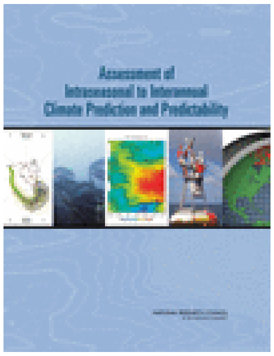

Assessment of Intraseasonal to Interannual Climate Prediction and Predictability

Committee on Assessment of Intraseasonal to Interannual Climate Prediction and Predictability; National Research Council

ISBN: 0-309-15184-8, 192 pages, 8 1/2 x 11, (2010)

Assessment of Intraseasonal to Interannual Climate Prediction and Predictability

Committee on Assessment of Intraseasonal to Interannual Climate Prediction and Predictability

Board on Atmospheric Sciences and Climate

Division on Earth and Life Studies

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS • 500 Fifth Street, N.W. • Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by the National Oceanic and Atmospheric Administration under contract number DG133R-08-CQ-0062, TO# 2. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsoring agency or any of its sub agencies.

International Standard Book Number-13: 978-0-309-15183-2
International Standard Book Number-10: 0-309-15183-X

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

COMMITTEE ON ASSESSMENT OF INTRASEASONAL TO INTERANNUAL CLIMATE PREDICTION AND PREDICTABILITY

ROBERT A. WELLER (*Chair*), Woods Hole Oceanographic Institution, Massachusetts
JEFFREY L. ANDERSON, National Center for Atmospheric Research, Colorado
ALBERTO ARRIBAS, Met Office Hadley Centre, United Kingdom
ROBERT E. DICKINSON, University of Texas, Austin
LISA GODDARD, Columbia University, New York, New York
EUGENIA KALNAY, University of Maryland, College Park
BENJAMIN KIRTMAN, University of Miami, Florida
RANDAL D. KOSTER, National Aeronautics and Space Administration, Greenbelt, Maryland
MICHAEL B. RICHMAN, University of Oklahoma, Norman
R. SARAVANAN, Texas A&M University, College Station
DUANE WALISER, Jet Propulsion Laboratory, California Institute of Technology, Pasadena
BIN WANG, University of Hawaii, Honolulu

NRC Staff:

MARTHA MCCONNELL, Study Director
JOSEPH CASOLA, Postdoctoral Fellow
LAUREN BROWN, Research Associate
SHELLY FREELAND, Senior Program Assistant
DAVID REIDMILLER, Christine Mirzayan Science and Technology Fellow

Acknowledgments

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the NRC's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their participation in their review of this report:

Brian Hoskins, Imperial College London, UK
Richard Kleeman, New York University
Robert A. Knox, University of California, San Diego
Arthur Lee, Chevron Corporation, San Ramon, CA 94583
Ruby Leung, Pacific Northwest National Laboratory, Richland, WA
Robert E. Livezey, National Oceanic and Atmospheric Administration (retired), Silver Spring, MD
Andrew M. Moore, University of California, Santa Cruz
Sumant Nigam, University of Maryland, College Park
Tim Palmer, European Centre for Medium-Range Weather Forecasts, Reading, UK
Matthew C. Wheeler, Centre for Australian Weather and Climate Research, Melbourne

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by **Roger B. Lukas** (University of Hawaii). Appointed by the National Research Council, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Contents

Summary 1

1 Introduction 11
Scope and Purpose of this Report 11
ISI Predictability: The Example Of El Niño-Southern Oscillation 16
Organization of This Report 20

2 Climate Prediction 21
The Concept of Predictability 21
Sources of Predictability 26
Methodologies Used To Quantitatively Estimate Prediction Skill 42
Challenges to Improving Prediction Skill 53

3 Building Blocks of Intraseasonal to Interannual Forecasting 54
Historical Perspective for Intraseasonal to Interannual Forecasting 54
Observations 56
Statistical Models 66
Dynamical Models 73
Data Assimilation 79
Use of Forecasts 87
Example of an ISI Forecast System 96
Potential Improvements to ISI Forecast Systems 98

4 Case Studies 101
El Niño-Southern Oscillation (ENSO) 101
Madden-Julian Oscillation (MJO) 108
Soil Moisture 116

5 Best Practices 124
Public Archives 124
Metrics 125
More Useful Forecast Products 126
Accelerated Synergy with the Research Community 127

6 Recommendations and Remarks on Implementation 129
Recommendations 129
Remarks on Implementation 135

Closing Remarks 141

References 143

Appendix A Background Information on Statistical Techniques 170

Appendix B Committee Members' Biographical Information 176

Summary

Many important decisions regarding water management, agriculture, and energy are made on weekly, monthly, seasonal, and annual timescales. These decisions can benefit from high quality, reliable predictions. Yet making useful predictions about the climate system on these timescales is a challenge. The purpose of this report is to examine current capabilities for making such intraseasonal to interannual (ISI)[1] predictions for the climate system, to analyze how past improvements in these capabilities have been achieved, and to recommend opportunities for future improvement.

ISI climate predictions occupy an intermediate timescale between traditional weather forecasts, which are useful for the coming days, and global climate simulations associated with climate change, which relate to changes occurring over decades and centuries (see Box S.1). Predicting the climate at this intermediate timescale poses unique challenges since it involves many processes that operate among the atmosphere, ocean, and land surface. These processes are often incompletely understood and difficult to measure with available observational platforms. There are limits to the spatial and temporal resolution of our observations, and a "perfect" or complete observation of the climate system will never be achieved. Numerical models of the climate system demonstrate large sensitivity to initial conditions that cause errors or uncertainties to grow with increases in forecasting lead times. Moreover, models are known to have errors in formulation and are limited in resolution, which can also lead to forecast errors.

BOX S.1
MODELS FOR PREDICTING THE WEATHER VERSUS MODELS FOR PREDICTING CLIMATE

To understand climate prediction, it is useful to distinguish it from weather forecasting, which is a familiar concept to many from everyday experience. Weather models derive their prediction skill from accurate knowledge of initial conditions in the atmosphere. They produce deterministic forecasts, often with high enough skill that they can be used for simple everyday decisions, such as choosing proper clothing, or to warn us of short-term weather threats, such as lightning, severe winds, or intense precipitation. Climate models, on the other hand, derive much of their prediction skill from knowledge of the initial conditions in slowly evolving components of the climate system, such as the ocean or the cryosphere. The predictions produced by climate models are inherently probabilistic and have considerably lower skill than 1-2 day weather forecasts. They are usually of little use in planning everyday activities. However, climate predictions are very useful to government agencies, non-governmental organizations, and private companies for policy and longer-term planning purposes. Examples of applications include drought mitigation, malaria prevention, farming, pricing of insurance, and managing energy resources.

[1] Intraseasonal to interannual is defined as extending from roughly two weeks to several years.

The committee was requested to assess how researchers and forecasters have addressed these challenges and to recommend avenues for further progress. Specifically, the committee was tasked to review the current understanding of ISI predictability, describe how past improvements in forecast systems have occurred, identify gaps in our current understanding of ISI predictability, assess the performance of current ISI forecast systems, and recommend strategies and best practices for future improvements to ISI forecasts and our overall understanding of ISI predictability.

The committee begins from the premise that the ability to predict the climate accurately at ISI timescales stems from our knowledge of "sources of predictability," the variables or processes operating within and among the atmosphere, ocean, and land that affect the state of the climate on ISI timescales. The sources of predictability are measured, represented, and simulated by ISI forecast systems through an assemblage of "building blocks," namely observational systems, statistical and dynamical models, and data assimilation schemes. This report illustrates the relationship between the sources of predictability and the building blocks of ISI forecast systems. In addition, this report discusses techniques and protocols for the verification and dissemination of ISI forecasts by operational forecasting centers, highlighting the impact that these practices can have on forecast quality and opportunities for improvement. This report concludes with recommendations for improving ISI forecast systems, targeting both operational forecasting centers and the broader research community.

SOURCES OF PREDICTABILITY

This report explores three interrelated categories of predictability sources that exist within the climate system. The first of these sources of predictability is related to particular variables that exhibit inertia or memory, such as ocean heat content, in which anomalous conditions can take relatively long periods of time (days to years) to decay. The second type of source of predictability is related to patterns of variability or feedbacks. Coupling among processes in the climate system can give rise to characteristic patterns that explain some portion of the spatial and temporal variance exhibited by key climate variables, such as temperature or precipitation. An example is the El Niño-Southern Oscillation (ENSO), where anomalous conditions in the tropical Pacific Ocean influence seasonal climate in the mid-latitudes around the globe. The third source of predictability is due to external forcing. Volcanic eruptions, changes in solar activity, and the accumulation of greenhouse gases in the atmosphere are all examples of external forcing. These events or processes can affect the climate on ISI timescales in predictable ways that can be exploited for making climate predictions.

It is important to note that the processes that affect the climate on ISI timescales can themselves operate on a variety of timescales. This is depicted in Figure S.1, which provides many examples of processes that affect the climate at ISI timescales and can serve as sources of predictability. These sources can be related to phenomena that occur in, on, or among the ocean, atmosphere, and land surface components of the climate system.

The ability of ISI forecast systems to represent these sources of predictability accurately partially determines the quality of the predictions. Past improvements in prediction quality have accompanied increased understanding of the sources of predictability and incorporation of this understanding into forecast systems. Future advances in the quality of ISI predictions are closely

FIGURE S.1 Processes that act as sources of ISI climate predictability extend over a wide range of timescales and involve interactions among the atmosphere, ocean, and land. CCEW: convectively coupled equatorial waves; TIW: tropical instability wave; MJO/MISV: Madden-Julian Oscillation/Monsoon intraseasonal variability; NAM: Northern Hemisphere annular mode; SAM: Southern Hemisphere annular mode; AO: Arctic oscillation; NAO: North Atlantic oscillation; QBO: quasi-biennial oscillation, IOD/ZM: Indian Ocean dipole/zonal mode; AMOC: Atlantic meridional overturning circulation. For the *y*-axis, "A" indicates "atmosphere;" "L" indicates "land;" "I" indicates "ice;" and, "O" indicates "ocean."

tied to exploiting new sources of predictability or improving the representation of known sources of predictability in current forecast systems.

THE BUILDING BLOCKS OF AN ISI FORECAST SYSTEM

ISI forecasting systems are composed of several "building blocks:" observations, statistical and dynamical models, and data assimilations systems. Observations are required to measure the state of the variables that contain memory in the climate system; to monitor the evolution of key processes that operate within and among the atmosphere, ocean, and land; and to identify the magnitude of external forcing. These observations can be utilized by a data assimilation system to provide the current state (an "initial state") of the climate system, and that information can be utilized by statistical and/or dynamical models to make predictions. Observations are required to validate models, verify forecasts, and expand understanding of underlying climate processes. These activities can then feedback into identifying model deficiencies and improving model formulations.

The performance of ISI forecast systems can be enhanced through improvements to these building blocks, which are intimately connected. For example, if new observations are made available, then it is likely that new components of statistical and dynamical models or data assimilation algorithms need to be developed in order to incorporate these observations into forecasts. Conversely, a comparison of existing models to novel observations or models may identify underlying deficiencies in our understanding of important climate processes, motivating further model development. Thus, improvements to ISI forecast systems stem from synergistic improvements across each of the building blocks, where upgrades to one component enhance, or are enhanced by, upgrades to the other components.

Based on its examination of the literature, the committee concludes that incremental increases in ISI forecasting quality are to be expected as the building blocks of ISI forecast systems are improved and upgraded. The committee also concludes that there are no "silver bullets;" there is no single action that will lead to a revolutionary leap forward in ISI predictions. As past improvements to ISI predictions and weather predictions have shown, progress forward can be achieved by a concerted effort to address the shortcomings of the various building blocks of forecast systems.

CASE STUDIES

Much can be learned about ISI predictions by exploring case studies for ENSO, the Madden-Julian Oscillation (MJO), and soil moisture. Such case studies demonstrate the role that observations, models, data assimilation techniques, and verification protocols play in making ISI forecasts. For ENSO, the perspective is somewhat historical, as many previous advances in ISI forecasting have come from an improved observational capacity that accompanied expanded understanding of physical processes and model development. For the MJO and soil moisture, the perspective is more forward-looking. There remains the potential to exploit the MJO and soil moisture to improve ISI forecasts.

ISI FORECASTING INSTITUTIONS

ISI forecasts, along with the building blocks for forecasting, are developed, produced, and processed by a variety of institutions around the world, including the National Centers for Environmental Prediction (NCEP), the European Centre for Medium Range Forecasting (ECMWF), the International Research Institute for Climate and Society (IRI), and many universities and research laboratories. The committee draws a distinction between "operational centers" and "research institutions." The former issue forecasts in real time on a fixed, regular schedule and are associated with a national meteorological and hydrological service; the latter are more research-oriented and are often associated with universities and academic scientists. Programs that can foster collaboration between these two types of institutions have been successful in advancing ISI forecast quality, and several of the recommendations aim to encourage further collaboration and enhance existing mechanisms for cooperation.

USE OF FORECASTS

Quality forecasts can contribute to societally-relevant decisions. However, a variety of metrics can be used to determine overall forecast quality; a single metric rarely encapsulates all the information regarding forecast quality. Likewise, different decision makers will rely on a diverse set of variables (e.g. agricultural planners may be most interested in precipitation, while wind power operators may be most interested in wind forecasts), and may have varying demands for the forecast quality associated with these variables.

The report describes the procedure of making ISI forecasts, outlining how information from various sources, both objective (e.g., predictions from dynamical or statistical models) and subjective (e.g., expert opinions of forecasters), are combined. The wide range of forecast formats and accessible forecast documentation makes it challenging to compare the performance of forecast systems or detect how changes in forecast inputs and practices affect overall forecast performance. Similarly, the variables and formats of forecasts may not correspond to the needs of decision makers, acting as a barrier to the use of forecasts, regardless of their quality.

RECOMMENDATIONS

The committee identified three general categories of actions to advance ISI predictions: Best Practices, Improvements to the Building Blocks of ISI Forecast Systems, and Research for Sources of Predictability. Best Practices are largely focused on the activities of the operational forecast centers and aim to improve the delivery and dissemination of forecast information for both decision makers and researchers. Although adopting Best Practices may require some additional resources on the part of the operational centers, the barriers to adoption are relatively minimal; many of the recommendations involve modification to current protocols or expansion of current programs rather than a novel set of initiatives.

The Improvements to the Building Blocks of ISI Forecast Systems pertain to both the operational and research communities and focus on the continued development of observations, statistical and dynamical models, and data assimilation systems. The benefits associated with these recommendations have a longer time horizon than those associated with Best Practices and may require several years to achieve.

Research for Sources of Predictability, the final category of the recommendations, is aimed primarily toward the research community. These recommendations constitute specific goals for current and future academic exploration of ISI processes. Although the committee agrees that these goals should be pursued with the intent that they contribute to operational ISI forecasts, the initial efforts to investigate these unexploited sources of predictability fall largely on research scientists.

Best Practices

(1) The synergy between operational ISI forecasting centers and the research community should be enhanced.

Establishing connections between the operational and research communities is critical to further progress in ISI forecasting. Fostering dialog and exchange between these communities permits identification of common problems and expands the sets of tools available for finding solutions. Specific activities include holding workshops focused on specific areas of model and forecast development, encouraging scientists that work at operational centers to participate in scientific meetings focused on modeling and the use of observations, granting of short term positions in operational centers to academic researchers, and improving the speed and manner by which new data sets generated by operational centers are shared with the broader research community.

(2) Operational ISI forecasting centers should establish public archives of all data used in forecasts including observations, model code, hindcasts, analyses, forecasts, re-analyses, re-forecasts, verifications, and official forecast outlooks.

Archives of the inputs to, outputs of, and tools used in ISI forecasts are needed in order to quantify and identify sources of forecast error, provide the baseline for forecast assessment and model fidelity, develop metrics and diagnostics for model assessment, calibrate model predictions, and document model and forecast improvements.

Archives can serve as an important mechanism for making ISI forecasts more readily useable for management decisions and societally relevant research. Although it is not possible for operational centers to foresee or address all possible needs of the forecast users, archives will permit the development of tailored forecast products for decision systems and risk management by users and researchers. Once engaged, these groups can also provide valuable feedback for further improvements in ISI forecasting.

(3) Operational ISI forecasting centers should broaden and make available the collection of metrics used to assess forecast quality.

No perfect metric exists that conveys all the information about a forecast. Multiple metrics should be used when assessing forecasts, including graphical techniques; metrics that assess the quality of probabilistic information and that from multi-model ensembles. Some of these metrics should include information on the distribution of forecast skill in space and time.

(4) The subjective components of operational ISI forecasts should be minimized.

Recent research suggests that the subjective component of many present-day forecasts can reduce forecast quality. The subjective component generally comes from qualitative discussion and interpretation by forecasters regarding the state of the climate system and forecasting tools. The subjective component also limits reproducibility, restricting retrospective comparison of forecast systems.

Improvements to the Building Blocks of Forecast Systems

(5) Statistical techniques, especially nonlinear methods, should be pursued in order to better characterize processes that contribute to ISI forecasts.

Statistical methods provide important tools for comparing model predictions and observations and subsequently identifying model deficiencies. Historically, linear statistical analyses of observational data have provided an awareness of many patterns of variability that have been useful for making ISI forecasts. Recent research demonstrates that nonlinear methods can yield statistically significant increases in prediction skill on ISI time scales when compared to traditional linear techniques. However, these techniques have not been incorporated operationally. Therefore, nonlinear alternatives should be explored to augment our current knowledge.

(6) Systematic errors in dynamical models should be identified.

Current state-of-the-art ISI prediction models have relatively large errors in their representation of the mean climate, the climate variability, and their interaction. These errors reduce prediction quality. Some classic examples include: (1) the so-called double intertropical convergence zone (ITCZ) problem, (2) the excessively strong equatorial cold tongue, (3) weak or incoherent intraseasonal variability, (4) failure to represent the multi-scale organization of tropical convection, and (5) poorly represented cloud processes, particularly low level stratus. These errors have both regional and global impacts and could be indicative of errors in the model formulations that are limiting prediction quality.

Sustained observations are needed to quantify model errors. Examples of sustained observations include those related to describing the properties or fluxes among the atmosphere, ocean, and land surface (e.g., boundary layer humidity, exchange of heat between the atmosphere and ocean).

(7) To reduce errors produced by dynamical models, the representation of physical processes should be improved.

The physical processes underlying ISI variability are often poorly understood. Process studies that are closely tied to operational ISI model improvement should be carried out with the goal of transferring improvements into operational ISI forecasts. Targeted, novel observations will likely play a role in these types of studies. Studies could focus on specific components of the climate system (e.g., sea ice, aerosols, snow cover), specific processes and variability (e.g., triggering the onset of an MJO), and the interactions among components of the climate system (e.g., air-land coupling strength, stratosphere-troposphere interactions).

The CLIVAR climate process teams (CPTs), which exist currently, provide a mechanism for accomplishing this. The CPTs focus modelers and process scientists on poorly-represented or unrepresented physical processes in models.

Work should be carried out to move toward more complete inclusion of climate processes in the models. Computing capabilities should be improved to permit the explicit simulation of subgrid-scale processes and remove as much reliance on parameterization as possible. The role of increasing model resolution in improving ISI forecasts should continue to be explored.

(8) Statistical and dynamical models should continue to be used in a complementary fashion by operational ISI forecasting centers.

Using multiple prediction tools leads to improved and more complete ISI forecasts. Statistical tools should continue to be developed and employed in an effort to improve dynamical model output. Examples of statistical techniques include stochastic physics, interactive ensembles, empirical corrections or empirically-based parameterizations and process models.

The use of statistical and dynamical downscaling methods is another application that should be explored to address the information mismatch between the coarse spatial resolution of operational climate forecasts and the fine resolution needs of some end users.

(9) Multi-model ensemble (MME) forecast strategies should be pursued, but standards and metrics for model selection should be developed.

Continued work is necessary to develop techniques of optimally selecting and weighting ensemble members. Experimentation with MME should not compete with model improvement, but rather, should contribute to the process of identifying areas for model improvement.

(10) To enable assimilation of all available observations of the coupled climate system, operational centers should implement state-of-the-art 4-D Var, Ensemble Kalman Filters, or hybrids of these in their data assimilation systems.

The most advanced assimilation systems are typically not used in operational settings or are limited to atmospheric observations only. Assimilation systems should be upgraded.

There are many available observations that are not currently being utilized in data assimilation schemes that could contribute to the initialization of dynamical models. More observations should be assimilated into operational ISI forecast systems. The expansion of the variables assimilated in weather forecasts has contributed to improvements in forecast quality. Analogous gains could be made for ISI forecasting. Priority should be given to expanding operational data assimilation to ocean observations such as sea surface heights.

Research for Sources of Predictability

(11) Many sources of predictability remain to be fully exploited by ISI forecast systems. To better understand key processes that are likely to contribute to improved ISI predictions, the committee recommends that the scientific community pursue the following six areas as research goals.

Madden-Julian Oscillation (MJO)

The path forward on understanding and forecasting the MJO should include focused process studies, model improvement, and close collaboration between research and operational communities. It will be necessary to develop and implement standardized diagnostics and metrics to gauge model improvements and track improvements in forecast quality. MJO influences on other important components of the climate system, such as ENSO, monsoon onsets and breaks, and tropical cyclone genesis should continue to be explored and exploited for additional predictability.

Stratosphere-Troposphere Interactions

Relatively long-lived (up to two months) atmospheric anomalies can arise from stratospheric disturbances. In sensitive areas such as Europe in winter, experiments suggest that the influence of stratospheric variability on land surface temperatures can exceed the local effect of sea surface temperature. Additionally, while our weather and climate models do not often resolve or represent the stratospheric Quasi-Biennial Oscillation very well, it is one of the more predictable features in the atmosphere, and it has been found to exhibit a signature in ISI surface climate.

Ocean-atmosphere coupling

Due to the very large heat capacity of sea water, anomalous sea surface temperatures and upper ocean heat content can have significant impacts on the atmosphere above. The impacts of the anomalies associated with ENSO are well-known. However, further research is needed to examine the role of extratropical atmosphere-ocean coupling, to investigate the need to more realistically represent ocean-atmosphere coupling over a wide range of spatial scales (including down to the scales of the sharp SST gradients associated with fronts), and to better observe and more realistically represent air-sea fluxes in models.

Land-atmosphere feedbacks

The realistic initialization of soil moisture in models can increase the accuracy of precipitation and temperature predictions at intraseasonal timescales. The realistic initialization of snow amount may also yield better quality predictions, though this connection is relatively unexplored. To maximize the impact of land feedbacks on prediction quality, the mechanisms underlying the land-atmosphere coupling (e.g., evaporation, boundary layer dynamics, convection) need to be better understood and better represented in forecast systems.

