since 2003, the ensemble-based assimilation methods (e.g., ensemble Kalman filter and particle filters) have evolved very fast, and the ensemble Kalman filter has been implemented in several operational centers. So, the most important chapter, 5 "Data Assimilation" (Chapter 5), has been completely rewritten. We have also introduced a new toy model (a stone in space) that makes the complex data assimilation equations easy to understand. Chapter 5 now discusses fundamental aspects of variational and ensemble methods and how to combine them (hybrid methods) to further improve the analysis. A short introduction to particle filters is also included.

The other chapters have also been updated. Chapter 1, "An Overview of Numerical Weather Prediction," which includes a history of numerical weather prediction and major developments, was rewritten to be accessible to an undergraduate-level reader. In Chapter 2, "The Continuous Equations," we have added an introduction to the governing equations of the oceans, Kelvin waves, and equatorially trapped waves. Chapter 3, "Numerical Discretization of the Equations of Motion" has also been updated. We have added sections on nonhydrostatic models and on the need to replace the long and successfully used spectral models because of the increased model resolution, and we discuss the next-generation global model at the National Centers for Environmental Prediction (NCEP). In Chapter 4, "Introduction to the Parameterization of Subgrid-Scale Physical Processes," we have added a section on the Simplified Parameterizations primitivE-Equation DYnamics (SPEEDY) model as an introduction to the dynamical core and the physical parameterizations of a full spectral Earth system model. A new section on cumulus parameterizations and super parameterization has also been included. Although we generally use the atmosphere as an example, the methodologies and the way of thinking presented in Chapters 2–6 are also applicable to ocean and land modeling data assimilation.

To accompany this book, we have provided additional learning materials (such as the computational project, as well as additional book sections) on the companion website. We will also continue to add new projects here.

During the past two decades, I have had the privilege of teaching at the University of Maryland (UMD) and the good fortune of being the advisor of about 40 students who have earned their PhDs, learning from them, and seeing many of them become recognized as leading experts in data assimilation. I have also developed a keen interest in the interaction of the Earth system with the human system, the overwhelming growth of the impact of the human system on the Earth system, and the impacts of inequality. I started working with Safa Mote (formerly Safa Motesharrei), a brilliant student who was equally interested in this problem. Jorge Rivas also had a keen interest in and deep knowledge of this subject and soon joined us. Together, we have written several papers, the first one being the highly cited human and nature dynamical (HANDY) model, a simple approach to modeling the interaction between the human and natural systems (Motesharrei et al., 2014). In 2016, we published another paper on "Modeling sustainability" (Motesharrei et al., 2016) with 20 coauthors, most of them leaders in their fields. We pointed out in this paper that since in the real world the human system has become the main driver of change in many of the physical subsystems of the Earth system, in order to understand the dynamics of *either* system, Earth system models *must* be coupled with human system models through bidirectional couplings representing the positive, negative, and delayed feedbacks that exist in the real systems. We also discuss applications of data assimilation methods to these coupled Earth–human system models. These include parameter estimation, sensitivity analysis, and ensemble runs to quantify uncertainty. We thus decided to change the title of the book to *Earth System Modeling, Data Assimilation, and Predictability*. Earth system models are generally considered to be the coupled modeling of the atmosphere, land, and oceans, as well as other subsystems, such as the biosphere and the cryosphere. However, the human system, which has come to completely dominate the evolution of the Earth system, has not yet been bidirectionally coupled to these natural systems (Motesharrei et al., 2016; Calvin and Bond-Lamberty, 2018). Our first HANDY model (Motesharrei et al., 2014), which models the dynamic interactions between the human and natural systems before the Industrial Revolution, is a minimal example of a coupled Earth–human system. In Chapter 6, "Atmospheric Predictability and Ensemble Forecasting," we now include several sections discussing the development of coupled Earth–human system modeling and how to use data assimilation to improve this modeling. In order to address the problems of climate change, it is both feasible and necessary to replace fossil fuels with renewables (e.g., Jacobson et al., 2015, 2018; IPCC, 2022). Modeling the Earth and human systems bidirectionally coupled together is necessary to inform policies that allow a timely transition to renewables.

In the last few years, artificial intelligence and machine learning, and particularly deep learning, have advanced extremely rapidly and offer the promise of being as accurate but much faster than standard numerical methods used in modeling and DA (e.g., Krasnopolsky, 2013). We emphasize in Chapter 3 the need to evaluate their performance in terms of not just comparing their RMS errors but, very importantly, whether

the new methods have the essential ability to reproduce the atmospheric instabilities, measured, for example, using the breeding method.

I am very fortunate that Dr. Cheng Da and Dr. Safa Mote agreed to become my coauthors.

Cheng Da is an expert in (and has a passion for) data assimilation, something that he jokes is due to his last name. He earned his PhD degree under my supervision at the University of Maryland, focusing on precipitation assimilation to improve the prediction of tropical cyclones and a new multilayer observation localization method to accurately assimilate nonlocal observations in the local ensemble transform Kalman filter (LETKF). Before this, he earned his bachelor's and master's degree in meteorology at Florida State University. He is now working on coupled data assimilation as a postdoctoral research associate at the University of Maryland and the Global Modeling and Assimilation Office at NASA's Goddard Space Flight Center.

Safa Mote is an applied mathematician and Earth systems scientist who has expertise in dynamical systems and climate modeling, computational and data science, and data-driven prediction. He builds mathematical models to propose and assess holistic decisions and policies that lead to sustainability of the coupled climate–energy–water–food nexus. He also develops novel computational methods by combining dynamical systems, machine learning, and data assimilation to analyze large, diverse datasets of these complex systems and to create forecasts and projections. He is particularly interested in forecasting, and possibly mitigating, high-impact severe weather events such as wildfires, droughts, floods, and storms. I am deeply grateful to Professor Mote for presenting frequent guest lectures over the past decade in my graduate courses on applied statistics, Earth system modeling, and data assimilation.

Eugenia Kalnay

Reviews and Comments on the First Edition

Book Review by Dr. Andrew Lorenc

The original book review was published in the *Quarterly Journal of the Royal Meteorological Society*, 2003.

Eugenia Kalnay's enthusiasm for numerical weather prediction (NWP) shows through from the first chapter of her book. The introductory historical overview, from the first successes of Charney and others in the 1950s to modern ensemble forecasting systems, is designed to motivate graduate-level students by explaining how the rapid progress was achieved, and listing the outstanding challenges. Eugenia then sets out to give a solid grounding in all aspects of NWP, with chapters on the continuous equations, their numerical discretization, the parametrization of subgrid-scale physical processes, data assimilation and predictability, and ensemble forecasting. She uses simple examples and exercises to cover the ground from first principles to state-of-the-art NWP systems. Other exercises are more like essay topics, designed to make the reader think.

Chapter 2 is a concise but thorough exposition of the basic equations used for NWP, with discussion, through simple examples, of the types of waves which can occur, leading to understanding of the various filtering approximations. Chapter 3 is equally thorough in covering the numerical techniques used to construct forecast models. A student working through the examples and exercises should emerge confident to work on practical model development. Of course a book of this length cannot cover everything in depth; Chapter 4 is only a nine-page summary of parametrization methods, with references for more details but no exercises. In contrast, Chapter 5 on data assimilation is the longest in the book. Early empirical methods such as successive correction and nudging are described, before embarking, via simple examples, on the least-squares and Bayesian derivations of modern statistical methods. Advanced methods with evolving forecast error covariance are covered, with a mention of current research interest in the ensemble Kalman filter. Chapter 6 on predictability is another highlight, going from Lorenz's butterfly, through chaos, to modern ensemble forecasting systems.

Although she has succeeded in covering most aspects of NWP in a single book, Eugenia's background and particular enthusiasms show through in the emphasis given to each topic. For instance, 4D-Var is just described in the advanced data-assimilation methods section, without mentioning its success in operational use, and breeding is given more space than singular vectors in the section on ensembles. Yet her enthusiasm

is a strength of the book. One is left with a feeling for the interesting problems that remain. As a broad introduction to the basis of modern NWP, this book has no equal.