High impact events affecting atmospheric composition

Research efforts should study the consequences on the climate system at ISI timescales of unusual but high impact events, such as volcanic eruptions, limited nuclear exchanges, or space impacts that cause a sudden, drastic change to the atmospheric burden of aerosols and trace gases. ISI forecasts from operational centers following these types of events could have significant societal ramifications.

Non-stationarity

Statistical and dynamical models for ISI forecasting should be improved to better capture the predictability associated with long-term trends in atmospheric composition (e.g., increases in greenhouse gas concentrations) and land cover change. Current statistical techniques and dynamical models do not adequately deal with this non-stationarity. Improved statistical techniques should be developed for exploiting the predictability associated with such non-stationary behavior. The use of dynamical models that include a more comprehensive treatment of radiative processes such as aerosol effects, and also incorporate trends in land use, could help improve the quality of dynamical ISI forecasts on longer timescales.

CLOSING THOUGHTS

For the short term, operational ISI forecast centers can increase the value of forecasts to decision makers and researchers by modifying procedures for archiving and disseminating forecast information and enhancing collaborations with the external research community. Over the next several years and coming decades, improvements to observational capabilities, statistical and dynamical models, and data assimilations systems should permit ISI forecast systems to better represent the variables and processes that serve as sources of predictability. Research to characterize sources of predictability that are poorly understood should also offer opportunities to improve ISI predictions as well as our understanding of important underlying climate processes.

1

Introduction

SCOPE AND PURPOSE OF THIS REPORT

This study responds to a request by the National Oceanic and Atmospheric Administration (NOAA) to the National Academy of Sciences to review the current understanding of climate predictability on ISI timescales, including past improvements in our understanding of predictability; to identify remaining gaps in our understanding of predictability at these timescales; to assess the performance of current prediction systems; and to recommend strategies and best practices for improving estimates of predictability and prediction skill (see Box 1.1).

In preparing the report, the committee has drawn on published literature as well as from presentations from a variety of research scientists and forecasting experts representing both U.S. and international institutions. Due to the experiences of the committee members and the source of the report request, the report tends to focus most heavily on climate predictions for the United States and North America. In contrast to the U.S.-focus regarding predictions, the recommendations regarding forecasting procedures and protocols (Best Practices) have been crafted based upon forecasting experiences from the United States and abroad. These have been drawn from and could be applicable to many national and international institutions. Likewise, many of the physical processes discussed (e.g., ENSO, MJO, NAO) have significant impacts on non-U.S. climate phenomena, such as the Indian monsoon.

The committee feels that this report will inform and guide decisions regarding future opportunities in climate research and operational forecasting. Significant challenges remain in formulating and disseminating accurate and useful forecasts at the intraseasonal and interannual timescales. Significant opportunities exist for the research community to expand its knowledge of climate processes, especially with respect to the coupling among components of the climate system, and improving observational systems, statistical and dynamical models, and data assimilation techniques. Likewise, for the operational community, opportunities exist to verify, catalog, and share forecasts in a more systematic manner. Overall, better communication between the research and operational communities is required for all of these improvements to be achieved.

Introduction to the Climate System

The sun serves as the primary energy source for the climate system, and day-to-day and season-to-season changes in the solar radiation received by the Earth lead to some well-recognized changes in the climate system. For example, on a clear, calm day, sea surface temperature (SST) in the tropics and mid-latitudes can warm as much as 3°C during the day. At the same sites, we observe warming and an increase in SST through the spring and into the summer followed by cooling and a decrease in SST through the fall and winter. These changes in SST occur in concert with seasonal changes in surface winds. These types of day-to-day and season-to-season variability, caused by strong, regular, and periodic external forcing from the sun can be accurately predicted.

But beyond these daily and seasonal cycles, the dynamics of the climate system are more complex and incompletely understood, challenging our efforts to make predictions. For example, to answer a question like "Will the upcoming winter be colder or wetter than usual?" requires an understanding of climate variability on the timescales of weeks, months, and years. This variability stems from the atmosphere, the ocean, the land, and the coupling between them. How these components of the climate system interact and affect one another can be understood by examining how they exchange heat, moisture, and momentum. For example, the ocean absorbs heat from the sun and can also transport that heat and release it elsewhere on the earth's surface. At mid- and high-latitudes, cooling and evaporation make surface water denser and, through convection, force surface water into the ocean's interior. Both the density differences in the ocean and the action of the wind on the sea surface drive a global, three-dimensional circulation in the ocean that results in spatial and temporal variability in SST. Likewise, solar heating and turbulent heat and moisture fluxes at the ocean and land surfaces drive atmospheric circulations on a wide range of scales from global to local. Moist, warm parcels of air near the surface become buoyant, and this convection can communicate the influence of the surface broadly through the atmosphere and, in turn, to remote surface locations. In contrast, cooling or evaporation within the lower atmosphere stabilizes the atmospheric boundary layer locally and limits the ability of the surface to force the atmosphere elsewhere.

The ability of the atmosphere, ocean, and land to interact and affect one another occurs over a broad range of spatial scales and timescales. These interactions give rise to complex, often nonlinear, dynamics making it difficult to understand and predict the climate variability that we observe. While much progress has been made extending weather forecast skill to a week or more, the ability to make predictions on timescales longer than two weeks is still limited. At shorter timescales, most of the important dynamics reside within the atmosphere. But for longer timescales, the storage of heat and moisture by the ocean and the land becomes more important. Unfortunately, we have less information about the ocean and the land than we have about the atmosphere, and we often lack a full understanding of the interactions among the three.

Committee Approach to Predictability

Historically, deterministic "predictability" of chaotic systems like day-to-day weather processes has referred to how relatively small errors in the initial conditions lead to relatively large forecast errors some time later—typically 10–14 days. Although developed in the context of weather prediction, this concept of deterministic predictability has also been applied to predictions of the entire climate system, including those on ISI timescales. However, over time, the term "predictability" has been used in confusing ways in the atmospheric and oceanic literature. In this report, the term "predictability" is used *qualitatively* to describe the extent to which the representation of a physical process can contribute to and perhaps even improve prediction quality. There are two important aspects of the committee's approach to the concept of predictability:

- It is not possible to quantify a true *limit of predictability* for the climate system.
- Quantitative statements can be made regarding the *lower bounds* of predictability, as derived by the performance of existing forecast systems. If a forecast system shows quantitative skill according to some metric, then at least that much predictability must exist in nature.

The approach that the committee has pursued impacts its ability to fulfill its requested tasks. Underlying several parts of the Statement of Task is an assumption that nature contains inherent predictability limits that can be accurately and quantitatively estimated through the analysis of observations and/or model results. In particular, Task 4 (see Box 1.1) asks the committee to:

> Assess the performance of current prediction systems in relation to the estimated predictability of the climate system on intraseasonal to interannual timescales, and recommend strategies (e.g.,observations, model improvements, and research priorities) to narrow gaps that exist between current predictive capabilities and estimated limits of predictability.

The committee finds that presently observational estimates of predictability are severely limited—the observational record is too short and the estimates require assumptions about the observational data (e.g. stationarity) that are difficult to satisfy. Model-based estimates of the intrinsic predictability[2] can also be made but are severely limited by the fidelity of the model. For example, model predictability estimates of the ENSO cycle could in principle span the gamut from zero predictability (modeling the cycle as a white noise process) to perfect predictability (modeling it as a sine wave). Of course, modelers use much more physically-based representations of ENSO; nevertheless, the predictability a model produces is unequivocally a function of the underlying model assumptions—the discretization of flow equations, the parameterizations of physical processes, and so on. Model-based ENSO predictability estimates vary widely among models, and for this and any other such process a higher estimate of predictability is not intrinsically a more accurate one.[3]

The committee finds that model-based estimates of the intrinsic limit of predictability are useful in a qualitative sense. While the studies themselves may very well be quantitative in implementation and analysis, they are best used to identify physical processes that impact the model-based estimate and therefore provide qualitative guidance in how to attack the forecast improvement problem in that model. (A simple example: if Model A shows no intrinsic predictability for a variable in a region where Model B shows some real forecast skill for that variable, then process formulations underlying that variable in Model A are deficient and could be a focus of improvement.) In fact, the committee recommendations are specifically designed to identify the infrastructure needs (i.e., observations, models, best practices) that will accelerate

[2] Intrinsic predictability is the extent to which the prediction is possible if an optimum procedure is used (see “The Concept of Predictability” in Chapter 2; Lorenz, 2006).
[3] These considerations and conclusions are not limited to ENSO but could also be said of the MJO and other sources of predictability.

the process of transitioning the qualitative guidance into quantitative forecast improvements. This process will necessarily involve rigorous forecast verification.

Despite the utility of model-specific estimates of the limits of intrinsic predictability for individual model development, the committee finds that *the only quantitative statements that can be made regarding predictability in nature involve its lower bounds, as provided by verifying forecasts from existing prediction systems.* In other words, if a forecast system shows true quantitative skill at some level, then at least that much predictability must exist in nature. This sentiment underlies much of the analysis in the report, and can be illustrated with an example. Suppose a statistical prediction of a measure of the strength of ENSO such as the Nino 3.4 index (the departure of the monthly mean SST inside a box bounded by 120°W–170°W and 5°S–5°N from its long-term mean) is of higher quality than a dynamical prediction. It could then be concluded that additional forecast quality can be obtained with a more accurate dynamical method. If estimates for the upper bound of predictability in nature could be derived, they would be uniquely valuable since they would indicate how much quality may yet be derived through future improvements in forecasting systems. In other words, such estimates could indicate how much potential quality is waiting to be tapped. Unfortunately, such estimates are inaccessible. The true limits of predictability cannot be quantified with any certainty because there is no way of estimating predictability without models or, in the case of observational data, *ad hoc* assumptions.

Despite the inability to unambiguously quantify the intrinsic or upper "limit of predictability," the committee was able to assess the performance of forecast systems. The quantitative assessment of forecast quality is a useful lower bound on predictability. It is clear that the skill of models has improved over time (Fig 1.1), at least with respect to the types of ENSO-based metrics that are usually discussed in the literature and has recently been evident with respect to advances in MJO forecasting as well (see "Dynamical Models" section in Chapter 3 and the MJO and ENSO case studies in Chapter 4 for a more specific discussion). With regard to the current generation of forecast systems, attempts to perform a rigorous evaluation of forecast quality have been made using available archives and multi-model ensemble systems (e.g., Climate-system Historical Forecast Project (CHFP), ENSEMBLES). However, these initiatives are relatively recent. The multitude of available forecast formats and metrics[4], and the lack of openly available data and information regarding past forecasts and verifications can make it difficult to compare across, or even conduct, such studies. The Best Practice recommendations, especially with respect to archiving forecast information and metrics, have been designed to help facilitate the establishment of a framework for comparing and evaluating estimates of prediction quality (i.e. lower bounds on predictability) derived from forecast models.

[4] Although the World Meteorological Organization has a recommended set of metrics for forecast verification, these have not been applied consistently by modeling and forecast centers. In addition, the WMO estimates of forecast quality may be relevant to the climate prediction community, but may not relate directly to the types of information a decision-maker might need, such as the occurrence of anomalous temperature or precipitation at a more local or regional level, or the occurrence of an extreme event (e.g. heat wave, flood).

BOX 1.1
STATEMENT OF TASK

This study will review the current state of knowledge about estimates of predictability of the climate system on intraseasonal to interannual timescales, assess in what ways current estimates are deficient, and recommend ways to improve upon the current predictability estimates. The study will also recommend research and model development foci and efforts that will be most beneficial in narrowing the gap between the current skill of predictions and estimated predictability limits. The review of predictability estimates to be addressed will include oceanic and atmospheric variables such as sea surface temperature, sub-surface heat content, surface temperature, precipitation, and soil moisture, as well as indices like Nino3.4 sea surface temperatures or the phases of the Madden-Julian Oscillation.

Specifically, the study committee will:

1. Review current understanding of climate predictability on intraseasonal to interannual time scales, including sources of predictability, the methodologies used to estimate predictability, current estimates of predictability, and how these estimates have evolved over time;

2. Describe how improvements in modeling, observational capabilities, and other technological improvements (e.g., analysis, development of ensemble prediction systems, data assimilation systems, computing capabilities) have led to changes in our understanding and estimates of predictability;

3. Identify any key deficiencies and gaps remaining in our understanding of climate predictability on intraseasonal to interannual timescales, and recommend research priorities to address these gaps;

4. Assess the performance of current prediction systems in relation to the estimated predictability of the climate system on intraseasonal to interannual timescales, and recommend strategies (e.g.,observations, model improvements, and research priorities) to narrow gaps that exist between current predictive capabilities and estimated limits of predictability; and

5. Recommend strategies and best practices that could be used to quantitatively assess improvements in both predictability estimates and prediction skill over time.

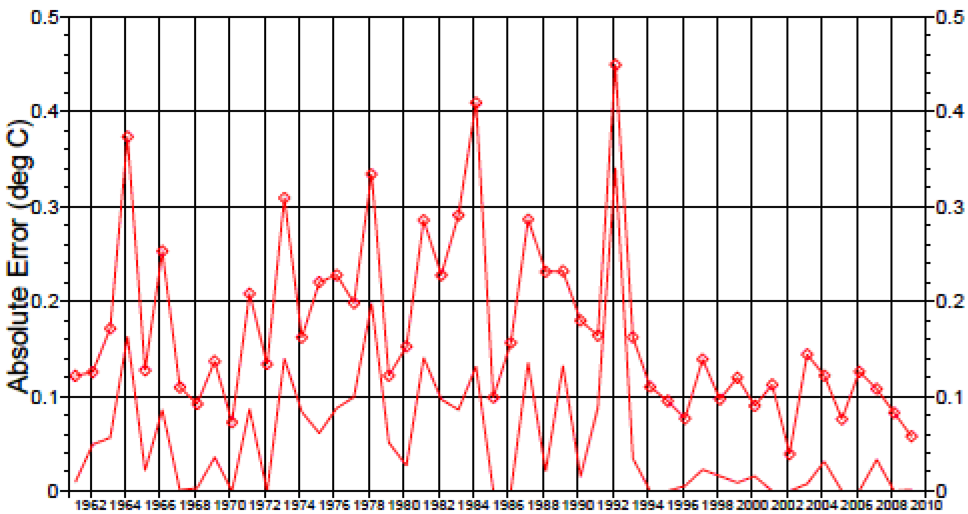

FIGURE 1.1 Time series of Mean Absolute Error (MAE) (thicker line with symbols) for the first three months of NINO3.4 predictions starting 1st February each year. Also shown (thin line, no symbols) is what is referred to as the Best Absolute Error (BAE), which is defined at each lead time as either zero (if the observations lie within the predicted range) or the distance between the observed value and the closest ensemble member, and then averaged over lead times. For a perfect forecasting system with a modest ensemble size, the BAE would be mostly zero, with occasional small positive values. The step change in skill after 1993 is evident. SOURCE: Stockdale et al. (2010), Fig 7a.

ISI PREDICTABILITY: THE EXAMPLE OF EL NIÑO-SOUTHERN OSCILLATION

The El Niño-Southern Oscillation (ENSO) serves as a prime example of a process that contributes to forecasts on intraseasonal to interannual (ISI) timescales, which extend from roughly two weeks to several years (see Box 1.2). Figure 1.2 shows the SST anomalies associated with one of the largest El Niño or warm ENSO events observed during the twentieth century. These anomalies tend to be at a maximum during the Northern Hemisphere winter and can persist on the order of months to a year. Although these anomalies are strongest in the equatorial Pacific Ocean, they affect winter temperature and precipitation globally, as shown in Figure 1.3. Current ISI forecast systems, which draw upon observations of the atmosphere and ocean as well as the physical and statistical relationships that describe the coupling between them, can often provide accurate predictions of the SST anomalies associated with ENSO. Figure 1.4 shows the predictions from a number of dynamical and statistical models for the SST anomaly in the equatorial Pacific several months in advance. Although the predictions track the behavior of observed SST anomalies relatively well, the spread among the models is substantial, sometimes even differing in the sign of the SST anomaly.

BOX 1.2
GENERAL TERMINOLOGY

Intraseasonal to interannual timescale (ISI)—roughly**,** two weeks to several years; the report focuses on predictions of the climate system on this timescale and the physical processes that are used to make these predictions.

Skill—the statistical evaluation of accuracy. Skill is most often determined by comparison of the disseminated forecast with a reference forecast, such as persistence, climatology, or objective guidance. Skill estimates can encompass deterministic estimates of skill, which are related to accuracy, or probabilistic estimates of skill, which are related to frequency of occurrence of specific events or thresholds. Skill is expressed quantitatively in terms of a specific metric.

Quality—the broad assessment of forecast performance encompassing a range of metrics, presumably related to the fidelity of physical processes (see also Kirtman and Pirani, 2008; Gottschalck et al. 2010).

Prediction—information on future climate (deterministic or probabilistic) from a specific tool (statistical or dynamical).

Forecast—issued guidance on future climate, which may take the form of quantitative outcomes, maps, and/or text. A forecast is usually (though not always) based on a "forecast system" that incorporates several prediction inputs or, at least, is based on the interpretation of an individual prediction input against past experience.

Model validation—comparison between observed and model-simulated climate. This may consider characteristics of climatology, variability or specific model processes.

Forecast verification—comparison between observations and forecasts over a specific time period, which typically involves more than one quantitative metric of skill.

Ensemble—a set of dynamical model runs from a single model, or from multiple models, that can be used to make a forecast. Within a single model, each model run differs from other members of the ensemble by a small perturbation in the initial state. For multiple models, it is assumed that the models differ in their physics and/or their parameterizations of sub-grid scale processes.

Note: these definitions are generally consistent with those appearing in the American Meteorological Society's Glossary of Meteorology (*http://amsglossary.allenpress.com/glossary*); in some cases, detail has been added to clarify usage in this report.

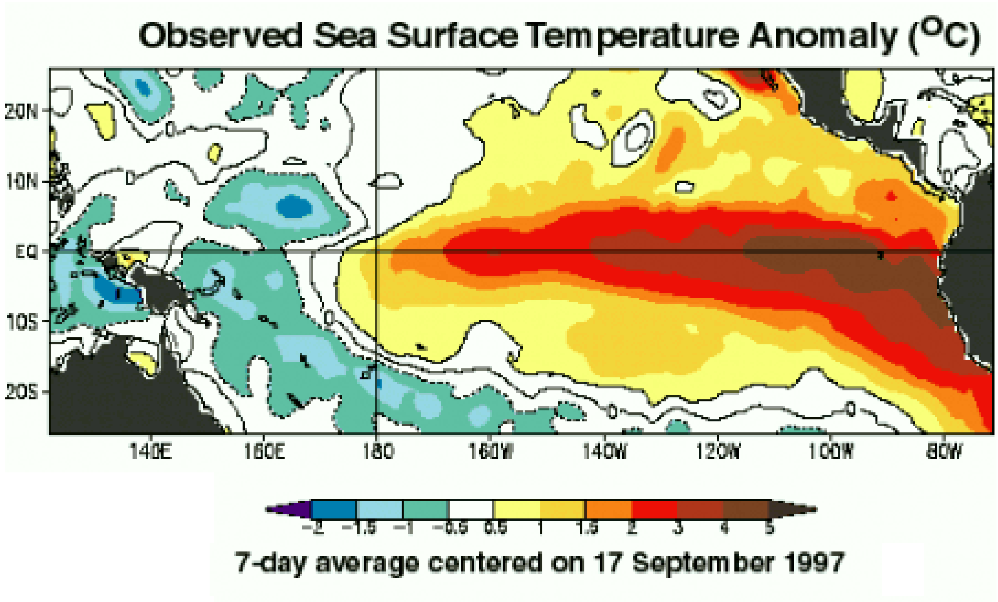

FIGURE 1.2 Sea surface temperature (SST) anomalies in the equatorial Pacific Ocean for a period during Fall 1997. This pattern is characteristic of a large amplitude El Niño event. SOURCE: CPC/NCEP/NOAA.

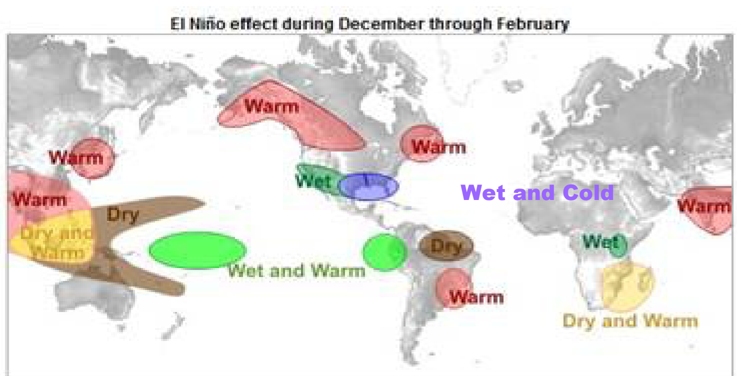

FIGURE 1.3 Patterns of anomalous temperature and precipitation during an El Niño episode for the Northern Hemisphere winter. SOURCE: Adapted from CPC/NCEP/NOAA.