Book Review by Dr. Peter Lynch

The original book review has been published in *Splanc*, Met Éireann Newsletter, 2003.

Computer modelling is now the primary means of forecasting the weather. The accuracy of Numerical Weather Prediction models has improved steadily over the half-century since the first tentative experiments. There are a huge number of technical papers and reports devoted to NWP, but very few books. So, a frisson of excitement accompanied the rumour that Eugenia Kalnay was writing a new book. Expectations were high, since she is a renowned expert in the field. She has not disappointed us.

Eugenia Kalnay got a PhD in MIT in 1971. Her advisor was Jule Charney (Ray Bates, a former Met Éireann Assistant Director, was another of his doctoral students). She was Director of the Environmental Modeling Center, U.S. National Weather Service, for ten years to 1997, so she knows a thing or two about NWP. Her book is called Atmospheric Modeling, Data Assimilation and Predictability, and covers all three topics in the title. It is a mathematical book, but don't stop reading, because it has a considerable amount of expository material, accessible to all. I am reviewing it here in the hope of persuading mathophobes (pace Mr. Gates) to look through the book. They will gain much by perusing the discursive sections even if they skip the sums.

The first of the six chapters is a historical survey, tracing the development of NWP from the ENIAC integrations up to the present, and ending with a peek into the future. It is almost completely non-mathematical and is eminently readable. The next two chapters deal with the equations of motion and the numerical methods of solving them. These are tough going, but are worth scanning, as there are some descriptive sections. In chapter 3.5, on regional models, Eugenia refers to a report by Aidan McDonald as an excellent review of the lateral boundary conditions used in operational regional NWP. Chapter 4 is an admirably succinct (i.e., mercifully short) introduction to the huge area of subgrid-scale physical processes. Again, much of it is accessible to the general reader.

Chapter 5, the longest in the book, is also the toughest, dealing with the vital subject of data assimilation. The material here is inherently difficult. However, a superficial scan will give you an overview of the subject, and many buzz-words to amaze your friends. The final chapter is on predictability and ensemble forecasting.

Forecasters should find this particularly useful and relevant. We have emerged from denial (or was it the Amazon?), and recognise the limitations on our predictive abilities. The chaotic nature of the atmosphere is unavoidable. We can forecast with probability but never with certainty, particularly at longer lead-times. Operational ensemble forecasting is reviewed in this chapter. The concluding sections are on the role of the oceans in long-range forecasting and on climate change. The bibliography is comprehensive, running to 44 pages. More than a dozen references are to work of current or former Met Éireann scientists. This is gratifying, indicating that a small,

dedicated team can have an international impact. Some of the numerical techniques developed in Dublin, and now in widespread use, are described in this excellent book.

Book Review by Dr. Takemasa Miyoshi

The original book review in Japanese was published in the December 2003 issue of *Tenki*, the monthly bulletin of the Meteorological Society of Japan. ©Meteorological Society of Japan. Used with Permission. (Translated by Dr. Takemasa Miyoshi.)

Wide range of topics on numerical weather prediction (NWP) including its history, techniques in atmospheric modeling, data assimilation, and predictability are discussed in clear, self-contained, and precise way. The outstanding advantage is that basic ideas as well as details of techniques are included, which makes the discussion clearer. Not only readers new to NWP but also knowledgeable readers must be satisfied with the richness of the contents.

There has been a well-known textbook on NWP written by Haltiner and Williams (1980), but it is true that much progress including the realization of ensemble forecasting has been made thanks to the progress in the computational capability. Based on the latest situation, the present work describes the basics of atmospheric dynamics and dynamical equations in NWP, and in addition, the details of techniques in numerical integration methods. For instance, various numerical discretization schemes, Galerkin and spectral methods, semi-Lagrangian schemes, staggered grid structures, which are important in implementing NWP models, are precisely described. Furthermore, parameterization of subgrid-scale physical processes is described briefly. As for data assimilation, the discussion covers the basic idea of least square method, and the practical methods in realistic high-dimensional system including the optimum interpolation (OI), the three-dimensional variational method (3D-VAR), the four-dimensional variational method (4D-VAR), and the Kalman filtering, some of which are operationally used and some of which are latest research topics. There is no need to mention that initialization methods are included in the discussion. In the last chapter, predictability and ensemble forecasting, in which the author has been deeply involved, are discussed. The general discussion of chaos of nonlinear dynamical systems using a simple model is introduced, and the purpose of ensemble prediction is clearly explained. It follows that the basics of the typical error vectors such as singular vectors and Lyapunov vectors are explained, and then, the ensemble forecasting in the actual NWP cases is discussed. The discussion includes the Monte-Carlo method, the Lagged Averaged Forecasting (LAF) method, the breeding method, the singular vector method, and multimodel method. Climate change is also included, which is surprising because of the wide range of its topics.

This clear, precise, and comprehensive discussion in NWP was possible thanks to the author who had been involved in the development of NWP models at the National Centers for Environmental Prediction (NCEP) for many years. The author, Eugenia Kalnay, was a mentee of the famous Jule Charney, the "father" of NWP, and she is currently involved in the education and research in NWP at the University of Maryland,

College Park. The present work is based on her lecture notes, which is why the discussion is kind and precise. Moreover, representative papers in each field are included in the reference, which makes it best for beginners. For readers who have knowledge and experience in NWP, it helps to organize their knowledge, and the scales fall from their eyes because of the author's wisdom which has influenced places all over the work.

In conclusion, the present work crystallizes the author's wisdom, and a must in studying NWP.

Book Review by Dr. Christopher K. Wikle

The original book review was published in *Technometrics*, 2005.

With increasing concerns about potential climate change and its associated impacts, questions of uncertainty in climate prediction are becoming paramount. Naturally, such concerns about uncertainty suggest the need for statistical expertise. In addition, statisticians interested in problems in the environmental sciences (e.g., meteorology, oceanography, ecology) are incorporating deterministic dynamical models (e.g., partial differential equations) in their statistical models for spatiotemporal processes. Because it is relatively uncommon for statisticians to receive training in dynamical systems and partial differential equations (PDEs) in their formal academic courses, there is a need for a concise, yet thorough book that describes the critical aspects of these subjects, at least those related to environmental processes. This book does exactly that.

In essence, this book is an overview of modern numerical weather forecasting. After an accessible historical overview in Chapter 1, Chapter 2 provides a summary of the continuous system of equations that describe the state of the atmosphere. Although somewhat terse for a statistician with no previous exposure to the governing equations of the atmosphere, it is an excellent overview of atmospheric dynamics. Chapter 3 then describes the numerical solution of such systems of PDEs. This is an excellent chapter and would be a nice reference for statisticians interested in numerical solutions of PDEs, as well as those interested in the connection between PDEs and difference equations. Chapter 4 is a very short introduction to the parameterization of subgrid-scale processes that are necessary in atmospheric models as a result of the discretization limits imposed on the continuous system. Such parameterizations are critically important in weather forecasting and climate modeling, yet are rarely considered from a statistical perspective (either in terms of estimation of free model parameters or in terms of the uncertainty in the model specification).

Chapters 5 and 6 are more directly relevant to statistics. Chapter 5 gives a comprehensive overview of "data assimilation," which involves combining uncertain observations with deterministic (or quasi-deterministic) dynamical models to obtain an estimate of the state of the system (and its uncertainty). For numerical weather forecasting, this is critical. To initialize the discretized deterministic weather forecasting model, one must have an estimate of the system state variables at each grid location (in three-dimensional space). Obviously, atmospheric observations do not occur at such

resolution, so one must effectively interpolate the observations to obtain the initial state. The problem is complicated by the presence of large gaps of missing data and the fact that the initial state must be in some sort of dynamical balance (i.e., must be a physically realistic representation of the system state). This requires using a dynamical model to help fill in the missing information. Statisticians are familiar with these ideas in the context of spatial prediction (i.e., kriging). More generally, data assimilation can be posed from a Bayesian perspective in which one has a prior distribution for the state process that might be obtained from a numerical weather prediction model. The "likelihood" is then the distribution of the data conditional on this true process. One of the biggest challenges with this procedure is obtaining the "prior" covariance matrix. In the case of normal assumptions for both the prior and "likelihood," both the state vector and the associated covariance matrix can be obtained via the Kalman filter. However, with the high-dimensional state vectors common to atmospheric systems (on the order of millions!), and the nonlinearity of the dynamical evolution equations, standard Kalman filter methods are not practical. This book presents the outline of so-called "ensemble" Kalman filters, which overcome many of these problems and are currently the focus of intense research and development because of their outstanding potential for practical assimilation problems.