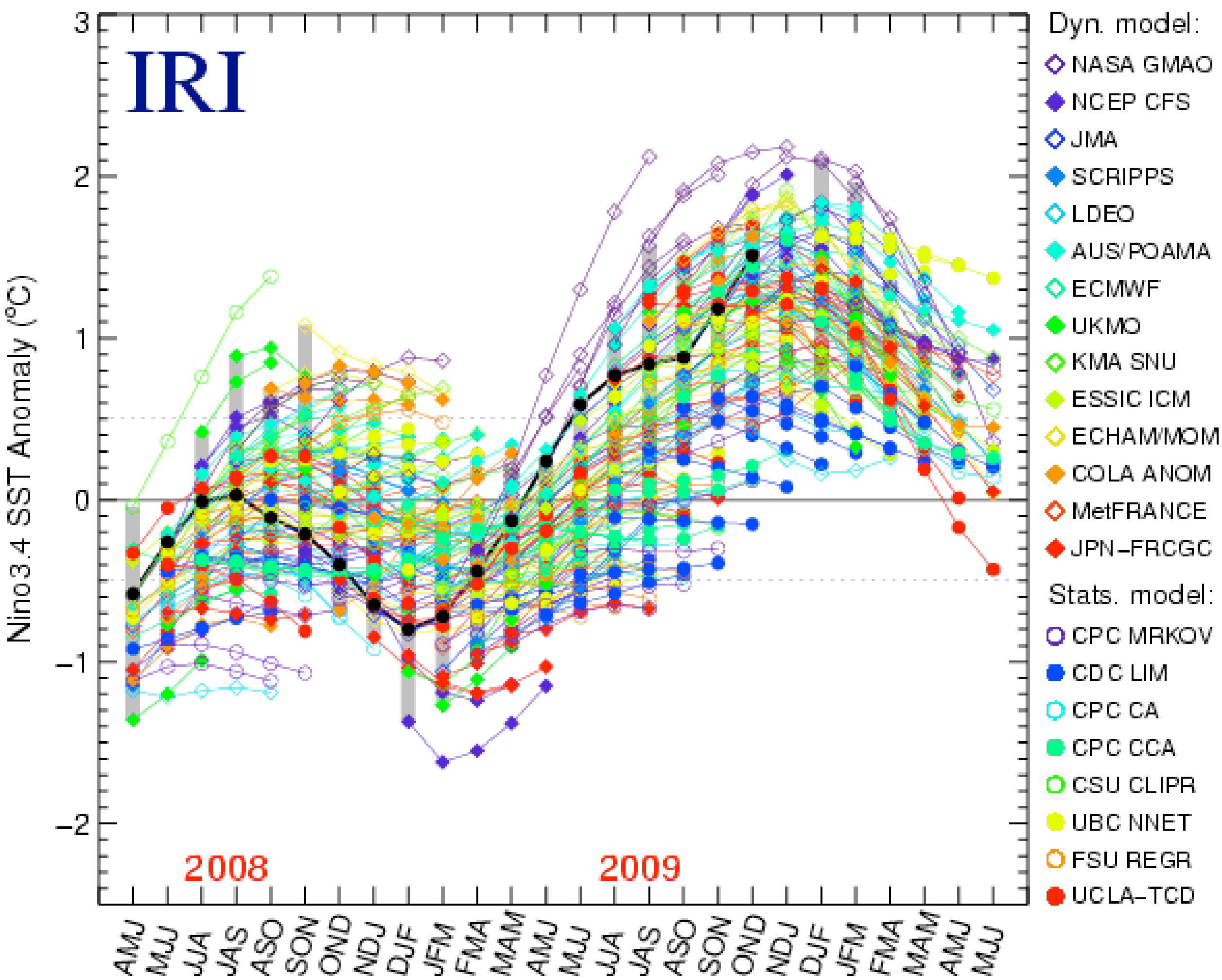

FIGURE 1.4 Predictions from various statistical (colored circles) and dynamical (colored diamonds) models along with observations (black circles) for SST anomalies in the equatorial Pacific Ocean. Many of the model predictions track the evolution of the anomalies, but the spread among the models is still rather large. The starting points of the models vary, and each prediction extends for approximately 5 months, since the predictions tend to diverge significantly after this period. SOURCE: International Research Institute for Climate and Society (IRI).

There is significant potential for societal benefit from improvements in ISI prediction quality and the provision of ISI forecasts. Many management decisions regarding water supplies, energy production, transportation, agriculture, forestry, and fisheries are made routinely on sub-seasonal, seasonal, or annual schedules. For example, during an El Niño winter, coastal areas in California may experience a heightened risk of flooding caused by an increase in precipitation as well as sea level height, while mountainous areas in the Pacific Northwest of the United States may experience less snowfall, reducing subsequent water availability. Thus, knowledge of the climate system at ISI timescales can be a useful input to making resource management and planning decisions.

Expanding our knowledge of processes affecting the climate on ISI timescales is an important priority. Many such processes have been identified (e.g., the Madden-Julian

Oscillation [MJO] and variability related to monsoonal circulations), but are not completely understood. In addition, ISI variability can be observed at a relatively high frequency (multiple times per year) when compared to longer-term phenomena (e.g., decadal or multi-decadal oscillations such as the Pacific Decadal Oscillation), providing researchers with a relatively greater number of "realizations" to exploit within the observational record.

ORGANIZATION OF THIS REPORT

The remainder of this report is organized into five chapters:

Chapter 2 reviews the concept of predictability, starting with an initial review of the historical background for climate prediction. Lorenz's work on weather prediction in the 1960s and 1970s is a foundation for present efforts; work in the 1980s extended prediction timescales by exploiting ENSO variability in the tropical Pacific and its associated teleconnections. Chapter 2 also introduces the view that a meaningful definition associates predictability with sources of variability, such as: 1) the inertia, or memory, of that state of the environment; 2) the patterns of interaction or coupling between variables, which include "teleconnections"; and 3) the response to external forcing. Various processes in the atmosphere, ocean, and land offer such sources of predictability. However, many gaps remain in our understanding of these processes. Chapter 2 also introduces the reader to the methodologies used to quantitatively estimate prediction skill and discusses model validation and forecast verification. Appendix A provides more technical detail about statistical methods.

Chapter 3 presents the reader with an introductory review of ISI forecasting followed by the committee's understanding of its critical components: observations, statistical models, dynamical models, and data assimilation. The processes for making and disseminating forecasts are also discussed, as well as their use by decision makers. It closes with the committee's summary of the potential improvements to current ISI forecast systems.

Chapter 4 uses three case studies to amplify and illustrate the state of and challenges facing efforts to improve ISI prediction. The three examples are ENSO, MJO, and soil moisture.

Chapter 5 defines the "Best Practices" that could be implemented to improve ISI predictions. This section also discusses some of the synthesizing issues given the content of the preceding chapters, exploring how the suggested activities could improve forecast quality, lead to more effective use of observations, and relate to the concept of "seamless" forecasting. In addition, realistic expectations for the speed and extent of improvements are discussed.

Chapter 6 presents the committee's recommendations and some remarks on their implementation.

2

Climate Prediction

This part of the report begins by reviewing the concept of predictability, starting with a summary of the historical background for climate prediction. Lorenz's work on weather prediction in the 1960s and 1970s is a foundation for present efforts. Progress in the 1980s extended prediction timescales, exploiting improved observational awareness of ENSO variability in the tropical Pacific and its associated teleconnections. Future improvements in prediction quality depend upon the ability to identify and understand patterns of variability and specific processes that operate on ISI timescales. Various processes in the atmosphere, ocean, and land offer sources of predictability; several are introduced in the following sections. Gaps in our present understanding of predictability are summarized to lay the foundation for discussion later in the report on how the future improvements are likely to be realized. In going forward, it will be necessary to assess the incremental skill gained from new sources of predictability. The methodologies to be used to quantitatively estimate prediction skill, validate models, and verify forecasts are discussed.

THE CONCEPT OF PREDICTABILITY

Lorenz in 1969 defined predictability as "a limit to the accuracy with which forecasting is possible" (Lorenz, 1969a). He later refined his view, providing two definitions of predictability (Lorenz, 2006): "intrinsic predictability—the extent to which the prediction is possible if an optimum procedure is used" and "practical predictability—the extent to which we ourselves are able to predict by the best-known procedures, either currently or in the foreseeable future." The forecasting that interested Lorenz and others during the 1960s and 1970s, which focused on weather and the state of the mid-latitude troposphere, provided much of the framework regarding forecasting and predictability that remains applicable to longer-range forecasts of the climate system and is reviewed here.
.

Atmospheric Predictability

Lorenz noted that practical predictability was a function of: (1) the physical system under investigation, (2) the available observations, and (3) the dynamical prediction models used to simulate the system. He noted in 2006 that the ability to predict could be limited by the lack of observations of the system and by the dynamical models' shortcomings in their forward extrapolations. While estimates of the predictability of day-to-day weather have been made by investigating the physical system, analyzing observations, and experimenting with models (Table 2.1), no single approach provides a definitive and quantitative estimate of predictability.

TABLE 2.1 Historical methods for evaluating predictability and their advantages and disadvantages.

Method and References	Description	Analysis
Physical System: Analytic closure (Leith, 1971)	Assuming that the atmosphere is governed by the laws of two-dimensional turbulence, a predictability limit can be estimated from the rate of error growth implied by the energy spectrum.	• Estimates are rough due to numerous assumptions. • Assumptions are stringent (e.g., atmospheric flow is non-divergent and moist processes are not important for error). • Difficult to extend to other aspects of real atmosphere.
Model: (Lorenz, 1965; Tribbia and Baumhefner, 2004; Buizza, 1997; Kalnay, 2003)	Using a dynamical model, experiments are designed to answer: How long is it expected to take for two random draws from the analysis distribution for this model and observing system to become practically indistinguishable from two random draws from the model's climatological distribution?	• Predictability results are highly dependent on the quality of the model being used. • Predictability is a function of the uncertainty in analyses used as model initial conditions.
Observations: Observed Analogs (Lorenz, 1969a; Van den Dool, 1994; Van den Dool et al., 2003)	The observed divergence in time of analogs (i.e., similar observed atmospheric states) provides an estimate of forecast divergence.	• Difficult to identify analogs and extrapolate the results to real atmosphere. Close analogs are not expected without a much longer observational record.

The studies listed in Table 2.1 demonstrate that for practical purposes (i.e., using available atmospheric observations and dynamical models), the limit for making skillful forecasts of mid-latitude weather systems is estimated to be approximately two weeks[5], largely due to the sensitivity of forecasts to the atmospheric initial conditions (see Box 2.1)[6]. However, their focus on weather and the state of the atmosphere excludes processes that are valuable for climate prediction. For instance, many factors external to the atmosphere were ignored, such as incoming solar radiation and the state of the ocean, land, and cryosphere. Single events, such as a volcanic eruption, that might influence predictability were not considered; nor were long-term trends in the climate system, such as global warming. In addition, the models were unable to replicate many features internal to the atmosphere, including tropical cyclones, the Quasi-Biennial Oscillation (QBO), the Madden Julian Oscillation (MJO), atmospheric tides, and low frequency atmospheric patterns of variability like the Arctic and Antarctic Oscillations. These additional features are important for the impacts that they may have on the estimates of weather predictability, as well as for their influence on predictability on longer climate timescales.

[5] The limit also depends on the quantitative skill metric being used.
[6] Model error also contributes to errors in weather prediction (e.g., Orrell et al., 2001).

BOX 2.1
WEATHER AND CLIMATE FORECASTS AND THE IMPORTANCE OF INITIAL CONDITIONS

Forecasts are computed as "initial value" problems: they require realistic models and accurate initial conditions of the system being simulated in order to generate accurate forecasts. Lorenz (1965) showed that even with a perfect model and essentially perfect initial conditions, the fact that the atmosphere is chaotic[7] causes forecasts to lose all predictive information after a finite time. He estimated the "limit of predictability" for weather as about two weeks, an estimate that still stands: it is generally considered not possible to make detailed weather predictions beyond two weeks based on atmospheric initialization alone. Lorenz's discovery was initially only of academic interest since, at that time, there was little quality in operational forecasts beyond two days, but in recent decades forecast quality has improved, especially since the introduction of ensemble forecasting. Useful forecasts now extend to the range of 5 to 10 days (see Figure 2.1).

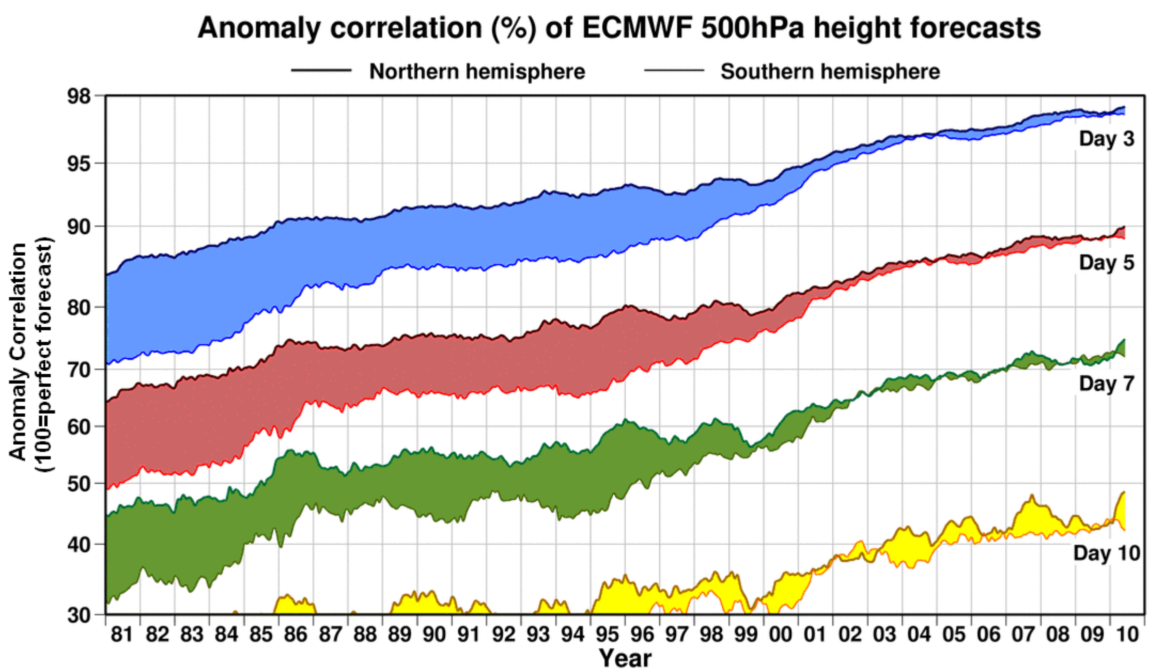

FIGURE 2.1. Evolution of ECMWF forecast skill for varying lead times (3 days in blue; 5 days in red; 7 days in green; 10 days in yellow) as measured by 500-hPa height anomaly correlation. Top line corresponds to the Northern Hemisphere; bottom line corresponds to the Southern hemisphere. Large improvements have been made, including a reduction in the gap in accuracy between the hemispheres. SOURCE: courtesy of ECMWF, adapted from Simmons and Hollingsworth (2002).

[7] Here, "chaotic" refers to a system that contains instabilities that grow with time.

The initial conditions for atmospheric forecasts are obtained through data assimilation, a way of combining short-range forecasts with observations to obtain an optimal estimate of the state of the atmosphere. Figure 2.2 shows the three factors on which the quality of the initial conditions depends: (1) good observations with good coverage, (2) a good model able to accurately reproduce the evolution of the atmosphere, and (3) an analysis scheme able to optimally combine the observations and the forecasts. The impressive improvement in 500-hPa geopotential height anomaly correlation in recent decades (Figure 2.1) has been due to improvements made in each of these three components. Since atmospheric predictability is highly dependent on the stability of the evolving atmosphere, ensemble forecasts made from slightly perturbed initial conditions have given forecasters an additional tool to estimate the reliability of the forecast. In other words, a minor error in an observation or in the model can lead to an abrupt loss of forecast quality if the atmospheric conditions are unstable.

For climate prediction on ISI timescales, the initial conditions involve phenomena with much longer timescales than the dominant atmospheric instabilities. For example, the SST anomalies associated with an El Niño event need to be known when establishing the initial conditions. Essentially, the initial conditions extend beyond the atmosphere to include details on the states of the ocean and land surface. From these long-lived phenomena, predictability of atmospheric anomalies can theoretically be extended beyond approximately two weeks to at least a few seasons.

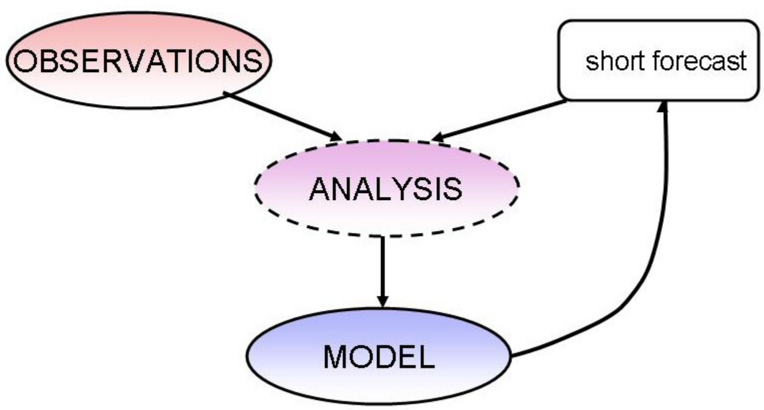

FIGURE 2.2 Schematic for data assimilation for an analysis cycle. The diagram shows the three factors that affect the initial conditions: observations, a model, and an analysis scheme.

Predictability of the Ocean and Atmosphere

As better observations led to an improved understanding of the climate system in the 1970s and 1980s, predictions of the atmosphere beyond the limits of the "classical" predictability proliferated. Statistical forecast systems had already demonstrated that predictions for time averages of some mid-latitude atmospheric quantities could be made at well past two weeks (Charney and Shukla, 1981). Observations of ENSO made it clear that some aspects of the tropical atmosphere could be predicted at longer lead-times as well. Observational, theoretical, and modeling studies (Horel and Wallace, 1981; Sarachik and Cane, 2010) demonstrated that there were relationships between variability observed in the tropical oceans and the variability of the extratropical atmosphere. It became clear that longer-range forecasts of atmospheric quantities could be made using predictions of the coupled ocean-atmosphere system.

Although operational, extended forecasts continued to focus on surface temperature and precipitation over continents, the atmospheric initial conditions were no longer considered important for making these forecasts; atmospheric ISI prediction was now considered a boundary value problem (Lorenz, 1975; Chen and Van den Dool, 1997; Shukla, 1998; Chu, 1999). Boundary forcing, initially from the ocean but later from the land and cryosphere (Brankovic et al., 1994), was used as the source of predictive information. This was appropriate because coupled models of the atmosphere, ocean, and land surface were still in their infancy and were not competitive with statistical prediction models (Anderson et al., 1999).

Given this context, researchers asked: if there exists a perfect prediction of ocean or land conditions, how well could the state of the mid-latitude atmosphere be predicted (Yang et al., 2004)? This question has been addressed observationally by estimating the signal-to-noise ratio. In this case the portion of the climate variance related to the lower boundary forcing is the signal, the portion of the climate variance related to atmospheric internal dynamics is the noise, and the ratio of the two represents one possible measure of predictability (e.g., Kang and Shukla, 2005). Such studies can lead to overly optimistic estimates of predictability because they assume that the boundary conditions are predicted perfectly.

There is an additional problem with this boundary-forced approach. These estimates assume that feedbacks between the atmosphere and the ocean do not contribute to the predictability. However, coupling between the atmosphere and the ocean can also be important in the evolution of SST anomalies (Wang et al., 2004; Zheng et al., 2004; Wu and Kirtman, 2005; Wang et al,. 2005; Kumar et al., 2005; Fu et al., 2003, 2006; Woolnough et al., 2007). Because the boundary-forced approach ignores this atmosphere-ocean co-variability (or any other climate system component couplings), these boundary-forced predictability estimates are of limited use.

Climate System Predictability

The techniques for estimating predictability shown in Table 2.1 can be applied to the coupled prediction problem (e.g., Goswami and Shukla, 1989; Kirtman and Schopf, 1998). However, each method is still subject to limitations similar to those mentioned in Table 2.1. Given the complexity of the climate system, estimates based on analytical closure are somewhat intractable (i.e., how can error growth rates from a simple system of equations relate to the real climate system?); approaches based on observations are limited by the relatively short length of the observational record, combined with the difficulty in identifying controlled analogs for a particular state of the climate. Non-stationarity in the climate system further reduces the chance that observed analogs would become useful in the foreseeable future, if ever. Model-based estimates are thus the most practical, but are still limited by the ability to measure the initial conditions for the climate and the mathematical representation of the physical processes.

As discussed in Chapter 1, most efforts to estimate prediction quality (or hindcast quality) are relatively recent, and involve analysis of numerous model-generated predictions for a similar time period (Waliser, 2005; Waliser, 2006; Woolnaugh et al. 2007; Pegion and Kirtman, 2008; Kirtman and Pirani, 2008; Gottschalck et al. 2010). For example, Kirtman and Pirani (2008) reported on the WCRP Seasonal Prediction Workshop in Barcelona where the participants discussed validating and assessing the quality of seasonal predictions based on a number of international research projects on dynamical seasonal prediction (e.g., SMIP2/HFP, DEMETER, ENSEMBLES, APCC). This collection of international projects includes a variety of different experimental designs (i.e., coupled vs. uncoupled), different forecast periods, initial condition start dates, and levels of data availability. Despite these differences, there was an attempt to arrive at consensus regarding the current status of prediction quality. Several different deterministic and probabilistic skill metrics were proposed, and it was noted that no single metric is sufficiently comprehensive. This is particularly true in cases where forecasts are used for decision support. Nevertheless, the workshop report includes an evaluation of multi-model prediction for Nino3.4 SSTA, 2m-temperature and precipitation in 21 standard land regions (Giorgi and Francisco, 2000). While it was recognized that the various skill metrics used were incomplete[8] and that there were difficulties related to the different experimental designs and protocols, the consensus was clear that multi-model skill scores were on average superior to any individual model (Kirtman and Pirani, 2008). Systematic efforts along the above lines for the intraseaosnal time scale have only recently begun with the development of an MJO forecast metric and a common approach to its application amongst a number of international forecast centers (Gottschalck et al. 2010) as well as the establishment of a multi-model MJO hindcast experiment (see www.ucar.edu/yotc/iso.html).

SOURCES OF PREDICTABILITY

Overview of Physical Foundations

Climate reflects a complex combination of behaviors of many interconnected physical and (often chaotic) dynamical processes operating at a variety of time scales in the atmosphere, ocean, and land. Its complexity is manifested in the varied forms of weather and climate variability and phenomena, and in turn, in their fundamental (if unmeasurable) limits of predictability, as defined above. Yet, embedded in the climate system are sources of predictability that can be utilized. Three categories can be used to characterize these sources of weather and climate predictability: inertia, patterns of variability, and external forcing. The actual predictability associated with an individual phenomenon typically involves interaction among these categories.

The first category is the “inertia” or “memory” of a climate variable when it is considered as a quantity stored in some reservoir of nonzero capacity, with fluxes (physical climate processes) that increase or decrease the amount of the variable within the reservoir over time, e.g., soil moisture near the land-atmosphere interface. Taking the top meter of soil as a control volume and the moisture within that volume as the climate variable of interest, the soil moisture increases with water infiltrated from the surface (rainfall or snowmelt), decreases with evaporation or transpiration, and changes further via within-soil fluxes of moisture through the sides and bottom of the volume. For a given soil moisture anomaly, the lifetime of the anomaly

[8] The particular metrics used to evaluate prediction quality were the multi-model Brier Skill Score for 2m-temperature and rainfall and the Mean Square Skill Score for the Nino3.4 SSTA.

(and thus our ability to predict soil moisture with time) will depend on these fluxes relative to the size of the control volume. Soil moisture anomalies at meter depth have inherent time scales of weeks to months. As panel (a) of Figure 2.3 shows, soil moisture anomalies exist considerably longer than the precipitation events that cause them.

Arguably, many variables related to the thermodynamic state of the climate system have *some* inertial memory that can be a source of predictability. Surface air temperature in a small regional control volume, for example, is a source of predictability that is very short given the efficiency of the processes (winds, radiation, surface turbulent fluxes, etc.) that affect it. If the air temperature at a given location is known at noon, its value at 12:05 PM that day can be predicted with a very high degree of certainty, whereas its predicted value days later is much more uncertain. In stark contrast, the inertial memory of ocean heat content can extend out to seasons and even years, depending on averaging depth. Examples of other variables with long memories include snowpack and trace gases (e.g., methane) stored in the soil or the ocean.

The second category involves patterns of variability—not variables describing the state of the climate and their underlying inertia, but rather *interactions* (e.g., feedbacks) between variables in coupled systems. These modes of variability are typically composed of amplification and decay mechanisms that result in dynamically growing and receding (and in some cases oscillating) patterns with definable and predictable characteristics and lifetimes. With modes of variability, predictability does not result from the decay of an initial anomaly associated with fluxes into and out of a reservoir, as in the first category, but rather with the prediction of the next stage(s) in the life cycle of the dynamic mode based on its current state and the equations or empirical relationships that determine its subsequent evolution. In many examples related to inertia or memory within the climate system, the atmosphere plays a "passive" and dissipative role in the evolution of the underlying anomaly. On the other hand, for the patterns of variability or feedbacks discussed here, the atmosphere plays a more active role in amplifying or maintaining an anomaly associated with processes occurring in the ocean or on land.

"Teleconnections" is a term used to describe certain patterns of variability, especially when they act over relatively large geographic distances. Teleconnections illustrate how interaction among the atmosphere, ocean, and land surface can "transmit" predictability in one region to another remote region. For example, during ENSO events, features of the planetary scale circulation (e.g., the strength and location of the mid-latitude jet stream) interact with anomalous convection in the tropical Pacific. These interactions can lead to anomalous temperature and precipitation patterns across the globe (panel b of figure 2.3). Thus, predictions of tropical Pacific sea surface temperature due to ENSO can be exploited to predict air temperature anomalies in some continental regions on the time scales of months to seasons. For air temperature, this teleconnection pattern offers enhanced predictability compared to memory alone, which would only be useful for minutes to hours. It should be noted that the predictability of teleconnection responses (in the above example, air temperature in a location outside of the tropical Pacific) will be lower than that of the source (in the above example, tropical Pacific SST) because of dynamical chaos that limits the transmission of predictability.

The third category involves the response of climatic variables to external forcing, and it includes some obvious examples. Naturally, many Earth system variables respond in very predictable ways to diurnal and annual cycles of solar forcing and even to the much longer cycles associated with orbital variations. Other examples of external forcing variations that can provide predictability include human impacts—long-term changes in atmospheric aerosols, greenhouse gas concentrations, and land use change.

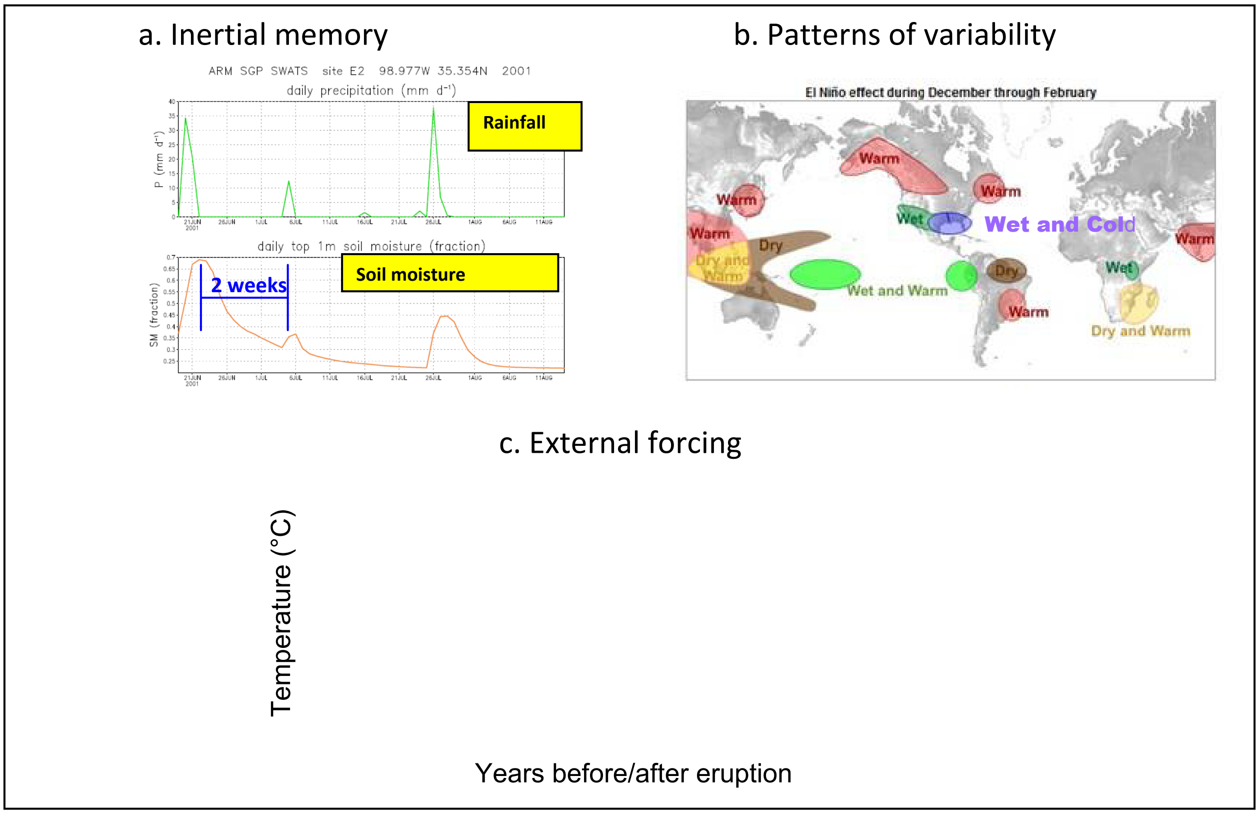

FIGURE 2.3 (a) Example of inertial memory. A positive soil moisture anomaly at the Atmospheric Radiation Measurement/Cloud and Radiation Testbed (ARM/CART) site in Oklahoma decreases with a time scale much longer than the atmospheric events that caused it. SOURCE: Greg Walker, personal communication. Soil moisture time scales measured at other sites are even longer than this (Vinnikov and Yeserkepova, 1991). (b) Example of teleconnections. Map of El Niño impacts on global climate, for December–February. SOURCE: Adapted from CPC/NCEP/NOAA (c) Example of external forcing. Global mean temperature anomaly prior (negative *x*-axis values) and following (positive *x*-axis values) volcanic eruptions, averaged for 6 events. Substantial cooling is observed for nearly 2 years following the date of eruption. The dark line has the ENSO events removed; the light line does not. SOURCE: Robock and Mao (1995).

Examples of Predictability Sources

Figure 2.4 provides a quick glimpse of various predictability sources in terms of their inherent time scales. This view, based on time scale, is an alternative or complement to the three-category framework (inertia, patterns of variability, and external forcing). Provided in the present section is a broad overview of predictability sources relevant to ISI time scales. Some of the examples will be discussed more comprehensively in later chapters.

It is important to realize that the timescales associated with sources of predictability often arise from a combination of inertia and feedback processes. Also, it should be noted that the

FIGURE 2.4 Processes that act as sources of ISI climate predictability extend over a wide range of timescales, and involve interactions among the atmosphere, ocean, and land. CCEW: convectively coupled equatorial waves (in the atmosphere); TIW: tropical instability wave (in the ocean); MJO/MISV: Madden-Julian Oscillation/Monsoon intraseasonal variability; NAM: Northern Hemisphere annular mode; SAM: Southern Hemisphere annular mode; AO: Arctic oscillation; NAO: North Atlantic oscillation; QBO: quasi-biennial oscillation, IOD/ZM: Indian Ocean dipole/zonal mode; AMOC: Atlantic meridional overturning circulation. For the *y*-axis, "A" indicates "atmosphere;" "L" indicates "land;" "I" indicates "ice;" and, "O" indicates "ocean."

timescales in Figure 2.4 indicate the timescale of the *variability* associated with a particular process. This is distinct from the timescale associated with a *prediction*. For example, ENSO exhibits variability on the scale of years; however, information about the state of ENSO can be useful for making ISI predictions on weekly, monthly, and seasonal time scales.

As discussed in Chapter 1 (see Committee Approach to Predictability), it can be difficult to quantify the intrinsic predictability associated with any of the individual processes depicted in Figure 2.4 (i.e., for what lead-time is an ENSO prediction viable? And to what extent would that prediction contribute to skill for predicting temperature or precipitation in a particular region?). As mentioned earlier (see Climate System Predictability), prediction experiments form the foundation of our understanding. However, these experiments are rarely definitive in quantifying such limits of predictability. For example, for ENSO, there are three competing theories (inherently nonlinear; periodic, forced by weather noise; and the damped oscillator) that underlie various models of ENSO, each with its own estimate of predictability (see Kirtman et al., 2005 for a detailed discussion). At this time we are unable to resolve which theory is correct since all yield results that are arguably "consistent" with observational estimates. To further complicate the understanding of the limits of predictability for ENSO, there are important interactions with other sources of predictability that may enhance or inhibit the predictability associated with ENSO (see Chapter 4). ENSO is just one example of how understanding what "sets" the predictability associated with a particular process is a critical challenge for the ISI prediction community. The challenge of improving forecast quality necessitates enhancing the individual building blocks (see Chapter 3) that make up our predictions systems, but it also requires a deeper understanding of the physical mechanisms and processes that are the sources of predictability.

Inertia

Upper ocean heat content

On seasonal-to-interannual time scales upper ocean heat content is a known source of predictability. The ocean can store a tremendous amount of heat. The heat capacity of 1 m^3 of seawater is 4.2×10^6 joules m^{-3} K^{-1} or 3,500 times that of air and 1.8 times that of granite. Sunlight penetrates the upper ocean, and much of the energy associated with sunlight can be absorbed directly by the top few meters of the ocean. Mixing processes further distribute heat through the surface mixed layer, which can be tens to hundreds of meters thick. As Gill (1982) points out, with the difference in heat capacity and density, the upper 2.5 m of the ocean can, when cooling 1°C, heat the entire column of air above it that same 1°C. The ocean can also transport warm water from one location to another, so that warm tropical water is carried by the Gulf Stream off New England, where in winter during a cold-air outbreak, the ocean can heat the atmosphere at up to 1200 W m^{-2}, a heating rate not that different from the solar constant. Stewart (2005) shows that a 100 m deep ocean mixed layer heated 10°C seasonally stores 100 times more heat than 1 m thick layer of rock heated that same 10°C; as a result the release of the heat from the ocean mixed layer can have a large impact on the atmosphere. Thus, the atmosphere acts as a "receiver" of any anomalies that have been stored in the ocean, and predictions of the evolution of air temperature over the ocean can be improved by consideration of the ocean state.

Soil moisture

Soil moisture memory spans intraseasonal time scales. Memory in soil moisture is translated to the atmosphere through the impact of soil moisture on the surface energy budget, mainly through its impact on evaporation. Soil moisture initialization in forecast systems is known to affect the evolution of forecasted precipitation and air temperature in certain areas during certain times of the year on intraseasonal time scales (e.g., Koster et al., 2010). Model studies (Fischer et al., 2007) suggest that the European heat wave of summer 2003 was exacerbated by dry soil moisture anomalies in the previous spring.

Snow cover

Snow acts to raise surface albedo and decouple the atmosphere from warmer underlying soil. Large snowpack anomalies during winter also imply large surface runoff and soil moisture anomalies during and following the snowmelt season, anomalies that are of direct relevance to water resources management and that in turn could feed back on the atmosphere, potentially providing some predictability at the seasonal time scale. The impact of October Eurasian snow cover on atmospheric dynamics may improve the prediction quality of northern hemisphere wintertime temperature forecasts (Cohen and Fletcher, 2007). The autumn Siberian snow cover anomalies can be used for prediction of the East Asian winter monsoon strength (Jhun and Lee, 2004; Wang et al., 2009).

Vegetation

Vegetation structure and health respond slowly to climate anomalies, and anomalous vegetation properties may persist for some time (months to perhaps years) after the long-term climate anomaly that spawned them subsides. Vegetation properties such as species type, fractional cover, and leaf area index help control evaporation, radiation exchange, and momentum exchange at the land surface; thus, long-term memory in vegetation anomalies could be translated into the larger Earth system (e.g. Zeng et al., 1999).

Water table variations

Water table properties vary on much longer timescales (years or more for deep water tables) than surface soil moisture. Some useful predictability may stem from these variations, though the investigation of the connection of these variations to the overall climate system is still in its infancy, in part due to a paucity of relevant observations in time and space.

Land heat content

Thermal energy stored in land is released by molecular diffusion and thus over all time scales, but with a rate of release that decreases with the square root of the time scale. In practice, there is strong diurnal storage (up to 100 W m^{-2}) of heat energy and a still significant amount over the annual cycle (up to 5 W m^{-2}). This is particularly strong in relatively unvegetated regions where solar radiation is absorbed mostly by the soil, since vegetation has much less thermal inertia, or in higher latitudes where soil water seasonally freezes.

Polar sea ice

Sea ice is an active component of the climate system and is highly coupled with the atmosphere and ocean at time scales ranging from synoptic to decadal. When large anomalies are established in sea ice, they tend to persist due to inertial memory and to positive feedback in the atmosphere-ocean-sea ice system. These characteristics suggest that some aspects of sea ice may be predictable on ISI seasonal time scales. In the Southern Hemisphere, sea ice concentration anomalies can be predicted statistically by a linear Markov model on seasonal time scales (Chen and Yuan, 2004). The best cross-validated skill is at the large climate action centers in the southeast Pacific and Weddell Sea, reaching 0.5 correlation with observed estimates even at 12-month lead time, which is comparable to or even better than that for ENSO prediction. We have less understanding of how well sea ice impacts the predictability of the overlying atmosphere.

Patterns of Variability

Different components of the climate system, each with their own inertial memory, interact with each other in complex ways. The dynamics of the feedbacks and interactions can lead to the development of predictable modes, or patterns, of variability.

It should be noted that the descriptions for the patterns of variability provided in the following subsections describe their "typical" behavior, focusing on commonalities among observed events and the mechanisms that drive the phenomena. In reality, the manifestation or impact of a pattern may differ from these "typical" cases since the various patterns of variability can be affected by one another as well as by the unpredictable "noise" inherent to the climate system, especially in the atmosphere. For example, not all ENSO events have the same features, and in some cases, these differences among events can be understood from interactions between ENSO and the MJO (see the MJO case study in Chapter 4).

Low-frequency equatorial waves in the atmosphere and ocean

The equator provides an efficient wave guide by which tropical dynamical energy is organized, propagated, and dissipated. In the atmosphere, equatorial Kelvin and Rossby waves and mixed Rossby-Gravity waves (Matsuno, 1966) are observed. Due to the moist and vertically unstable nature of the tropics, these low-frequency waves are often associated with convection and are referred to as convectively-coupled equatorial waves (CCEWs) (Wheeler and Kiladis, 1999; Kiladis et al., 2009). The spatial scales of these disturbances can be quite large (on the order of thousands of kilometers), and their time scales for propagating across ocean basins can be of the order of days to weeks. Figure 2.5 shows a time-longitude plot of equatorial outgoing longwave radiation (OLR) anomalies, produced following a wavenumber-frequency analysis. OLR is a good proxy for deep tropical convection, and the colors in Figure 2.5 show areas of enhanced (hot colors) or suppressed (cool colors) convection. These patterns in OLR correspond to characteristic types of waves (green, blue, and black ovals), illustrating that variability in the tropical atmosphere is consistent with the simplified theory of Matsuno (Kiladis et al., 2009). Figure 2.5 also demonstrates the manner in which these waves are manifest in relation to the typical background variability. Although complicated by their coupling to atmospheric convection, the organization and propagation of these low-frequency waves provides an element of predictability for the tropical atmosphere and possibly the extra-tropics via teleconnections.

Analogous to the discussion of the atmosphere above, the equatorial ocean supports the presence of equatorial wave modes, such as the Kelvin, Rossby and mixed-Rossby gravity modes. One simplifying aspect for their presence in the ocean is that, in contrast to the atmosphere, no convection or phase changes are involved. Because the equivalent depth of the ocean is considerably smaller than that for the atmosphere, its propagation speeds are much slower, and thus the time scale (and the predictability that arises from it) is much longer (e.g., a Kelvin wave takes about 2–3 months to cross the Pacific Ocean). These waves play a crucial role in the ocean thermocline adjustment and ENSO turnaround, as discussed below.

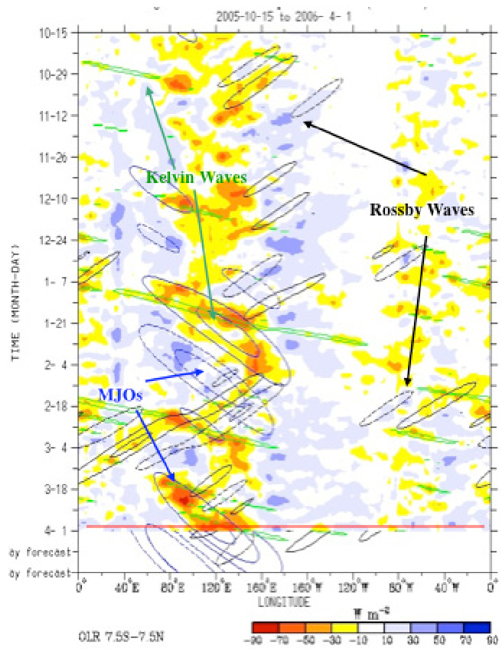

FIGURE 2.5 Some atmospheric waves offer an important source of predictability. The time-longitude diagram depicts the speed and direction of propagation that Kelvin waves (green ovals), Rossby waves (black ovals), and waves associated with the MJO (blue ovals) can exhibit in the tropics. The dense shading, which often overlaps with the position of the ovals, corresponds to anomalies in outgoing longwave-radiation (OLR); positive OLR anomalies indicate clear skies and suppressed convection; negative OLR anomalies indicate enhanced convection. SOURCE: Adapted from Wheeler and Weickmann (2001).

Madden-Julian Oscillation (MJO)

Another fundamental mode of tropical convectively-coupled wave-like variability is the Madden-Julian Oscillation (MJO; Madden and Julian, 1972, 1994). MJOs operate on the planetary scale, with most of the convective disturbance and variations occurring in the Indo-Pacific warm pool regions. The typical time scale of these quasi-periodic disturbances is of the order of 40–50 days. They tend to propagate eastward in boreal winter and north and/or northeastward in boreal summer. They strongly influence the onsets and breaks of the Australian and Asian monsoons and are sometimes referred to as monsoon intraseasonal variability (MISV) or oscillation (MISO). As with the CCEWs mentioned above, they are thought to be a source of both local predictability and predictability in the extra-tropics. The MJO and its associated predictability are discussed in more detail in Chapter 4 of this report. Figure 2.6 illustrates composite MJO events for boreal summer (May–October).

SST and mixed layer feedback on subseasonal time scales

The large/planetary spatial and subseasonal time scales of the CCEWs and MJO discussed above, along with the often strong impact of these phenomena on surface fluxes via wind speed and cloudiness, can result in significant modulation of the ocean surface mixed layer,

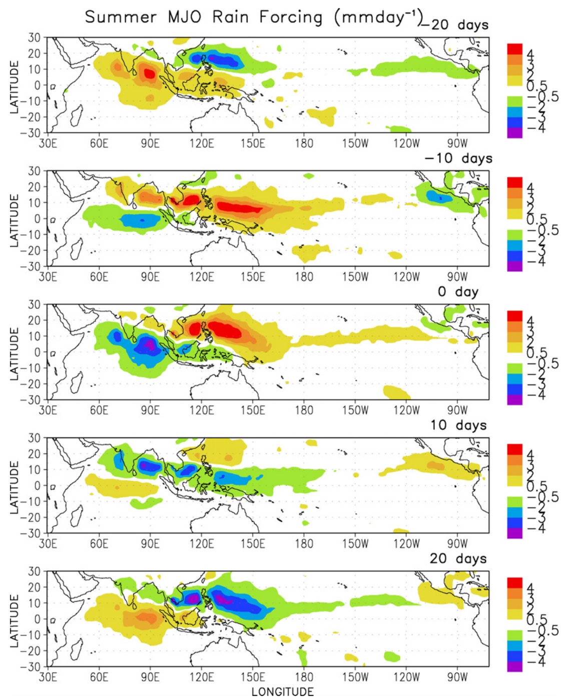

FIGURE 2.6 Characteristic rainfall patterns (mm per day) before, during, and following an MJO event during the boreal summer (May–October). Dry anomalies are indicated by "cool" colors (green, blue, purple) and wet anomalies are indicated by "hot" colors (yellow, orange, red). SOURCE: Waliser et al. (2005).

with variations in depth on the order of tens of meters and in temperature of the order of a degree. This process can impart a feedback onto the atmospheric wave processes which influences their subsequent evolution (e.g. amplitude, propagation speed).

Annular Modes (Northern or Southern, NAM or SAM)

The Annular Modes, also refereed to as the Arctic Oscillation (Figure 2.7) in the Northern Hemisphere, or the Antarctic Oscillation in the Southern Hemisphere, are dominant modes of variability outside the tropics. They are established on a weekly time scale due to atmospheric internal dynamics (such as mean flow-wave interaction or stratosphere-troposphere interaction). They offer some predictability on seasonal time scales through longer-timescale persistence of stratospheric winds (Baldwin and Dunkerton, 1999). The modes can influence surface temperature and precipitation, especially the frequency of extreme events (Thompson and Wallace, 2001).

The manifestation of the Arctic Oscillation in the Atlantic sector is commonly referred to at the North Atlantic Oscillation (NAO). An index for the NAO is typically formed from the difference in sea-level pressure between the Azores and Iceland. High index values correspond to stronger westerly flow across the North Atlantic, an intensification and northward shift of the storm track (Rogers, 1990) and warmer and wetter winters in northern Europe (Hurrell 1995). Covarying with the NAO there is an associated tripole pattern of sea surface temperature anomalies (Deser and Blackmon 1993).

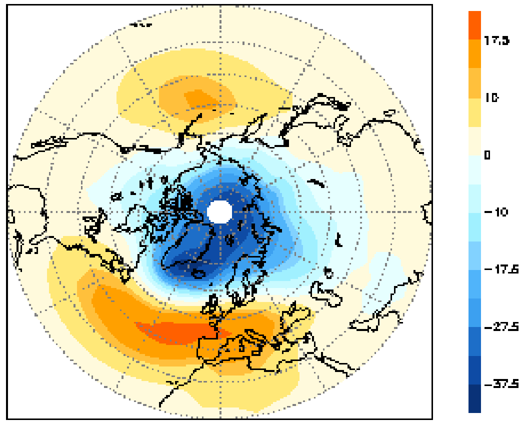

FIGURE 2.7 Characteristic pattern of anomalous sea level pressure (SLP; in hPa) associated with the positive polarity of the Arctic Oscillation (AO) in the winter. Blue indicates lower than normal SLP and red indicates higher than normal SLP; this phase of the AO exhibits an enhanced westerly jet over the Atlantic Ocean in the mid-latitudes. The North Atlantic Oscillation can be thought of as the portion of the AO pattern that resides in the Atlantic sector. SOURCE: Adapted from Thompson and Wallace (2000).