Chapter 6 is concerned with atmospheric predictability and ensemble forecasting. From a statistical perspective, this chapter provides a very nice and concise review of the basics of chaotic systems and such important concepts as singular vectors and Lyapunov vectors. The growth of errors in nonlinear dynamical systems is critical for prediction, and these concepts provide a fundamental way to evaluate predictability. There is also a discussion of the necessity of accounting for the uncertainty in the initial state for nonlinear prediction models due to chaos. This leads to the idea of ensemble forecasting, which is closely tied to the ensemble methods for data assimilation presented in Chapter 5.

In summary, this book is an excellent reference for statisticians interested in dynamical systems and/or spatiotemporal processes in the environment. The technical level is reasonable for statisticians with advanced degrees. Although a few token exercises are included, to use this book as a textbook the instructor would need to provide supplemental problems.

Short Reviews and Endorsements

. . . quite wonderful, achieving a tremendous balance between comprehensiveness and readability. I am especially pleased with the numerical analysis part, which is crystal clear and shows the benefits of classroom testing. I also like the tiny little touches, like the stepped-on butterfly story and the mention that Poincaré knew about chaos in celestial mechanics. Your book fills an enormous hole in the literature of NWP [numerical weather prediction].
Richard C . J. Somerville, Scripps Institution of Oceanography, San Diego

Fantastic . . . in content, format and practicability.
Kelvin K. Droegemeier, Regents' Professor of Meteorology, and Director, Center for Analysis and Prediction of Storms, University of Oklahoma

[I] admire the clarity and pedagogic superiority of [this] presentation.
Anders Persson, Swedish Meteorological and Hydrological Institute (SMHI)

. . . much better for learning about data assimilation than anything else currently available.
Richard Swinbank, United Kingdom Meteorological Office

. . . [the] presentation is impeccable and is very accessible to non-meteorologists like me.
Eric Kostelich, University of Arizona

. . . what a great wealth of historical information.
Lawrence Takacs, NASA, Data Assimilation Office

. . . [the] method in the [data] assimilation section of starting with 'baby' examples, and then working up through the full analysis, is great for understanding. On the predictability part, the history, and the explanations of how the unstable perturbations grow is the best I've seen.
Alexander E. MacDonald, Director, NOAA Forecast Systems Lab

. . . this book . . . is extremely useful, informative, and well-written . . . there are many instances where items that were only marginally familiar beforehand have now become very clear.
Brian O. Blanton, Senior Scientist/Oceanographer, University of North Carolina at Chapel Hill

Acknowledgments from the Second Edition

This book would not have been possible without tremendous support and helpful advice and feedback from numerous colleagues and students. We are grateful to numerous colleagues at the University of Maryland, Portland State University, the National Aeronautics and Space Administration (NASA), and many other institutions: Tony Busalacchi, Bob Cahalan, Mark Cane, Alberto Carrassi, Jim Carton, Rita Colwell, Dacian Daescu, Russ Dickerson, Paolo D'Odorico, Bill Dorland, Kelvin Droegemeier, Julien Emile-Geay, Clark Evans, Kuishuang Feng, Rachel Franklin, Inez Fung, Michael Ghil, Mitch Goldberg, Chris Jarzynski, Eric Hackert, Milt Halem, Pedram Hassanzadeh, Ross Hoffman, Klaus Hubacek, Brian Hunt, Eric Kostelich, Fred Kucharski, Doron Levy, Michel Loreau, Andrew Lorenc, Peter Lynch, Michael Mann, Fernando Miralles-Wilhelm, Takemasa Miyoshi, Sumant Nigam, Carlos Nobre, Ed Ott, Tim Palmer, George Philander, Drew Rice, Jorge Rivas, Raj Roy, Alfredo Ruiz-Barradas, Matthias Ruth, Roald Sagdeev, June Sherer, Adel Shirmohammadi, Jagadish Shukla, Jelena Srebric, Zoltan Toth, Mike Wallace, Christopher Wikle, Victor Yakovenko, Jim Yorke, Ning Zeng, and Aleksey Zimin.