The NAO is the single largest contributing pattern to European interannual variability and plays an important role in predictions of European winter climate. However, the ability to predict the NAO on seasonal timescales is limited in current generation of models used for seasonal forecasting. There is some evidence that variability in the Atlantic Gulf Stream can influence the long-term variability of the NAO (Wu and Gordon, 2002). In addition, there is evidence of forcing of the NAO by ENSO (Bronniman et al., 2007; Ineson and Scaife 2009) and the stratospheric Quasi-Biennial Oscillation (Boer and Hamilton 2008).

Stratosphere-Troposphere Interaction, Quasi-Biennial Oscillation (QBO)

Since the stratosphere can interact with the troposphere, knowledge of the state of the stratosphere can serve almost as a boundary condition when attempting to simulate the troposphere. The stratospheric circulation can be highly variable, with a time scale much longer than that of the troposphere. The variability of the stratospheric circulation can be characterized mainly by the strength of the polar vortex, or equivalently the high latitude westerly winds. Stratospheric variability peaks during Northern winter and Southern late spring. When the flow just above the tropopause is anomalous, the tropospheric flow tends to be disturbed in the same manner, with the anomalous tropospheric flow lasting up to about two months (Baldwin et al., 2003a, 2003b). Generally, the surface pressure signature looks very much like the North Atlantic Oscillation or Northern Annular Mode. Surface temperature signals are also similar to those from the NAO and SAM and there are associated effects on extremes (Thompson et al., 2002). In sensitive areas such as Europe in winter, experiments suggest that the influence of stratospheric variability on land surface temperatures can exceed the local effect of sea surface temperature.

Sudden stratospheric warming events serve as an extreme example of how the stratosphere could serve as a source of predictability. During a sudden warming event the polar vortex abruptly (over the course of a few days) slows down, leading to an increase in polar stratospheric temperature of several tens of degrees Kelvin. Although attenuated, over the course of the following weeks the warming signal migrates downward into the troposphere with signals that can be detected in the surface climate approximately a month following the warming event (Limpasuvan et al., 2004).

Additionally, the Quasi-Biennial Oscillation (QBO) of the stratospheric circulation offers a source of predictability for the tropospheric climate. The stratospheric QBO in the tropics arises from the interaction of the stratospheric mean flow with eddy fluxes of momentum carried upward by Rossby and gravity waves that are excited by tropical convection. The result is an oscillation in the stratospheric zonal winds having a period of about 26 months. While our weather and climate models do not often resolve or represent the QBO well, it is one of the more predictable features in the atmosphere, and it has been found to exhibit a signature in surface climate (Thompson et al., 2002).

Tropical Instability Waves (TIWs) in the ocean

TIWs are most prevalent in the eastern Pacific Ocean and are evident in SST and other quantities such as ocean surface chlorophyll and even boundary-layer cloudiness, particularly just north of the equator. They arise from shear-flow and baroclinic instabilities and result in westward propagating wave-like features having length scales on the order of 1000s of km and time scales of about 1–2 weeks. There is evidence that they may affect the overlying atmosphere. (Hoffman et al., 2009). That the strong SST gradients associated with the TIWs affect the surface winds has been documented by Chelton et al. (2001). Song et al. (2009) suggest that atmospheric models should improve the realism of their coupling between the atmosphere and ocean mesoscale variability in SST in order to correctly capture small scale variability in the wind field.

El Niño-Southern Oscillation (ENSO), involving subsurface ocean heat content

The evolution of ENSO can be predicted one to a few seasons in advance using coupled atmosphere-ocean models (e.g., Zebiak and Cane, 1987), and SST anomalies in the tropical Pacific Ocean can contribute to predictions of the global atmospheric circulation at seasonal leads (e.g., Shukla, 1998). Accurate prediction of ENSO is a key objective and benchmark of many operational seasonal forecasting efforts, as discussed in more detail in Chapter 4.

Rossby wave energy propagation in the atmosphere underlies important teleconnections involving ENSO, and can be understood by considering relatively basic climate dynamics. Rossby waves conserve absolute vorticity. Hence away from sources and sinks, north-south transport of the vorticity from the earth's rotation is balanced by advective transport of wave disturbance vorticity. As a consequence, in the mid-latitudes where large scale winds are predominantly westerly, wavelike disturbances are possible in the wind, pressure, and temperature patterns, whereas in the tropics where easterly winds are dominant, forced disturbances remain localized. These waves provide teleconnections between the tropics and the midlatitudes.

The teleconnections associated with ENSO can be profoundly important because many populated areas would otherwise not be affected by this source of predictability. Figure 2.3 (panel b), for example, illustrates some of the teleconnections associated with the ENSO cycle. The positive phase of the ENSO cycle tends to promote warm conditions in the northeast United

States, even though this area is not involved in the dynamics underlying ENSO. In essence, the northeast United States is a passive recipient of ENSO predictability through a global-scale teleconnection process.

ENSO itself can be related to other patterns of variability. For example, westerly wind bursts associated with the MJO may help to trigger ENSO events (see the ENSO case study in Chapter 4). Also, Yuan (2004) describes a teleconnection process between ENSO and the Antarctic Dipole, a separate climate mode. ENSO forcing triggers the Antarctic Dipole, with implications for sea ice prediction at seasonal timescales.

Indian Ocean Dipole/Zonal Mode (IOD/IOZM)

A coupled mode of interannual variability has been found in the equatorial Indian Ocean in which the normally positive SST gradient is significantly weakened or reversed for a period on the order of a season (Saji et al., 1999; Webster et al., 1999). It can result in significant regional climate impacts, such as in east Africa and southern Asia. The independence of this mode and its connections to ENSO are still being investigated, but in any case the Indian Ocean Dipole/Zonal Mode (IOD/IOZM), like ENSO, appears to offer an intermittent source of interannual predictability. Similar to ENSO, the IOD/IOZM involves equatorial SST-wind-thermocline/upwelling feedback (Bjerknes, 1969); however, in contrast to ENSO, it also involves off-equatorial, SST-convection-atmospheric Rossby wave interaction (Li et al., 2003; Wang et al., 2003). The latter is strongly regulated by seasonal reversal of the monsoon circulation, hence the IOD/IOZM lasts only a season or two. The off-equatorial, SST-convection-Rossby wave interaction can maintain cooling of western North Pacific SST and anomalous anticyclonic circulation during the decaying phase of ENSO, providing a source of predictability for the East Asian summer monsoon (Wang et al., 2000).

External Forcing

Greenhouse gases (CO_2, etc.)

Greenhouse gases have a direct impact on the radiation balance of the atmosphere: increases in greenhouse gases warm the global climate. The non-stationarity associated with this climate change is an important component of climate forecasts even on ISI timescales. For example, the NOAA Climate Prediction Center uses optimal climate normals and other empirical techniques to capture this non-stationarity in climate forecasts (Huang et al., 1996a; Livezey et al., 2007). However, regional details of this climate change are difficult to model numerically due to the myriad important feedbacks that need to be taken into account. These include feedbacks due to the enhancement of water vapor in the warming atmosphere and the associated changes in cloudiness and snow/ice amount, all of which can affect the radiation budget. In addition, there are feedbacks from the carbon cycle itself (including the release of additional greenhouse gases in northern latitudes as permafrost melts), the ocean thermohaline circulation, changes in the biosphere, and so on.

Anthropogenic Aerosols

Atmospheric aerosols, which affect the radiation budget of the Earth, include major human-related components that change with the nature of human activities, and thus which may be predictable. The human-related components include sulfate aerosols from fossil fuel combustion and organic aerosols from biomass burning and land use change.

The effect of changes in aerosols on precipitation at ISI timescales could be important. Bollasina and Nigam (2009) have shown that elevated aerosol concentrations over the Indian subcontinent can accompany periods of reduced cloudiness, increased downward shortwave radiation, and ultimately a delayed onset of the monsoon. However, the role of aerosols as the "cause" of a decrease or delay in precipitation is not yet confirmed—more research on sub-seasonal timescales is required to isolate the effect of aerosols from the influence of the large-scale synoptic flow and associated changes in precipitation.

Land use change

Humans have had a marked impact on the character of the land surface through deforestation, agricultural conversion, and urbanization. This change in surface character can have a long-term impact on surface energy and water budgets (e.g., deforested land may generate less evapotranspiration than forested land), which in turn can have a long-term impact on the rest of the climate system.

Fluctuations in solar output

The sun provides the energy that powers the Earth's climate system. Its output varies slightly with an 11-year cycle that is highly predictable because it is nearly periodic. Larger changes may occur on longer time scales, but in the absence of measurements, these changes cannot be quantified beyond a statement that they appear to be small compared to the signal seen from greenhouse gases. As discussed by Haigh et al. (2005), it is likely that the mechanism that links solar fluctuations to surface climate involves the communication of anomalies between the stratosphere and troposphere, which is discussed in the "Gaps in Our Knowledge" section in this chapter.

Volcanoes and other high-impact modifications of atmospheric composition

There are a number of rare external forcing events that can cause a sudden drastic change to the atmospheric burden of aerosols, trace gases, and particulates. Volcanoes are one example. Major eruptions are relatively rare (less than one per decade) but can quickly inject large volumes of material high into the atmosphere. The effects on the climate system can be felt for years afterward, typically as a cooling of the global mean temperature (Robock and Mao, 1995). The impacts of major eruptions on temperature distributions over the continental United States can be larger than those from internal variations of the climate system discussed earlier in this section (Bradley, 1988). Shortened growing seasons caused by an overall reduction of solar radiation reaching the ground could have negative impacts on particular crops in some regions.

Forest fires provide another example of relatively rapid changes in atmospheric composition that can affect climate on ISI time scales. The Indonesian fires in 1997–98 helped to exacerbate the very strong El Niño drought. The aerosol loading altered regional radiative balances (Davison et al., 2004), affecting precipitation in ways that were different from those predicted on the basis of the El Niño event (Graf et al., 2009; Rotsayn et al., 2010).

Other rare events that could significantly modify atmospheric composition include nuclear exchange (Robock et al., 2007) or impacts from space. While most of these high impact events cannot be predicted with any accuracy, intraseasonal to interannual climate predictions subsequent to the event would be affected by the changed atmospheric composition. Given a modern climate prediction system, it should be possible to observe and analyze the concentration of injected material in the atmosphere and produce ISI forecasts (Ramachandran et al., 2000).

Gaps in Our Knowledge

Our understanding of ISI climate predictability—both of its sources and extent—is still far from complete. Numerous gaps still exist in our observations of climate processes and variability, in our inclusion of the wide range of relevant processes in models, and in our knowledge of the sources of predictability:

- We cannot yet claim to have identified all of the reservoirs, linkages, and teleconnection patterns associated with predictability in the Earth system. For many of the predictability sources we *have* identified, we cannot claim to understand fully the mechanisms that underlie them. The observational record contains many non-stationary trends that may relate to predictability but are not yet adequately explained. The science is proceeding but is encumbered by the overall complexity of the system.
- The models that have been used to evaluate the known sources of predictability and to make forecasts are known to be deficient in many ways. Many key processes associated with predictability occur at spatial scales that cannot be resolved by current models. Examples in the atmosphere include cumulus convection, boundary-layer turbulence, and cloud-aerosol microphysics; examples in the ocean include horizontal transports associated with eddies and vertical mixing. In addition, processes associated with the coupling of the ocean or the land surface to the atmosphere through the exchanges of heat, fresh water, and other constituents can be difficult to resolve. The models thus rely on parameterizations, which are simple approximations that often have to be "tuned", making them undependable in untested situations. A wealth of literature is available on the deficiencies of current, state-of-the-art climate models; it indicates that currently available dynamical models do not always outperform simple empirical models or persistence metrics.
- Even if the models used were perfect—even if they included and represented accurately all physical and dynamical processes relevant to predictability at adequate spatial and temporal resolution—they would still be limited in their ability to make accurate forecasts by deficiencies in our ability to initialize prognostic fields.

The degree to which these gaps limit our ability to make forecasts—or, stated another way, the improvement we could make in forecasts if these gaps were fully addressed—is difficult to ascertain. Exploring sources of predictability, in particular addressing gaps in our understanding of these sources, might yield substantial improvements in forecast performance. Here we briefly outline several sources of predictability for which gaps in understanding can be clearly delineated.

Madden-Julian Oscillation

Gaps in our understanding of the MJO are discussed in detail in Chapter 4. Of primary importance, dynamical models exhibit systematic biases in representing the MJO. These biases could be a manifestation of many processes that are poorly represented and difficult to observe, including coupling between the ocean and atmosphere, shallow and deep convection, cloud microphysics, and cloud-radiation interactions. Additionally, it is still an open question as to how to best organize model experiments and pose model-observation comparisons for identifying model inadequacies. The MJO Working Group, formed under the auspices of U.S. CLIVAR, brought together representatives of the modeling community to address these gaps. This activity has since been extended through the formation of the WCRP-WWPR/THORPEX YOTC MJO Task Force (*see http://www.ucar.edu/yotc*).

Stratosphere

The stratospheric aspects of ISI prediction can only be captured by models that properly simulate stratospheric variability. Thus far, the stratosphere's potential to improve ISI forecasts is largely untapped. To take advantage of this predictability source, it is essential that models used for seasonal forecasting simulate the intense, rapid shifts in the stratospheric circulation, as well as the downward propagation of circulation anomalies through the stratosphere. In addition, models need to be able to simulate the poorly understood connections between lower stratospheric and tropospheric circulations.

To maximize predictability from stratospheric processes, forecasting systems also need to predict stratospheric warmings and other variability at as long a lead time as possible. There are coordinated international experiments underway to examine how stratospheric processes impact ISI forecast quality.

Ocean-atmosphere feedbacks

The two-way interaction between the ocean and the atmosphere plays a very important role in ISI predictability. This interaction can manifest itself as decreased thermal damping, as in the case of the ocean mixed layer response to atmospheric forcing, or as quasi-periodic evolution of upper ocean heat content, as in the case of ENSO. Subgrid-scale processes, such as deep convection and stratus clouds in the atmosphere, or coastal upwelling in the ocean, are important components of this ocean-atmosphere interaction. Our current computational capabilities are insufficient to fully resolve these processes in numerical models, and gaps in our scientific knowledge of these processes limit our ability to parameterize them accurately.

Poor simulation of ocean-atmosphere feedbacks degrades the skill of ISI predictions though systematic as well as random errors. The systematic errors affect the ensemble-mean forecast and also manifest themselves as climate biases in ISI forecast models. Errors in the representation of low-level stratus in the vicinity of the coastal upwelling regions off the coasts of South America and Africa are a well-known example of the climate bias associated with ocean-atmosphere feedbacks. The random errors associated with the feedbacks can lead to incorrect estimation of the spread in an ensemble forecast. In the case of ENSO prediction, random errors associated with the subgrid-scale parameterizations in the atmospheric model are believed to be responsible for the weak ensemble spread in ENSO forecasts.

Cryosphere

Generally, ISI prediction models use climatological sea ice or initialize a sea ice model with climatology. Despite the potential for prediction, the effects of sea ice are poorly included in ISI prediction models. From a simple energetic consistency perspective, active sea ice models that capture the relevant feedbacks need to be included in the ISI models. There is a clear need to identify the remote effects (and causes) of sea ice anomalies and understand associated processes and influence on forecast quality. Land ice (Antarctic and Greenland ice caps, glaciers) is similarly treated very simply, but it is considered to change on timescales too long to affect ISI variability.

Snow in the Northern Hemisphere is a highly variable surface condition in both space and time. Its impacts on the surface energy and water budgets make it a viable candidate for contributing to atmospheric anomalies and ISI predictability. The potential role for knowledge of snow cover anomalies to contribute to forecasts of the Northern Hemisphere winter temperature and East Asian winter monsoon has been mentioned in the Inertia subsection. However, operational models have yet to exploit this source of predictability. Models typically have some initialization and representation of snow cover effects, but the quality and impact of these effects are largely untested.

Soil Moisture

Soil moisture initialization as a source of subseasonal prediction quality is discussed in greater detail in Chapter 4. Although soil moisture can be initialized globally with current Land Data Assimilation System (LDAS, see Chapter 3) frameworks, the accuracy of the initialization could be improved with better measurements of (for example) rainfall and radiation, and it will presumably improve further with the advent of true and operational assimilation of land surface prognostic states. Improved estimates of land parameters (e.g., active soil depth, soil texture) would also help; accuracy in these estimates, which affect the simulation of surface hydrology and thus the surface energy balance, is currently limited by insufficient observational data. Statistical optimization of these parameter values may prove fruitful.

Arguably, in regard to ISI forecasting, the largest "gap" in the soil moisture realm is the degree of uncertainty in the strength of land-atmosphere coupling—our lack of knowledge of the degree to which soil moisture variations in nature affect variations in precipitation and air temperature. Such coupling strength cannot be measured directly with instruments (it can only be inferred indirectly at best), and the estimates of coupling strength quantified with modeling systems vary widely (Koster et al., 2006), indicating a substantial uncertainty in our knowledge of how best to model the relevant underlying physical processes such as evaporation, the structure of the boundary layer, and moist convection. Evaluating these individual components is thus important, but it is currently hindered by data availability. For example, direct joint measurements of evaporation and soil moisture at the hundred-kilometer scale are non-existent, making difficult the evaluation of whether model-generated evaporation fluxes respond realistically to soil moisture variations. Such gaps imply a need for joint model development and observational analysis, focusing on all of the physical processes connecting soil moisture to atmospheric variables.

Non-stationarity

Many statistical models employ assumptions regarding stationarity in the data (i.e., the statistics associated with the predictand do not depend on the time of sampling). However, long term trends associated with global warming and other alterations to the climate system (e.g., changes in land cover) demonstrate that the requirements of stationarity are unlikely to be met for many time series of climate data. In addition to trends, examples of non-stationarity include alterations in the variability of sources of predictability by the annual cycle (i.e., the impact of the annual cycle on ENSO), or by interactions among sources of predictability. Although optimal climate normals (e.g. Lamb and Changnon, 1981; Huang et al., 1996a) offer a valuable statistical tool for dealing with non-stationarity, a better physical understanding is needed regarding the effect of trends on the sources of predictability as well as the ways in which the variability associated with a source may be non-stationary.

METHODOLOGIES USED TO QUANTITATIVELY ESTIMATE PREDICTION SKILL

Prediction skill has been studied in depth since the early 1900s when Finley claimed considerable skill in forecasting tornadoes. Nearly simultaneously, recognition of the limitations of measuring skill surfaced (Murphy, 1991; 1993) resulting in the modern definition of skill as the accuracy of a forecast relative to some known reference forecast, such as climatology or persistence. In Finley's case, his claims of skill were relative to random forecasts.

The skill, or more generally the quality, of forecasts rests on all aspects of the forecast process: the assimilation of observed data for initialization, the completeness and accuracy of the observed data itself, the model(s), and the manner in which the models are employed and interpreted when producing the forecast. These elements and how they come together will be discussed in detail in Chapter 3. Here, the approaches and metrics commonly used to estimate the quality of a prediction or a model are described. In addition, discussion is offered for some newer and lesser-used metrics that could complement existing ones.

No single metric can provide a complete picture of prediction quality, even for a single variable. Thus, a suite of metrics needs to be considered (WMO SVS-LRF, 2002; Jolliffe and Stephenson, 2003), particularly when new models or forecast systems are compared with previous versions. Within any meaningful suite of metrics, one needs to consider the quality of the probabilistic information. The climate system is inherently chaotic, and our ability to measure its initial state is subject to uncertainty; thus any predicted or simulated representation of the climate needs to be probabilistic. While increased accuracy is often the goal of model or forecast improvements, the improved representation of some physical process in a model might not lead to increased accuracy, but it might better quantify the uncertainty in variables influenced by that process.

This section is broken into two subsections: model validation and forecast verification. The forecast verification section addresses the various metrics that are used to assess the quality of climate forecasts, typically for temperature and precipitation. As a good quality model is an obvious prerequisite to a good quality forecast, the section on model validation is presented first. Model validation may incorporate some of the same metrics as used in forecast verification, but it involves more because the model contains the physical processes in addition to the desired prediction fields. Model validation is typically performed prior to using a model for real-time forecasting in order to design statistical corrections for systematic model biases and to better understand the model's variability characteristics, such as accuracy of regional precipitation anomalies under El Niño conditions or the strength and timescale of the MJO.

Model Validation

Comparing a model environment to the observed conditions is a difficult task. Starting with the observed climate, there are numerous scales of variability that cannot be resolved, much less measured at regular intervals. As a result, any comparison would use an incomplete picture of the atmosphere or ocean. The "incompleteness" of the available observations constitutes sampling error. On a practical level, since measurement systems are not homogeneous in space and time, scientists select those variables that are considered most important. The problems become more complicated if a set of observations and model predictions are not available or valid over a common period. No numerical model is perfect, hence model errors are generated. Therefore, the observations and model predictions only can be collected to form two incomplete and imperfect probability density functions (PDFs), which provide a basis for comparison.

An example of several PDFs associated with predictions of SST in the tropical Pacific Ocean is shown in Figure 2.8. The initial prediction is shown in green. Relative to this PDF, the subsequent PDFs (red and blue) exhibit higher probabilities for warmer temperatures. This shift to warmer temperatures was in fact reflected by the verification: the observed temperature anomaly was just below 0.8°C, in between the peaks of the two later PDFs. The standard deviation of the three PDFs is a function of lead-time, although the dependence is perhaps surprisingly small here. The relatively small dependence reflects the relatively large set of available tools and the fact the forecast uncertainty with these tools is large.

A realistic model strives to capture the full variability of the climate system. In particular, such a model needs to capture the full PDF and the temporal and spatial correlations of the observations, even if the forecast information to be disseminated is only a summarized version of that PDF. Additionally, PDFs allow for identification of multimodal distributions, whereas summary statistics (e.g., means, variances, skewnesses) cannot. Several goodness-of-fit tests exist that can check for a significant agreement between the observed and simulated PDFs. When such a test is not practical, the mean and variance should be compared between the observations and model, as a minimum, and the skewness, if possible. Skewness differences can point to processes not being captured, as well as to nonlinearities. Any statistical analysis of these PDFs, particularly ones that attempt to assess skill or significance, will hinge on specific assumptions of the tests applied. A sufficiently large sample size will show the models' depiction of the atmosphere or ocean to differ from that of the observations. For those interested in comparing models to the observed world, this leaves us at an interesting juncture, one that may not have an answer without simplification (Oreskes et al., 1994). Can an experiment be designed to answer the question, "How good is a model or prediction?" Since we do not understand all of the processes and interactions that would lead to a perfect prediction, a strict validation is not possible. Additionally, it is possible that a model prediction can verify correctly

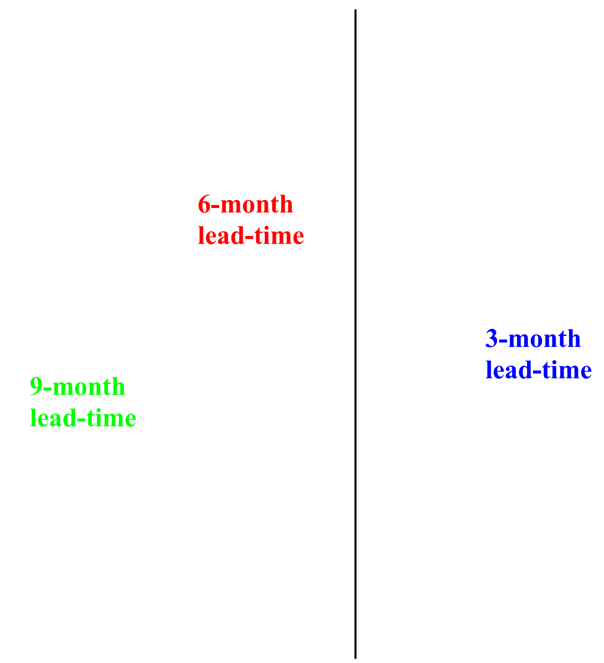

FIGURE 2.8 Examples of probabilistic predictions for Nino3.4 SST anomaly, represented by probability density functions (PDFs). The green curve is the prediction with the longest lead time (9 months), followed by the red (6 months) and blue (3 months) curves. With shorter lead times, the PDF for the prediction shifts to progressively warmer temperatures. The observed value for Nino3.4 SST anomaly is indicated by the vertical black line (0.78°C). SOURCE: International Research Institute for Climate and Society (IRI).

for the incorrect reason. Although a perfect validation is not possible, obtaining a useful comparison is possible if one recognizes the level of uncertainty associated with the observations and model. Moreover, constructing specific hypothesis tests is a viable alternative. One might pose the question as "what aspect of the distribution of the observed atmosphere matches that of the atmosphere simulated by numerical models?"

Statistical Techniques for Identifying and Predicting Modes of Variability

The climate system is characterized by recurrent patterns of variability, sometimes referred to as modes of variability, which include ENSO, NAO, etc. Often, the identification of modes linking remote locations in the atmosphere or the ocean-atmosphere is useful for medium- to long-range prediction (Reynolds et al., 1996; Saravanan et al., 2000). Numerous methodologies have been applied to identify such modes, ranging from linear correlation to multivariate eigentechniques (Montroy et al., 1998) and nonlinear methods (Lu et al., 2009; Richman and Adrianto, 2010). Definitions of these techniques may be found in Wilks (2006) and a summary of their use for mode identification is contained in Appendix A. Often, as the time scale increases, the nonlinear contribution to the modes tends to be filtered. However, Athanasiadis and Ambaum (2009) note that the maintenance and evolution of low frequency variability arise from inherently nonlinear processes, such as transient eddies, on the intraseasonal time scale. This suggests that linear techniques may not fully capture the predictability associated with the modes, and the use of nonlinear techniques needs to be explored (e.g., Kucharski et al., 2010).

Merits of Nonlinear Techniques

To date, nearly all mode identification has been limited to linear analyses. In fact, we have defined our concept of modes through linear correlations and empirical orthogonal functions (EOFs)/principal component analysis (PCA) as those were the techniques that were computationally feasible at the time. Recently, nonlinear mode identification has begun to emerge as efficient nonlinear classification techniques have developed and as computational power has increased. To assess the degree of nonlinearity, the skill of nonlinear techniques can be compared to that derived from traditional linear methods (Tang et al., 2000). Forecasters can investigate if extracting the linear part of the signal is sufficient for prediction. On the intraseasonal time scale, when monsoon variability has been probed by a nonlinear neural network technique (Cavazos et al., 2002), a picture emerges with nonlinear modes related to the nonlinear dynamics embedded in the observed systems (Cavazos et al., 2002) and model physics (Krasnopolsky et al., 2005). Nonlinear counterparts to PCA, such as neural network PCA, have been shown to identify the nonlinear part of the ENSO structure (Monahan and Dai, 2004). By using a nonlinear dimension reduction method that draws on the thermocline structure to predict the onset of ENSO events, Lima et al. (2009) have shown increased skill at longer lead times in when compared to traditional linear techniques, such as EOF and canonical correlation analysis (CCA). Techniques have also been applied to cloud classification (Lee et al., 2004), wind storm modeling (Mercer et al., 2008) and classification of tornado outbreaks (Mercer et al., 2009). Some nonlinear techniques, such as neural networks, are sensitive to noisy data and exhibit a propensity to overfit the data that they are trained on (Manzato, 2005), which can limit their utility in forecasting. Careful quality control of data is essential prior to the application of such methods. To assess the signal that is shared between the training and testing data, some form of cross-validation is typically required (Michaelson, 1987). Techniques include various forms of bootstrapping (Efron and Tibshirani, 1993), permutation tests (Mielke et al., 1981), jackknifing (Jarvis and Stuart, 2001) and n-fold cross validation (Cannon et al., 2002).

Kernel techniques, such as support vector machines, kernel principal components (Richman and Adrianto, 2010), and maximum variance unfolding, avoid the problem of finding a local minimum and overfitting. Kernel techniques have a high potential for mode identification where linear modes provide ambiguous separability (e.g., the overlapping patterns of the Arctic Oscillation and the North Atlantic Oscillation).

Forecast Verification

Finley's tornado forecasts were more skillful than random forecasts, according to the metric he was using, which tabulated the percentage of correct forecasts. It credited forecasts of 'no tornado' on days with no tornadoes, and tornadoes were present on less than 2% of the days.

It turns out that if he had always predicted "no tornado" his skill would have been even greater (Jolliffe and Stephenson, 2003). This example illustrates the value in considering what aspect of forecast quality a metric should measure, and the baseline against which it is assessed. The particular assessment of forecast quality often depends on what characteristics of the forecast are of greatest interest to those who would use the information. No one verification measure can capture all aspects of forecast quality. Some measures are complementary, while others may provide redundant information. This section outlines several recommended (WMO SVS-LRF, 2002) and commonly-used metrics for verifying forecasts, and which aspects of forecast quality they address (see Jolliffe and Stephenson, 2003).

WMO's Standard Verification System (SVS) for Long Range Forecasts (LRF) (2002) outlines specifications for long-range (monthly to seasonal) forecast evaluation and for exchange of verification scores. The SVS-LRF provides recommendations on scores for both deterministic and probabilistic forecasts. For deterministic forecasts, the recommended metrics are the mean square skill score and the relative operating characteristics (ROC; curve and area under the curve; Mason and Graham, 1999). For categorical forecasts the recommended metric is the Gerrity Score (Gerrity, 1992). For probabilistic forecasts, the recommended metrics are the ROC and reliability diagrams. While these metrics are the main ones advocated by the WMO, several others are in regular use by modeling and prediction centers, and still others are being promoted as potentially more interpretable, at least for forecast users. Below, a variety of metrics are discussed and evaluated.

Deterministic Measures

Correlation addresses the question: to what extent are the forecasts varying coherently with the observed variability? Correlation assessments are typically done with anomalies (Figure 2.9a), or deviations from the mean, and can be applied to spatial patterns of variability (pattern correlations) or to time series of variability (temporal correlations). The agreement of co-variability does not indicate if the forecast values are of the right magnitude, and so it is not strictly a measure of accuracy in the forecast.

The mean squared skill score (MSSS) addresses the question: How large are the typical errors in the forecast relative to those implied by the baseline? The baseline could be climatology, for example, assuming the next season's temperature will be that of the average value from the previous 30 years. The MSSS is related to the mean squared error (MSE) and summarizes several contributions to forecast quality, namely correlation, bias, and variance error. The root mean squared error (RMSE; equal to the square root of MSE) is much more widely used than either MSE or MSSS. For a predicted variable whose magnitude is of particular interest, such as the SST index of ENSO, the RMSE may be a preferred metric of forecast quality given its straightforward interpretation. However, RMSE alone is of limited information to the forecast community that wishes to identify the source of the error (Figure 2.9b).

The relative operating characteristics (ROC) addresses the question: can the forecasts discriminate an event from a non-event? The ROC curve effectively plots the hit rate, which is the ratio of correct forecasts to the number of times an event occurred, against the false alarm

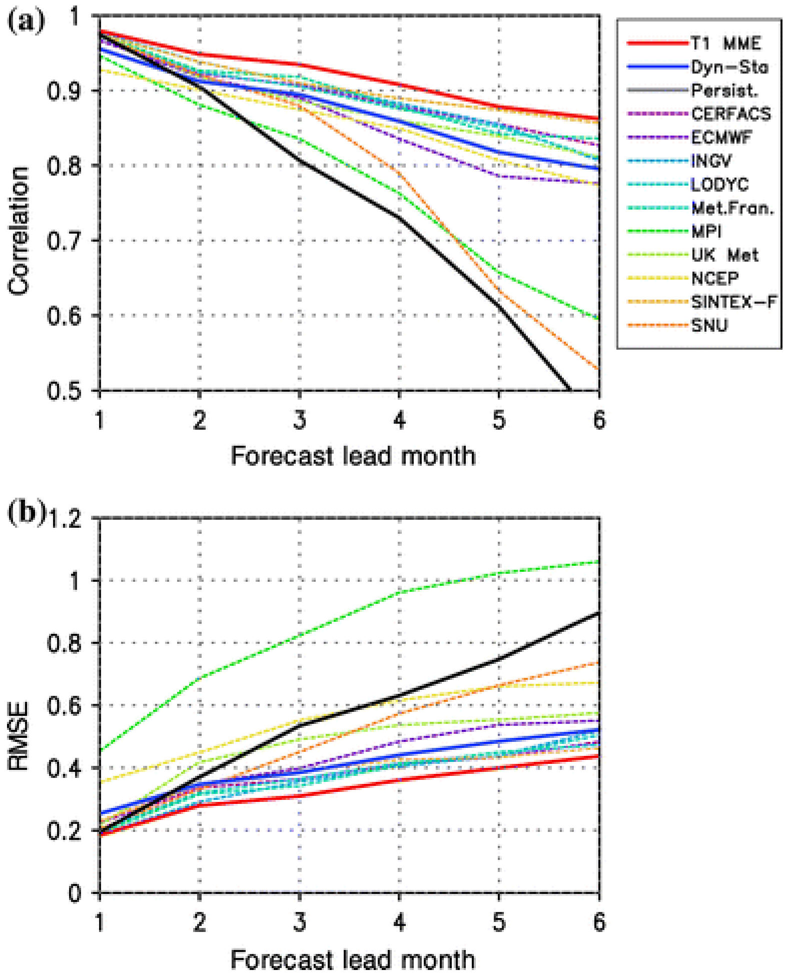

FIGURE 2.9 An example of a multi-model ensemble (MME) outperforming individual models in forecasting. (a) Anomaly correlation; the red line (MME) is above the individual models (colored lines), demonstrating that the pattern of anomalous temperature from the ensemble is a closer match to observations. (b) RMSE of NINO3.4; the red line is below the individual models, demonstrating that the magnitude of the errors associated with the ensemble is smaller. Black represents a persistence forecast. Names of the individual coupled models shown in the legend. SOURCE: Figure 7, Jin et al. (2008)

rate (probability of false detection), which is the ratio of false alarms to total number of non-occurrences. Therefore, one can assess the rate at which the forecast system correctly predicts the occurrence of a specific event (e.g. "above-normal temperature," El Niño conditions, etc.), relative to the rate at which one predicts the occurrence of an event incorrectly (Figure 2.10). If the forecast system has no skill, then the hit rate and false alarm rates are similar and the curve lies along the diagonal in the graph (area = 0.5). Positive forecast skill exists when the curve lies above the diagonal (0.5 < area <= 1.0) and the skill can be measured by the "area under the ROC curve".

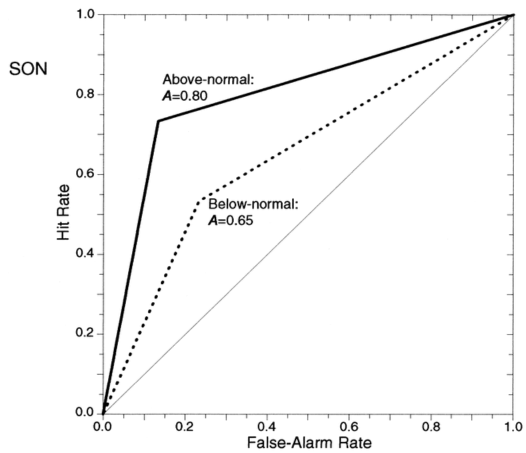

FIGURE 2.10 An example of a ROC curve that plots Hit Rates vs. False-Alarm Rates. In this case, the Hit Rate is the proportion of rainfall events (either above-normal, solid line, or below-normal, dotted-line) that were forecasted correctly; the False-Alarm Rate is the proportion of non-events (i.e., incidences of near normal rainfall) for which an event *was* forecasted. Since the curves are above the gray line, the forecast is considered skillful using this metric. The forecasts correspond to rainfall during September through November in a region in Africa during for the period 1950–1994 The areas beneath the curves, *A*, are indicated also. SOURCE: Figure 2, Mason and Graham (1999).

Probabilistic Measures

The key aspect of probabilistic forecasts is that they proffer quantitative uncertainty associated with the forecast. Thus, if a forecast includes uncertainty, it is important to assess the meaningfulness of that uncertainty; probabilistic forecasts need to be assessed probabilistically. Providing deterministic metrics as well, such as correlation or hit rate of the most likely outcome, may give additional information of use to decision makers, but provided alone, deterministic measures undermine the richness of the forecast information. For example, a deterministic measure such as a hit score based on collapsing the probabilistic forecasts to a deterministic forecast for the category with the largest probability, for purposes of verification, cannot then distinguish between a forecast of 100% likelihood and one of 40% likelihood of above-normal (e.g. for a 3-category system, with climatologically equal odds). However, the reaction of decision makers to such differing confidence in the predicted outcome would certainly be much different.

An important aspect of a quantitative probabilistic measure is that it is equitable. The term equitable means that a forecaster is not penalized for making a forecast that has a low climatological probability (e.g., forecasting a below normal temperature when the climatological probability of a below normal temperature is less than 10 percent).

The potential value of probabilistic assessment of forecasts is large for the model development, forecasting, and decision making communities. While users of the forecast information may stress the desire for "accurate" information, the climate system is inherently probabilistic. The most likely outcome, or equivalently the probabilistic median or the deterministic forecast, may give a general sense of expectations for the seasonal climate, but that information needs to be accompanied by an estimate of the uncertainty. Commercial decisions are often made, not on the basis of events which are likely to occur, but on the basis of events which are unlikely to occur, but which if they did occur, would involve serious financial loss (Palmer, 2002).

The Heidke skill score (HSS), which is actually appropriate to binary forecasts rather than probabilistic forecasts, has been applied in the context of probabilistic categorical forecasts where they have been collapsed into binary categorical forecasts by retaining the category with highest probability. The HSS can be interpreted to addresses the question: did the forecast indicate the correct shift in the probability distribution more often than would be expected by chance? This score may be seen as desirable to some because it is convenient and easily interpreted (Livezey and Timofeyeva, 2008), indicating how often the forecast is "correct" or not (Figure 2.11). However, if it is applied to a probabilistic forecast the HSS degrades the information content as described above.

The Heidke Skill Score is considered biased and may not be equitable. Jolliffe and Stephenson (2003) claim it is equitable for applications involving binary predictions (e.g., yes or no; event or non-event). Wilks (2006) claims it is not equitable for higher-order designs, since the correct forecasts of less likely events do not properly receive more weight (personal communication, Wilks, 2009). Thus, the forecaster may be discouraged from forecasting rare events on the basis of their low climatological probability. The reason for the bias in the Heidke skill is that the reference hit rate in the denominator is not constrained to be unbiased. This means the imagined random reference forecasts in the denominator have a marginal distribution that is not necessarily equal to that of the sample climatology. Peirce Skill (Wilks, 2006) is unbiased and can be substituted for Heidke skill.

The probabilistic ROC is a variant of the ROC described previously that considers the hit rates and false alarm rates for events forecast at varying levels of probabilistic confidence.

The Brier skill score (BSS) is a summary score of forecast quality that encapsulates both reliability and resolution measures of forecast quality. Reliability, discussed more below, addresses the question: to what extent do the probabilities mean what they say? Resolution addresses the question: can the probabilistic forecasts discern changes in the frequency of observed events relative to the underlying climatological distribution? An example of good forecast resolution would be that when forecasts were issued with high probability for an El Niño event, El Niño events were much more likely to happen than would be estimated from their observed frequency over all years.

A reliability diagram shows the complete joint distribution of forecasts and observations for a probabilistic forecast of an event or forecast category (such as the above-normal tercile) (Figure 2.12). They indicate to what degree the probabilities assigned to an event are representative of the likely occurrence of that event. In a reliable forecast system, the probability assigned to a particular outcome should be the frequency with which—given the same forecast—that outcome should be observed. The information supplied by reliability diagrams includes calibration, or what is observed given a specific forecast (e.g., under and overforecasting), as well as resolution and refinement which is the frequency distribution of each of the possible forecasts giving information on the degree of aggregate forecaster confidence (small inset graph in Figure 2.12). Reliability diagrams can further indicate whether there are systematic biases in the forecasts, such as not predicting enough occurrences of above-normal temperatures. Such probabilistic verification, as ROC scores or reliability diagrams, also can be useful for estimating event-specific prediction skill, for example if El Niño events were better predicted than La Niña

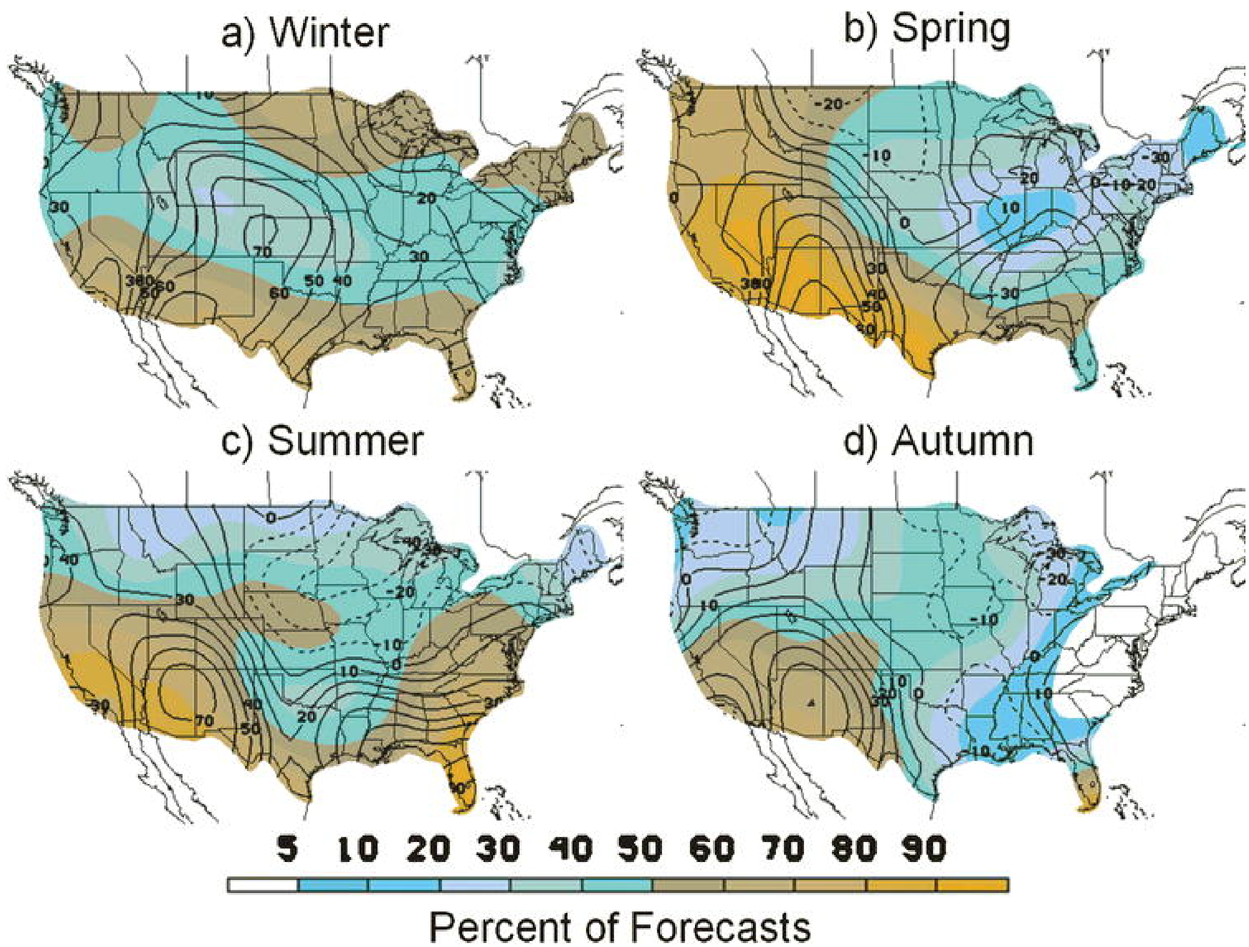

FIGURE 2.11 Seasonal differences in forecast skill (contours) for temperature and how frequently (shading) forecasts differ from climatological odds (i.e., equal chances of normal, above-normal, and below-normal). Blue (tan) corresponds to areas where forecasts are often similar to (often different from) the climatological odds. The skill metric is a Heidke skill score, and is calculated by including only those forecasts that differ from climatological odds. Areas with high-valued contours indicate where deviations from climatology have frequently been forecasted correctly. The forecasts are from CPC and are valid ½ month from issuance. SOURCE: Figure14, O'Lenic et al. (2008).

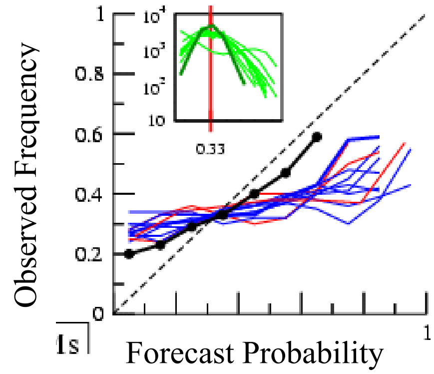

FIGURE 2.12 An example of a reliability diagram, which indicates the skill of probabilistic forecasts. The diagram compares the forecasted probability of an event (in this case, above-normal winter rainfall in North America) to its observed frequency. A perfect forecast is represented by the dashed line, a horizontal line represents a forecast identical to climatology, and sloped lines are potentially skillful. The blue and red lines correspond to individual CGCMs and AGCMs, respectively, and are more horizontal than the black line, which represents the mean of these models. While the mean of the models is more reliable than any of the individual models, it tends to be underconfident for rare events (the black line lies above the perfect forecast line for low-probability events). Typically, a histogram accompanies a reliability diagram (inset), indicating the number of times that forecasts of various confidence levels were issued. SOURCE: Adapted from Goddard and Hoerling (2006).

events or if drought conditions were better predicted than very wet seasons. A distinction in prediction skill between the cases of high and low variability calls for further examination of the physical causes of the discrepancy, and whether it is inherent to the climate system dynamics or a shortcoming of the model(s).