We would like to thank many former students in our classes and/or research group members: Javier Amezcua, Eviatar Bach, Kriti Bhargava, George Britzolakis, Ming Cai, Chu-Chun Chang, Tse-Chun Chen, Matteo Corazza, Chris Danforth, Eli Dennis, Steve Greybush, David Groff, Matt Hoffman, Daisuke Hotta, Ji-Sun Kang, Maia Karpovich, Hong Li, Yan Li, Guo-Yuan Lien, Junjie Liu, Yun Liu, Erin Lynch, Adrienne Norwood, DJ Patil, Malaquias Peña, Steve Penny, Zhaoxia Pu, Juan Ruiz, Tamara Singleton, Travis Sluka, Qianqian Song, Luyu Sun, Shu-Chih Yang, and Takuma Yoshida. Many of these former research group members are currently professors at major research universities or lead scientists in meteorological agencies worldwide.

A complete list, even if possible, will inevitably become very long, so we would like to express our sincere appreciation to all colleagues whose names do not appear in the above lists due to space limitation.

Acknowledgments from the First Edition

I drafted about two-thirds of this book while teaching the subject for the first time at Oklahoma University, during the fall of 1998. Oklahoma University provided me with a supportive environment that made it possible to write the first draft. I made major revisions and finished the book while teaching the course again in the fall in 1999 through 2001 at the University of Maryland. The students that took the course at University of Maryland and Oklahoma University gave me essential feedback and helped me find many (hopefully most) of the errors in the drafts.

In addition, several people helped to substantially revise one or more of the manuscript chapters, and their suggestions and corrections have been invaluable. Norm Phillips read an early draft of Chapter 1 and made important historical comments. Anders Persson wrote the notes on the early history of numerical weather prediction, especially in Europe, reproduced in an appendix. Alfredo Ruiz Barradas reviewed Chapter 2. Will Sawyer reviewed and made major suggestions for improvements for Chapter 3. Hua-lu Pan influenced Chapter 4. Jim Geiger reviewed Chapter 5 and pointed out sections that were obscure. Jim Purser also reviewed this chapter and not only made very helpful suggestions but also provided an elegant demonstration of the equivalence of the 3D-Var and OI formulations. Discussions with Peter Lyster on this chapter were also very helpful. D. J. Patil suggested many improvements to Chapter 6, and Bill Martin pointed out the story by Ray Bradbury concerning the "butterfly effect." Joaquim Ballabrera substantially improved the appendix on model output postprocessing. Shu-Chih Yang and Matteo Corazza carefully reviewed the complete book, including the appendices, and suggested many clarifications and corrections.

I am grateful to Malaquias Peña, who wrote the abbreviations list and helped with many figures and corrected references. Dick Wobus created the beautiful six-day ensemble forecast figure shown on the cover. Seon Ki Park provided the linear tangent and adjoint code in Appendix B. The help and guidance of Matt Lloyd and Susan Francis of Cambridge University Press, the editing of the text by Maureen Storey, and the kind foreword by Norm Phillips are also very gratefully acknowledged.

I began to learn numerical weather prediction (NWP) in the late 1960s from professors at the University of Buenos Aires, especially Rolando Garcia and Ruben Norscini, and from the inspiring book of P. D. Thompson. At the Massachusetts Institute of Technology (MIT), my thesis advisor, Jule Charney, and the lectures of Norm Phillips and Ed Lorenz, influenced me more than I can describe. The NWP class notes of Akio Arakawa at University of California, Los Angeles and the National Center for

Atmospheric Research text on numerical methods by John Gary helped me teach the subject at MIT. Over the last 30 years, I have continued learning from numerous colleagues at other institutions where I had the privilege of working. They include the University of Montevideo, MIT, NASA's Goddard Space Flight Center, Oklahoma University, and the University of Maryland. However, my most important experience came from a decade I spent as Director of the Environmental Modeling Center at the National Centers for Environmental Prediction (NCEP), where my extremely dedicated colleagues and I learned together how to best transition from research ideas to operational improvements.