Impacts of Non-Stationarity on Assessment of Skill

This section provides consideration of forecast verification in the context of a changing background climate. Many measures of prediction skill are sensitive to how much the prediction deviates from climatology; therefore, the assessment of seasonal predictions can be influenced by both changes in the drivers of climate predictability as well as trends or other slowly varying changes in the background state. ENSO exerts the greatest influence on seasonal-to-interannual climate variability globally (e.g. Glantz, 1996). As a result, climate predictions made during ENSO events yield much higher skill than those made during ENSO-neutral conditions (e.g. Goddard and Dilley, 2005; Livezey and Timofeyeva, 2008). Although an ENSO event typically occurs every 3–7 years, decadal modulation of the frequency and intensity of ENSO events is evident over the observational record of the 20th century (Zhang et al., 1997) and over the last millennium based on proxy coral data (Cobb et al., 2003). Therefore, there will be periods with higher prediction quality than others merely because there were more or stronger ENSO events during that period. Higher prediction skill will also appear in many metrics (e.g. correlation, Heidke skill score) when the background mean state climate is non-stationary, i.e. presence of trends; the non-stationarity could be due to anthropogenic climate change or natural variability

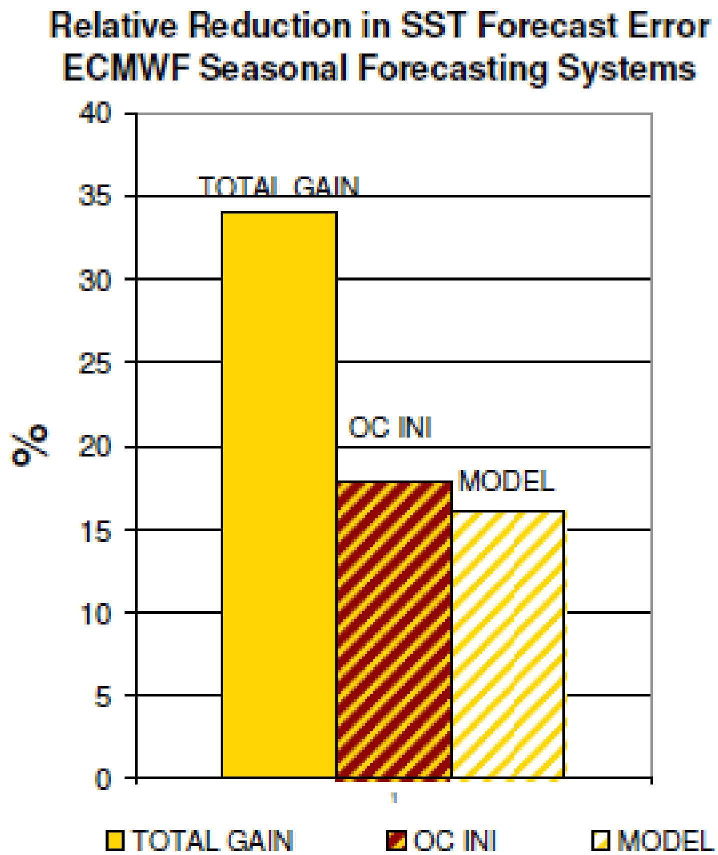

FIGURE 2.13 Progress in the seasonal forecast skill of the ECMWF operational system during the last decade. The solid bar shows the relative reduction in mean absolute error of forecast of SST in the Eastern Pacific (NINO3). The brown-striped bar shows the contribution from the ocean initialization, and the white-striped bar is the contribution from model improvement. SOURCE: Balmaseda et al. 2009.

on long multi-decadal timescales. Climate predictions are typically communicated as deviations from "climatology," or the background mean-state. If the mean-state is changing over time, the magnitude of the seasonal deviation will depend on the period used to define the climatology. Equivalently, predictions of deviations in the same direction as the "trend" can be credited with relatively high quality that is derived more from the slowly evolving trend than the interannual variability. For example, under anthropogenic climate change, the temperatures over most land areas are increasing relative to the mean state, say 1971–2000 (Trenberth et al., 2007). That does not necessarily mean that each year will be warmer than the year preceding it. However, predicting temperatures to be "above-normal," will appear skillful by many measures because temperatures in this decade are very likely to be warmer than those of 30 years prior.

A relevant question is then: can the forecast system discriminate between conditions in a pair of forecasts more often than not? For example, if year X is observed to be warmer than year Y, was that predicted to be so? Discrimination tests of forecast-observation pairs of cases addressing this type of question can be applied to deterministic or probabilistic forecasts. A generalization of such discrimination tests is outlined in Mason and Weigel (2009), and in many cases the metric becomes equivalent to those described above, such as generalized ROC areas for tercile probabilistic forecasts.

CHALLENGES TO IMPROVING PREDICTION SKILL

This chapter has provided the historical perspective on climate prediction, pointed to where there are opportunities to improve prediction quality by improving our understanding and representation in models of sources of predictability, and reviewed the methods available to quantify skill. From the 1980s to the 1990s, seasonal prediction quality improved dramatically, but then did not improve further (Kirtman and Pirani, 2008, 2009). The challenges in going forward are not only to determine where to gain further improvements but also to assess and understand the reasons for any incremental gains in prediction quality that have occurred.

In the following section we examine the building blocks of intraseasonal to interannual forecasting. Improvements may stem from better observations, better models, and improved assimilation. Recent analyses demonstrate how improvements in these components of forecast systems are the source for improvements in forecast quality (Stockdale et al., 2010; Balmaseda et al, 2009; Saha et al., 2006; Fig. 2.13), and thus predictability.

At the same time improvements may result from changes in the way in which the community works. Kirtman and Pirani's (2008, 2009) summary of the first World Climate Research Program Workshop on Seasonal Prediction indicates that that workshop recommended adoption of best practices in seasonal forecasting, including the adoption of common approaches to the production, use, and assessment of seasonal forecasts.

Thus, the challenges to improving intraseasonal to interannual prediction skill lie not only in improvements of the building blocks but also in how the community works together. Experimental modeling and examination of the incremental skill to be gained from new sources of predictability are needed. The three case studies provide examples of physical processes being examined as sources of predictability. A further challenge is to develop the community framework to nurture ongoing improvements to dynamical models.

3

Building Blocks of Intraseasonal to Interannual Forecasting

An ISI forecast is made utilizing observations of the climate system, statistical and/or dynamical models, data assimilation schemes, and in some cases the subjective intervention of the forecaster (see Box 3.1). Improvements in each of these components, or in how one component relates to another (e.g., data assimilation schemes expanded to include new sets of observations; observations made as part of a process study to validate or improve parameters in a dynamical model), can lead to increases in forecast quality. This portion of the report discusses these components of ISI forecasting systems, with an emphasis on assessing quality among forecast systems following a change in forecast inputs. Past advances that have contributed to improvements in forecast quality are noted, and the section ends by presenting areas in which further improvement could be realized.

HISTORICAL PERSPECTIVE FOR INTRASEASONAL TO INTERANNUAL FORECASTING

Scientific weather prediction originated in the 1930s, with the objective of extending forecasts as far into the future as possible. Studies at MIT under Carl Gustaf Rossby consequently included longer time scales than just the daily prediction issue. Jerome Namias became a protégé of Rossby, and took on the task of extending to longer scales as director of the "Extended Forecast Section" of the Weather Bureau/National Weather Service. The approaches developed emphasized upper level pressure patterns that could persist or move according to the Rossby barotropic model, and could provide "teleconnections" from one region to another. These patterns were then used to infer surface temperature and precipitation patterns. The latter were initially done by subjective methods, but soon statistical approaches were adopted through the work of Klein. For more than a few days in advance, prediction of daily weather would necessarily have low skill and so monthly or longer forecasts were obtained as averages. Work by Lorenz in the 1960s explained the lack of atmospheric predictability after more than about 10 days in terms of the chaotic nature of the underlying dynamics (see Chapter 2). At about the same time, Namias was emphasizing the need to consider underlying anomalous boundary conditions as provided by SSTs, soil moisture, and snow cover. The importance of changing tropical SSTs through ENSO was first identified by Bjerknes in the late 1960s. A first 90-day seasonal outlook was released by NOAA in 1974.

BOX 3.1
TERMINOLOGY FOR FORECAST SYSTEMS

Observation—measurement of a climate variable (e.g., temperature, wind speed). Observations are made *in situ* or remotely. Many remote observations are made from satellite-based instruments.

Statistical model—a model that has been mathematically fitted to observations of the climate system using random variables and their probability density functions.

Dynamical or Numerical model—a model that is based, primarily, on physical equations of motion, energy conservation, and equation(s) of state. Such models start from some initial state and evolve in time by updating the system according to physical equations.

Data assimilation—the process of combining predictions of the system with observations to obtain a best estimate of the state of the system. This state, known as an "analysis", is used as initial conditions in the next numerical prediction of the system.

Operational forecasting—the process of issuing forecasts in real time, prior to the target period, on a fixed, regular schedule by a national meteorological and/or hydrological service.

Initial conditions/Initialization—Initial conditions are estimations of the state (usually based on observational estimates and/or data assimilation systems) that are used to start or initialize a forecast system. Initialization can include additional modification of the initial conditions to best suit the particular forecast system.

Progress since the 1960s can be discussed in terms of advances in forecasting approaches (including their evaluation) and improved understanding and treatment of underlying mechanisms. One major direction of advancement in forecasting has been that of dynamical modeling (see "Dynamical Models"section in this chapter). Generally the dynamical models continued to improve according to advancements in computational resources and a growing knowledge of the key processes to be modeled. However, official forecasts in the United States depended on subjective interpretation of these objective products. In addition, various statistical (empirical) modeling approaches were developed and improved to remain as capable as the dynamical approaches in their validation. Other countries have been developing similar capabilities for seasonal prediction since the 1980s, largely depending on numerical modeling.

Recognition of the role of tropical SST anomalies, especially those associated with ENSO, in driving remote climate anomalies has led to much work in predicting tropical SST. Some of the key advancements in estimating these SSTs developed during the TOGA international study in conjunction with the deployment of the Tropical Atmosphere Ocean (TAO) array in the 1980s and 1990s (NRC, 1996; see "Ocean Observations" section in this chapter and "ENSO" section in Chapter 4).

Further expansion of the efforts in ISI forecasting have been undertaken by CLIVAR (Climate Variability and Predictability), a research program administered by the World Climate

Research Programme (WCRP). CLIVAR supports a variety of research programs[9] around the world focused on cross-cutting technical and scientific challenges related to climate variability and prediction on a wide range of time scales. CLIVAR also helps to coordinate regional, process-oriented studies (WCRP, 2010).

What follows is a description of the "building blocks" of an ISI forecasting system: observations, statistical and numerical models, data assimilation schemes. The quality and use of forecasts are also discussed. It is a broad overview, offering some historical context, an evaluation of strengths and weaknesses, and potential avenues for improvement. At the conclusion of Chapter 3, the key potential improvements are summarized; the Recommendations (Chapter 6) have been made with these improvements in mind.

OBSERVATIONS

Observations are an essential starting point for climate prediction. In contrast to weather prediction, which focuses primarily on atmospheric observations, ISI prediction requires information about the atmosphere, ocean, land surface, and cryosphere. Also in contrast to weather prediction, the observational basis for ISI prediction is both less mature and less certain to persist. Indeed, both continuing evolution and the need to sustain observations for ISI prediction are seen as issues at present and into the future. International cooperation and the governance of the World Meteorological Organization do much to ensure continuity of weather observations. Similar international cooperation is being developed for climate observations, but formal international commitments to these observations are not the general case. The following sections describe some of the platforms available for making these observations, and the increase in the number of observations over time.

Observations of quantities that end-users track, such as sea surface temperature and precipitation, and of quantities that record the coupling between elements of the climate system, such as soil moisture and air-sea fluxes, are particularly useful to assess both the realism of models and identify longer-term variability and trends that provide the context for ISI variability. However, current observational systems do not meet all ISI prediction needs, or are not always used to maximum benefit by ISI prediction systems. Some observations for the Earth system needed for initialization are not being taken, or are not available at a spatial or temporal resolution to make them useful. Some observations have not been available for a sufficiently long period of time to permit experimentation, validation, verification, and inclusion within statistical or dynamical models. In yet other cases, the observations are available, but they are not being included in data assimilation schemes. Additionally, regionally enhanced observations or studies that target learning more about the processes that govern ISI dynamics, including developing improved parameterizations of processes that are sub-grid scale in dynamical models, are needed.

New observations, both *in situ* and remotely sensed, may be available through research programs. Part of the challenge is to integrate these new observations, assess their utility and impact, and then, if the observations contribute to ISI prediction, develop the advocacy required to sustain them. Integration of observational efforts, as in CLIVAR climate process teams or by

[9] Programmatic evaluation of the U.S. CLIVAR project office can be found in NRC (2004). Historical strategic recommendations germane to ISI forecasting for CLIVAR's Global Ocean-Atmospehre-Land program (GOALS) can be found in NRC (1998).

bringing together observationalists with operational centers to engage in observing system simulation studies and assessments of the improvement stemming from observations have merit. Heterogeneous networks of observations, at times obtained by different organizations, may need better integration into accessible data bases and particular attention from partnered observationalists and modelers. For all observations, appropriate attention to metadata and data quality, including realistic estimates of uncertainty, are essential to ensuring their use and utility.

Atmosphere

Over the years the conventional meteorological observing system evolved from about 1,000 daily surface pressure observations in 1900 to about 100,000 at the present. Likewise, upper air observations (rawinsondes, pilot balloons, etc.) grew from less than 50 soundings in the 1920s to about 1,000 in the 1950s (of which most were pilot balloons). Today, there are about 1,000 rawinsondes used regularly (Dick Dee, personal communication). Satellite observations, introduced into operations in 1979, ushered in a totally new era of numerical weather prediction, although it was only in the 1990's that the science of data assimilation (see "Data Assimilation" section in this chapter) progressed enough to demonstrate that there was a clear positive impact from satellite data when added to rawinsondes in the Northern Hemisphere. Figure 3.1 illustrates the huge increase of different types of available satellite observations in the last two decades assimilated for ECMWF operational forecasts. These satellite products have not only grown in number, but also in diversity. They can provide information about atmospheric composition and hydrometeors, as well as vertical profiles of thermodynamic properties.

The assimilation of each of these observing systems poses a new challenge, and the full impact of each may not become clear for years because of the partial duplication of information among the different systems. It is often difficult to attribute an increase in prediction quality to the incorporation of a new set of observations in an ISI forecasting system. Some examples of improvements arising from the assimilation of specific observations, such as AMSU radiances, are discussed in the "Data Assimilation" section of this chapter.

The incorporation of targeted observations that focus on atmospheric processes that are sources of ISI predictability could also contribute to ISI forecast quality. In some cases, these observations exist for research purposes but are not being exploited by ISI forecast systems. In other cases, these observations do not exist. For example, high resolution observations of the vertical structure of the tropical atmosphere could improve the understanding of the MJO, the ability to validate current dynamical models, and perhaps the parameterization of these models. This is part of the mission of the Dynamics of the MJO experiment (DYNAMO; *http://www.eol.ucar.edu/projects/dynamo/documents/WP_latest.pdf*).

Oceans

As mentioned in Chapter 2, the oceans are a major source of predictability at intraseasonal to interannual timescales. The ocean provides a boundary for the atmosphere where heat, freshwater, momentum, and chemical constituents are exchanged. Large heat losses and evaporation at the sea surface cause convection and make surface water sink into the interior,

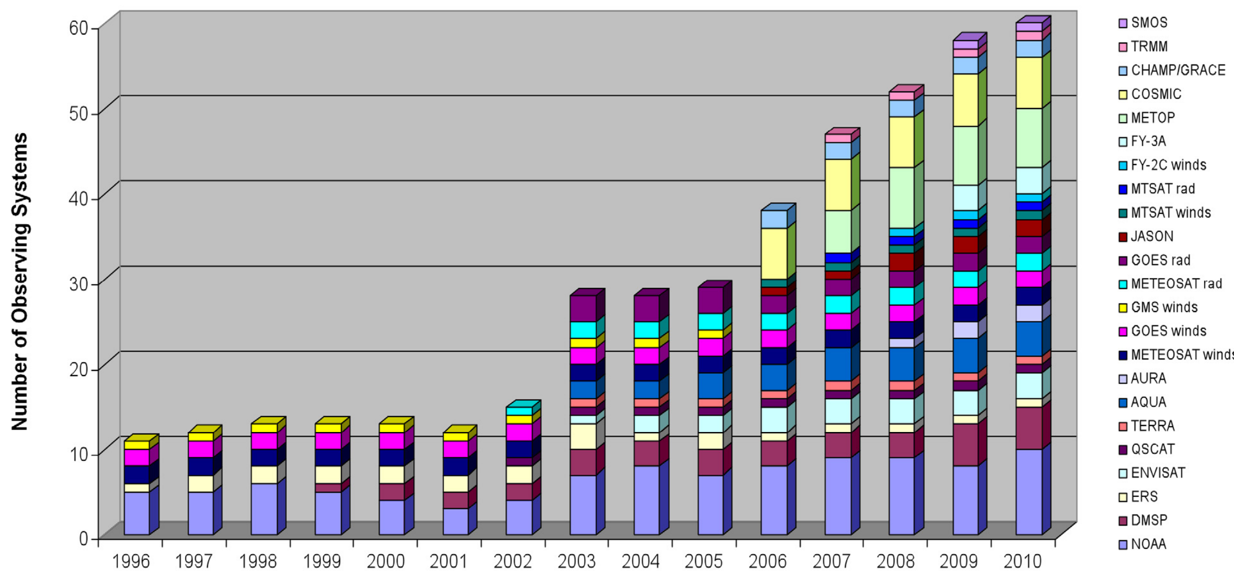

FIGURE 3.1: Number of satellite observing systems available since 1989 and assimilated into the ECMWF system. Each color represents a different source/platform. SOURCE: Courtesy of Jean-Noel Thepaut, ECMWF.

while surface heating and the addition of freshwater make surface water buoyant and resistant to mixing with deeper water.

Variability in the air-sea fluxes, in oceanic currents and their transports, and in large-scale propagating oceanic Rossby or Kelvin waves all contribute to the dynamics of the upper ocean and the sea surface temperature. In turn, the states of the ocean surface and sub-surface can force the atmosphere on intraseasonal to interannual timescales, as is clearly evident in the ENSO and MJO phenomena. Therefore, the initialization of sea surface and sub-surface ocean state is required for near-term climate prediction. Unfortunately, the comprehensive observation of the global oceans started much later than in the atmosphere and even today there are challenges that prevent collection of routine observations over large parts of the ocean.

The significant climatic impacts of ENSO, especially after the 1982–1983 event, demonstrated that a sustained, systematic, and comprehensive set of observations over the equatorial Pacific basin was needed. The TAO/Triangle Trans-Ocean Buoy Network (TRITON) array was developed during the 1985–1994 Tropical Ocean Global Atmosphere (TOGA) program (Hayes et al., 1991, McPhaden et al., 1998). The array spans one-third of the circumference of the globe at the equator and consists of 67 surface moorings plus five subsurface moorings. It was fully in place to capture the evolution of the 1997–1998 El Niño. In 2000, the original set called TAO was renamed TAO/TRITON with the introduction of the TRITON moorings at 12 locations in the western Pacific (McPhaden et al., 2001). TAO/TRITON has been the dominant source of upper ocean temperature and *in situ* surface wind data near the equator in the Pacific over the past 25 years and has provided the observational underpinning for theoretical explanations of ENSO such as the recharged oscillator (e.g. Jin, 1997). It provides a key constraint on initial conditions for seasonal forecasting at many centers around the world.

After the success of the TAO/TRITON array, further moored buoy observing systems have been developed over the Atlantic (PIRATA) and Indian (RAMA) oceans under the Global Tropical Moored Buoy Array (GTMBA) program. The moorings allow simultaneous observations of surface meteorology, the air-sea exchanges of heat, freshwater, and momentum, and the vertical structure of temperature, salinity, horizontal velocity, and other variables in the water column. Thus, they provide the means to monitor both the air-sea exchanges and the storage capacity of the upper ocean. The PIRATA array was designed for the purpose of improving the understanding of ocean-atmosphere interactions that affect the regional patterns of climate variability in the tropical Atlantic basin (Servain et al., 1998). The array, launched in 1997 and still being extended, currently has 17 permanent sites. The RAMA array was initiated in 2004 with the aim of improving our understanding of the east Africa, Asian, and Australian monsoon systems (McPhaden et al., 2009). It currently consists of 46 moorings spanning the width of the Indian Ocean between 15°N and 26°S. It is expected to be fully completed in 2012.

The maintenance of the GTMBA is absolutely essential for supporting climate forecasting. However, there are many difficulties in maintaining these arrays, not the least of which is identifying institutional arrangements that can sustain the cost of these observing systems (McPhaden et al., 2010). Away from the equator, the permanent *in situ* moored arrays are sparser and address sample the characteristic extra-tropical regions of the ocean-atmosphere system under the international OceanSITES program. Few such sites exist in high latitude locations, but efforts are underway in the United States (under the National Science Foundation Ocean Observatories Initiative) and in other countries to add sustained high latitude ocean observing capability.

In parallel to the development of the moored buoy arrays, the observation of SST has improved markedly over the last 20 years. SST is a fundamental variable for understanding the complex interactions between atmosphere and ocean. Since 1981, operational streams of satellite SST measurements have been put together with *in situ* measurements to form the modern SST observing systems (Donlon et al., 2009). Since 1999 more than 30 satellite missions capable of measuring SST in a variety of orbits (polar, low inclination, and geostationary) have been launched with infrared or passive microwave retrieval capabilities. New approaches to integrate remote sensing observations with *in situ* SST observations that help reduce bias errors are being taken (Zhang et al., 2009).

Despite the evident progress, an important issue remains: satellite observations of SST only started in the 1980s and satellites have a relatively short life span. Therefore, further work is necessary to ensure the “climate quality” of the data over long periods. This would facilitate the generation of SST re-analysis products for operational seasonal forecasting (Donlon et al., 2009).