Finally, I would like to express my gratitude for the tremendous support, patience, and encouragement that my husband, Malise Dick, my son, Jorge Rivas, and my sisters, Patricia and Susana Kalnay, have given me and for the love for education that my parents instilled in me.

Variables

a	radius of the Earth
$\mathbf{A}$	analysis error covariance matrix
$\mathbf{B}$	background error covariance matrix
$\mathbf{C}$	covariance matrix
C_p, C_v	specific heat at constant pressure, constant volume
$\mathbf{d}$	innovation or observational increments vector
D	fluid depth
$E()$	expected value
f	Coriolis parameter
g	gravitational constant
$\mathbf{H}$	linear observation operator matrix
H	observational operator, scale height of the atmosphere
$\mathbf{I}$	identity matrix
J	cost function
$J\mathrm{M}$	maximum number of grid points j
$\mathbf{K}$	Kalman gain matrix
$L(t_0, t)$	resolvent or propagator of TLM
$\mathbf{M}$	TLM matrix
N	Brunt–Väisälä frequency
$\mathbf{P}$	projection matrix
p	pressure, probability, distribution function
q	mixing ratio of water vapor and dry air mass
$\mathbf{Q}$	forecast model error covariance
$\mathbf{r}$	position vector
$\mathbf{R}$	observations error covariance matrix
R	root mean square error, gas constant
R_d	Rossby radius of deformation
R_0	Rossby number
RE	relative error
T	temperature
TS	threat score
u, v	eastward and northward wind components
$\mathbf{W}$	weight matrix
W	vertical wind component, optimal weight

x,y	horizontal coordinates
δ_{ij}	Kronecker delta
ε_a	analysis error
ε_b	background error
η	absolute vorticity
Φ	geopotential height
φ	geopotential, latitude
λ	longitude
λ_i	global Lyapunov exponent
ρ_{ij}	element i,j of the correlation matrix C
σ	standard deviation
σ^2	variance
ψ	streamfunction
ω	vertical velocity in pressure coordinates, spectral frequency
ζ	relative vorticity

Abbreviations

3D-Var	Three-dimensional variational analysis
4D-Var	Four-dimensional variational analysis
4DDA	Four-dimensional data assimilation
AC	Anomaly correlation
ADI	Alternating direction implicit
AGCM	Atmospheric general circulation model
AI	Artificial intelligence
AMIP	Atmospheric Model Intercomparison Project (frequently refers to long model runs in which the observed SST is used instead of climatology)
AMSR2	Advanced Microwave Scanning Radiometer 2
AMV	Atmospheric motion vectors
AO	Arctic Oscillation
ARPS	Advanced Regional Prediction System
AVHRR	Advanced Very High Resolution Radiometer
AVN	NCEP's aviation (global) spectral model
BLUE	Best Linear Unbiased Estimation
BV	Bred vector
CAPS	Center for Analysis and Prediction of Storms
CFL	Courant–Friedrichs–Lewy
COAMPS	US Navy's Coupled Ocean/Atmosphere Mesoscale Prediction System
CONUS	Continental USA
CPC	Climate Prediction Center (NCEP)
CRCP	Cloud Resolving Convective Parameterization
CSE	Control Simulation Experiment
CSI	Critical success index (same as threat score)
CSRM	Cloud System Resolving Model
DL	Deep learning
DWD	German Weather Service
EAKF	Ensemble adjustment Kalman filter
ECMWF	European Centre for Medium-Range Weather Forecasts
EDA	Ensemble data assimilation
EFSO	Ensemble Forecast Sensitivity to Observations

EKF	Extended Kalman filter
EMC	Environmental Modeling Center
ENIAC	Electronic numerical integrator and computer
ENSO	El Niño–Southern Oscillation
ETKF	Ensemble transform Kalman filter
EnKF	Ensemble Kalman filter
EnSRF	Ensemble square root filter
FASTEX	Fronts and Storm Track Experiment
FDE	Finite difference equation
FDR	Frequency dispersion relationship
FFSL	Flux-form semi-Lagrangian scheme
FSO	Forecast sensitivity to observations
FV3	Finite-volume, cubed-sphere
GCM	General circulation model
GEFS	Global Ensemble Forecast System
GFDL	Geophysical Fluid Dynamics Laboratory
GFS	Global Forecast System
GLE	Global Lyapunov exponents
GMI	Global Precipitation Measurement (GPM) Microwave Imager
GPS	Global Positioning System
GSI	Gridpoint Statistical Interpolation
hPa	Hectopascals (also known as millibars)
HFIP	Hurricane Forecast Improvement Program
HFSO	Hybrid forecast sensitivity to observations
HPC	Hydrometeorological Prediction Center (NCEP)
IAM	Integrated Assessment Model
ICTP	International Centre for Theoretical Physics
IFS	Integrated Forecasting System
IPCC	Intergovernmental Panel on Climate Change
JMA	Japan Meteorological Agency
JNWPU	Joint Numerical Weather Prediction Unit
KF	Kalman filter
LEKF	Local ensemble Kalman filter
LETKF	Local ensemble transform Kalman filter
LFM	Limited fine mesh
LLV	Local Lyapunov vectors
LPF	Local particle filter
MeteoFrance	National Meteorological Service for France
MCC	Mesoscale Compressible Community (model)
MJO	Madden–Julian Oscillation
ML	Machine learning
MM5	Fifth-Generation Penn State/NCAR Mesoscale Model
MMF	Multiscale modeling framework

MODIS	Moderate Resolution Imaging Spectroradiometer	
MPAS	Model for Prediction Across Scales	
MOS	Model Output Statistics	
NAO	North Atlantic Oscillation	
NASA	National Aeronautics and Space Administration	
NCAR	National Center for Atmospheric Research	
NCEP	National Centers for Environmental Prediction (US National Weather Service)	
NCI	Nonlinear computational instability	
NGGPS	Next Generation Global Prediction System	
NGM	Nested Grid Model	
NLNMI	Nonlinear normal mode initialization	
NMC	National Meteorological Center	
NOAA	National Oceanic and Atmospheric Administration	
NORPEX	North Pacific Experiment	
NWP	Numerical weather prediction	
NWS	National Weather Service	
OI	Optimal interpolation	
OSSE	Observing System Simulation Experiments	
PBL	Planetary boundary layer	
PDE	Partial differential equation	
PDO	Pacific Decadal Oscillation	
PF	Particle filter	
PIRCS	Project to Intercompare Regional Climate Systems	
PQC	Proactive quality control	
PQPF	Probabilistic Quantitative Precipitation Forecast	
PSAS	Physical space analysis scheme	
PVE	Potential vorticity equation	
RAFS	Regional Analysis and Forecasting System	
RAOB	Rawinsonde observation	
RDAS	Regional Data Assimilation System	
RF	Recursive filter	
RIKEN	RIkagaku KENkyūjyo (Institute of Physical and Chemical Research in Japan)	
RIP	Running in Place	
RSM	NCEP's Regional Spectral Model	
RUC	NCEP's Rapid Update Cycle	
SAC	Standardized anomaly correction	
SCM	Successive correction method	
SLAF	Scaled Lagged Average Forecasting	
SMAP	Soil Moisture Active Passive	
SOR	Successive overrelaxation	
SPEEDY	Simplified parameterizations, primitive-equation dynamics	
SSI	Spectral Statistical Interpolation	

SST	Sea surface temperature
SV	Singular vector
SWE	Shallow water equation
TAMC	Tangent and Adjoint Models Compiler
TOGA	Tropical Ocean, Global Atmosphere
TOVS	TIROS-N Operational Vertical Sounder
TS	Threat score
UKMO	United Kingdom Meteorological Office
UTC	Universal time or Greenwich time, e.g., 1200 UTC. Frequently abbreviated as 1200Z
WMO	World Meteorological Organization
WRF	Weather Research and Forecasting model