Even with the evident progress made with the tropical moored buoy arrays and the improvement of the satellite measurements of SST, as recently as the late 1990s there were still vast gaps in observations of the subsurface ocean. Such observations are needed for seasonal to interannual prediction. The ability of the ocean to provide heat to the atmosphere, the extent to which the upper ocean can be perturbed by the surface forcing, and the dynamics of the ocean that lead to changes in the distribution of heat and freshwater all depend on the vertical and horizontal structure of the ocean and its currents.

Surface height observations by satellite altimeters have added information about the density field in the ocean and thus, for example, the redistribution of water properties and mass along the equator associated with ENSO. Efforts to quantify the state of the ocean were further improved by the international implementation of the Argo profiling float program (*http://www.argo.ucsd.edu/*). Until then, most sub-surface ocean measurements were taken by expendable bathythermograph (XBT) probes measuring temperature versus depth and by shipboard profiling of salinity and temperature versus depth from research vessels, which are both limited in their global spatial coverage and depth. The Argo program was initiated in 1999 with the aim of deploying a large array of profiling floats measuring temperature and salinity to 1,500 to 2,000 meters deep and reporting in real time every 10 days. To achieve a 3° x 3° global spacing, 3,300 floats were required between 60°S and 60°N. As of February 2009, there are 3,325 active floats in the Argo array. After excluding floats from which data was not passing the quality control and those in high latitudes (beyond 60° latitude) or in heavily sampled marginal seas, the number of floats is only 2,600. Argo data is distributed via the internet without restriction and about 90% of the profiles are available within 24 hours of acquisition. Quality control continues after receipt of the data, particularly for the salinity observations. To improve the quality of data from Argo floats, ship-based hydrographic surveys obtaining salinity and temperature profiles are needed, and the process may require several years before the Argo data experts are confident that the best data quality has been achieved (Freeland et al., 2009). However, the real time Argo data is a critical contribution. With it, the depth of the surface mixed layer can be mapped globally, thus determining the magnitude (depth and temperature) of the oceanic thermal reservoir in immediate contact with the atmosphere.

Internationally, there is a coordinated effort under the Joint Commission on Oceanography and Marine Meteorology (JCOMM) of the World Meteorological Organization (WMO) and the International Oceanographic Commission (IOC) to coordinate sustained global ocean *in situ* observations, including Argo floats, surface SST drifters, Volunteer Observing Ship-based measurements, tropical moored arrays, and the extra-tropical moored buoys. Remote observations of surface vector winds combined with drifting buoy data can be used to identify the wind-driven flow of the upper ocean, thus complementing the ability of Argo floats and altimetry to observe the density-driven flow. Future satellite observations of interest include those of surface salinity.

The *in situ* ocean observing community will benefit from an ongoing dialog with those interested in improving prediction on intraseasonal and interannual timescales. Programs such as the World Climate Research Programme's (WCRP) CLIVAR work to coordinate sustained observations in the ocean with focused process studies that improve understanding of climate phenomena and processes. Distributed and sustained ocean and air-sea heat flux observations with global and full depth coverage are being used to identify biases and errors in coupled and ocean models. These include the surface buoys and associated moorings of OceanSITES and the repeat hydrographic survey lines done in each basin every 5–10 years. The moorings provide high temporal resolution sampling from the air-sea interface to the seafloor, while the surveys map ocean properties along basin-wide sections. Both programs provide data sets that quantify the structure and variability of the ocean that are often found in model fields. In contrast, denser sampling arrays are deployed for a limited duration as part of process studies. These studies are

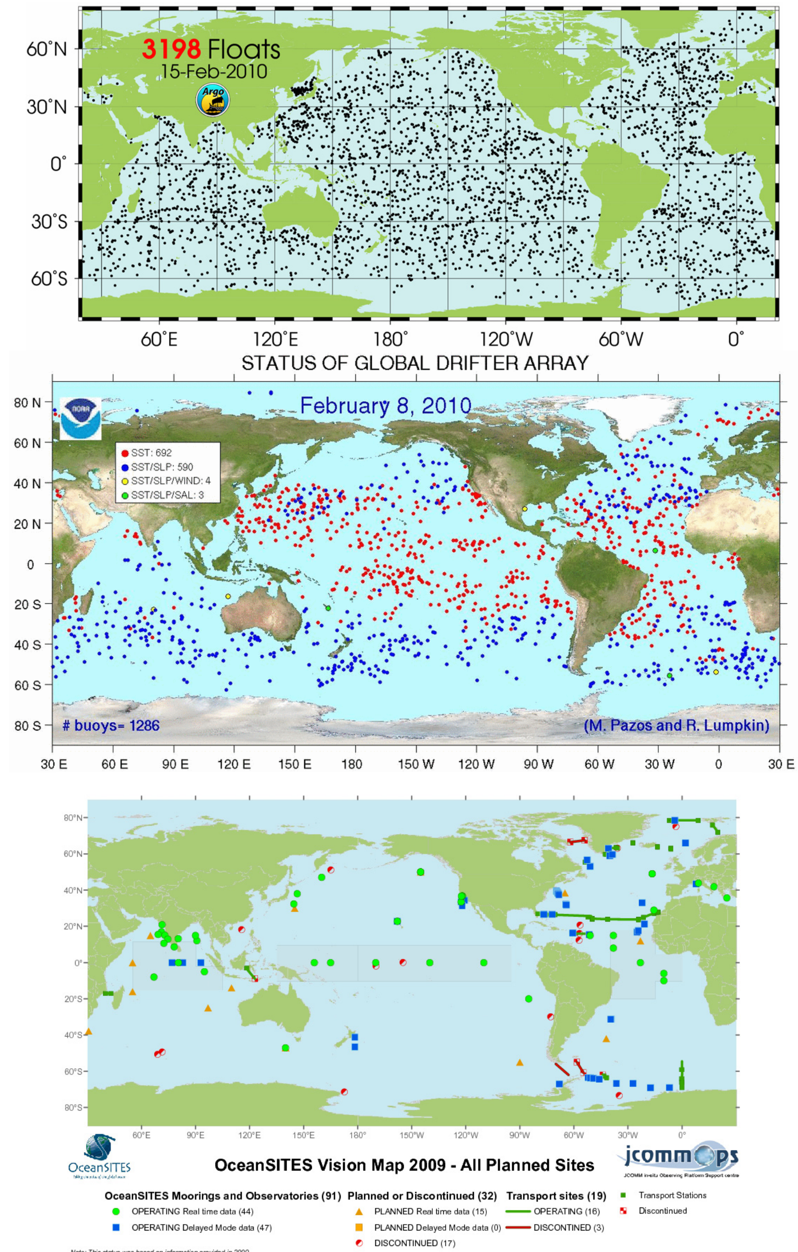

FIGURE 3.2 Examples of the spatial distribution of various ocean observations mentioned in the text. Top panel: Argo floats, which can provide surface and sub-surface information. SOURCE: Argo website (*http://www.argo.ucsd.edu/*) Middle panel: Drifters, which can provide SST, SLP, wind, and salinity information (see colors in legend). SOURCE: NOAA (*http://www.aoml.noaa.gov/phod/dac/gdp.html*). Bottom panel: OceanSITES, intended for long-term observations for depths up to 5000m in a stationary location. SOURCE: (OceanSITES *http://www.jcommops.org/FTPRoot/OceanSITES/maps/200908_VISION.pdf*)

designed to improve our understanding of physical processes and to aid in the parameterization of the processes not fully resolved by models. CLIVAR also works to build connectivity among the observing community, researchers investigating ocean processes and dynamics, and climate modelers. Process studies by CLIVAR and others add to understanding of ocean dynamics, develop improved parameterizations of processes not resolved in ocean models, and guide longer term investments in ocean observing.

Land

The land variables of potential relevance for seasonal prediction—the variables for which accurate initialization may prove fruitful—are soil moisture, snow, vegetation structure, water table depth, and land heat content. These variables help determine fluxes of heat and moisture between the land and the atmosphere on large scales and thus may contribute to ISI forecasts. In addition, some of these variables are associated with local hydrology and hydrological prediction (e.g., observations of snow in a mountain watershed in the winter can provide information on spring water supply). This evolution in the use of land and hydrological observations mirrors the emerging interest in new types of ocean observations, noted in the previous section.

Despite their importance to the surface energy and moisture balances and fluxes, our ability to measure such land variables on a global scale is extremely limited. Thus, alternative approaches for their global estimation have been, or still have to be, developed.

Soil Moisture

Of the listed land variables, soil moisture (perhaps along with snow) is probably the most important for subseasonal to seasonal prediction. For the prediction problem, however, direct measurements of soil moisture are limited in three important ways. First, each *in situ* soil moisture measurement is a highly localized measurement and is not representative of the mean soil moisture across the spatial scale considered by a model used for seasonal forecasting. Second, even if a local measurement was representative of a model's spatial grid scale, the global coverage of existing measurement sites would constitute only a small fraction of the Earth's land area, with most sites limited to parts of Asia and small regions in North America. Finally, even if the spatial coverage were suddenly made complete, the temporal coverage would still be lacking; long historical time series (decadal or longer) may be needed to interpret a measurement properly before using it in a model.

Satellite retrievals offer the promise of global soil moisture data at non-local scales. Data from the Scanning Multichannel Microwave Radiometer (SMMR) and Advanced Microwave Scanning Radiometer—Earth Observing System (AMSR-E) instruments, for example, have been processed into global soil moisture fields (Owe et al., 2001; Njoku et al., 2003). Figure 3.3 shows an example of the mean soil moisture as observed by the SMMR instrument. Such instruments, however, can only capture soil moisture information in the top few millimeters of soil, whereas the soil moisture of relevance for seasonal prediction extends much deeper, through the root zone (perhaps a meter). The usefulness of satellite soil moisture retrievals or their associated raw radiances will likely increase in the future as L-Band measurements come online

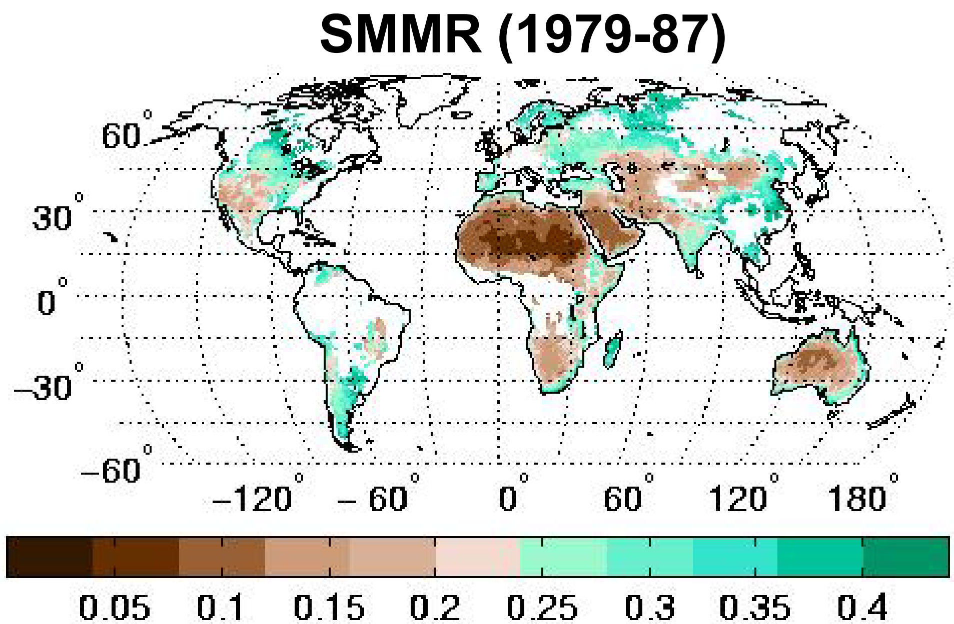

FIGURE 3.3 Mean soil moisture (m^3/m^3) in upper several millimeters of soil, as estimated via satellite with the SMMR instrument using the Owe et al. (2001) algorithm. SOURCE: Adapted from Reichle et al. (2007).

and data assimilation methods are further developed (see section on "Data Assimilation" in this chapter and the soil moisture case study in Chapter 4).

Currently, global soil moisture information for model initialization has to be derived indirectly from other sources. A common approach is to utilize the soil moisture produced by the atmospheric analysis already being used to generate the atmospheric initial conditions. This approach has the advantage of convenience, and the soil moisture conditions that are produced reflect reasonable histories of atmospheric forcing, as generated during the analysis integrations—if the analysis says that May is a relatively rainy month, then the June 1 soil moisture conditions produced will be correspondingly wet.

The main meteorological driver of soil moisture, however, is precipitation, and analysis-based precipitation estimates are far from perfect. Thus, a more careful approach to using model integrations to generate soil moisture initial conditions has been developed in recent years. This approach is commonly referred to as LDAS, for "Land Data Assimilation System", although the term is something of a misnomer; true land data assimilation in the context of the land initialization problem is discussed further in the "Data Assimilation" section below. LDAS systems are currently in use for some experimental real-time seasonal forecasts and are planned for imminent use in some official, operational seasonal forecasts.

An operational LDAS system produces real-time estimates of soil moisture by forcing a global array of land model elements offline (i.e., disconnected from the host atmospheric model) with real-time observations of meteorological forcing. (Here, real-time may mean several days to a week prior to the start of the forecast, to allow time for processing.) Real-time atmospheric data assimilation systems are the only reasonable global-scale sources for such forcings as wind speed, air temperature, and humidity. However, the evolution of the soil moisture state depends even more on precipitation and net radiation, whose reanalysis estimates are not reliable.

Consequently, LDAS systems use alternative sources such as merged satellite-gauge precipitation products (e.g., CMAP, or the Climate Prediction Center Merged Analysis of Precipitation) and satellite-based radiation products (e.g., AFWA AGRMET, or Air Force Weather Agency Agricultural Meteorology Modeling System). The LDAS system may still need atmospheric analysis data for the sub-diurnal time sequencing of the forcing, but the alternative data sources prove invaluable for "correcting" these precipitation and radiation time series so that their temporal-averages are realistic.

Such LDAS systems also require global distributions of surface parameters (vegetation type, soil type, etc.), currently available in various forms (e.g., Rodell et al., 2004). Consistency between the parameter set used for the LDAS system and that used for the full forecast system is an important consideration.

Snow

Real-time direct measurements of snow on the global scale do not exist, though some measurements are available at specific sites, for example, in the western United States (Snowpack Telemetry, SNOTEL) and through coded synoptic measurements made at weather stations (SYNOP). For global data coverage, satellite measurements are promising—certain instruments (e.g., MODIS) can estimate snow cover accurately at high resolution on a global scale. Satellite snow retrievals, however, also show significant limitations. For the seasonal forecasting problem, snow cover is not as important as snow water equivalent (SWE), which is the amount of water that would be produced if the snowpack were completely melted. Satellite estimates of SWE are made difficult by the sensitivity of the retrieved radiances to the morphology (crystalline structure) of the snow, which is almost impossible to estimate *a priori*—a given snowpack may have numerous vertical layers with different crystalline structures, reflecting the evolution of the snowpack with time through compaction and melt/refreeze processes. Compounding the difficulty of estimating SWE from space are spatial heterogeneities in snowpack associated with topography and vegetation.

The LDAS approach described above can provide SWE in addition to soil moisture states, assuming the land model used employs an adequate treatment of snow physics. In the future, the merging of LDAS products with the available *in situ* snow depth information and satellite-based snow data in the context of true data assimilation (see "Data Assimilation" section) will likely provide the best global snow initialization for operational forecasts.

Vegetation Structure

Current operational seasonal forecast models treat vegetation as a boundary condition, with prescribed time-invariant vegetation distributions and (often) prescribed seasonal cycles of vegetation phenology, e.g., leaf area index (LAI), greenness fraction, and root distributions. Early forecast systems relied on surface surveys of these quantities, and modern ones generally rely on satellite-based estimates.

Reliable dynamic vegetation modules would, for the seasonal prediction problem, allow the initialization and subsequent evolution of phenological prognostic variables such as LAI and rooting structure. A drought stressed region, for example, might be initialized with less leafy trees, with subsequent impacts on surface evapotranspiration, and the leaf deficit would only recover if the forecast brought the climate into a wetter regime. However, the use of dynamic vegetation models in seasonal forecasts is not on the immediate horizon for forecast centers, in light of other priorities and the need to develop these models further.

Water Table Depth

The use of water table depth information in historical and current operational systems is prevented by two things. First, outside of a handful of well-instrumented sites, such information does not exist (though GRACE satellite measurements of gravity anomalies can provide useful information at large scales); the global initialization of water table depth given current measurement programs is currently untenable. Second, even if such observations were available, land surface models used in current seasonal forecasting systems do not model variations in moisture deeper than a few meters below the surface, so that the observations, if they did exist, could not be used. The lack of deep water table variables in the models also prevents the estimation of water table depth through the LDAS approach. Given the long time scales associated with the water table, improvements in its measurement and modeling do have the potential to contribute to ISI prediction.

Soil Heat Content

Real-time *in situ* measurements of subsurface heat content are spotty at best and far from adequate for the initialization of a global-scale forecast system. Satellite data have limited penetration depth; they can only provide estimates of surface skin temperature. Global initialization of subsurface heat content can thus be accomplished in only two ways: (1) through an LDAS system, as described above, and (2) through a land data assimilation approach that combines the LDAS system information with observations of variables such as soil moisture, snow, and skin temperature. For maximum effectiveness, the land models utilized in these systems need to include temperature state variables representing at least the depth of the annual temperature cycle (i.e., a few meters).

Polar Ice

Polar regions are important components of the climate system. The most important parameters are those that influence the exchange of heat, mass, and momentum with the atmosphere and global oceans. NASA, NOAA, and DOE have polar-orbiting satellites that are collecting relevant data in the Arctic region. The National Snow and Ice Data Center (*http://nsidc.org/*) is supported by NOAA, NSF, and NASA to manage and distribute cryosphere data. The National Ice Center (*http://www.natice.noaa.gov/*) is funded by the Navy, NOAA, and the Coast Guard to provide snow, ice (ice extent, ice edge location), and iceberg products in the Arctic and in the Antarctic.

The NRC Report *Towards an Integrated Arctic Observing System* (NRC, 2006) advocated observation of “Key Variables” using *in situ* and remote sensing methods. These include albedo; elevation/bathymetry; ice thickness, extent, and concentration; precipitation; radiation; salinity; snow depth/water equivalent; soil moisture; temperature; velocity; humidity; freshwater flux; lake level; sea level; aerosol concentration; and land cover. Observing methods and recommendations are reviewed and presented in the Cryosphere Theme Report of the Integrated Global Observing System (IGOS 2007). Their cryospheric plan includes satellite-based, ground-based, and aircraft-based observations together with data management and modeling, assimilation, and reanalysis systems.

In terms of monitoring climate variability and change and weather and climate prediction, these reports identify the priority cryosphere observations as: long-term consistent records of cryosphere variables, high spatial and temporal resolution fields of snowfall, snow water equivalent, snow depth, albedo and temperature, and mapping of permafrost and frozen soil, lake and sea ice characteristics. Remote sensing methods can be used to address sea ice extent (recorded since the late 1970s) and ice thickness (recorded more recently, with IceSat) in order to investigate ice mass balance and the movement. Aerial sea ice reconnaissance needs to continue. Relevant *in situ* methods include the use of autonomous underwater vehicles (AUVs), moorings, and automated weather stations.

More specific recommendations are provided by Dickson et al. (2009), who assert that climate models do not represent Arctic processes well, limiting our ability to understand change in the Arctic seas and the impact of that change on climate. They also advocate that observations are needed for the Norwegian Atlantic current transport of heat, salt, and mass into the Arctic Ocean and of the amounts that enter by the Fram Strait and the Barents Sea; the change in sea ice in response to inflows into the Arctic of warmer water; the change and variability in temperature and salinity profiles beneath the ice, as by Ice-Tethered Profilers (ITPs); and, in general, all quantities relevant to the estimation of ocean-atmosphere heat exchange in the Arctic. For improving sea ice prediction, Dickson et al. (2009) recommends improved sea ice thickness measurements, especially in the spring. Such improved measurements could be obtained from below and above the ice as well as on the ice, using (for example) laser and radar altimetry, tiltmeter buoys on the ice surface, and floats or moorings below the ice with upward looking sonars.

STATISTICAL MODELS

Statistical and dynamical predictions are complementary. Advances in statistical prediction are often associated with enhanced understanding, which may lead to improved dynamic prediction, and vice versa. In addition, both techniques can ultimately be combined to provide better guidance for decision support.

Linear Models

What follows is a description of techniques used in statistical prediction models. Many of these techniques are similar to validation schemes for numerical models, for which the strengths and weaknesses are shown in Table 2.1. Also, a more in-depth description is provided in Appendix A.

Correlation and Regression

There is a long history of using correlation patterns to identify teleconnections, beginning with the landmark Southern Oscillation studies of Walker in the early 20th century (Katz, 2002) and increasing exponentially since (e.g., Blackmon et al., 1984; Wallace and Gutzler, 1981). The idea of teleconnections in meteorology is tied closely to that of correlations. Base points in locations, such as the center of the Nino3.4 box (Hanley et al., 2003), have served as the origin for teleconnection analysis throughout the globe (Ding and Wang, 2005) and have formed the basis for forecast quality (Johansson et al., 1998). The relationships between two locations can be calculated by measuring the mean squared error between the base point and the remote location. Large correlations correspond to a large degree of covariability and a correspondingly small mean squared error. Linear regression is an extension of correlation where directionality is assumed in diagnosing relationships between a predictor variable and a response variable or predictand.

Most often, the variance of the response variable is partitioned into components that are explained or unexplained by the predictor. The coefficient of determination (known as R^2) gives the amount of variance explained by the predictor and is often used to assess the goodness of fit for a given model, though it has been criticized as a forecast performance index for verification as it ignores bias (Murphy, 1995). If the assumptions regarding the distribution of the data are met, significance of the model parameters can be assessed through t-tests. In cases when the assumptions are not met, bootstrapping of the (x,y) pairs or the model residuals has been shown to be effective (Efron and Tibshirani, 1993).

Often the problems addressed by regression require multiple predictors to give meaningful answers. The statistical model, multiple regression, is a generalization of simple regression. Rather than pairs of data measured simultaneously, n-tuples of data are used where all of the predictors, $x_1, x_2, \ldots, x_m$ and the predictand, y, form the training data that are observed over n cases.

Historically, the most common application of regression methodology has been for relating numerical model output to some predictand at a future time using linear regression. The method is called "model output statistics" (MOS) by Glahn and Lowry (1972). This is related to another regression technique, known as the perfect prog (PP) method (Klein et al., 1959) where both predictors and response variables are observed quantities in the training dataset. These methods are popular, as they utilize information at relatively larger scales to represent sub grid scale processes. MOS has the advantage over PP of correcting for forecast model biases in the mean and variance. The disadvantages of MOS include rebuilding the equations with changes in models and assimilation systems. Brunet et al. (1988) offer detailed comments on the relative advantages of each method, claiming PP was superior for shorter-range forecasts and MOS for longer time leads.

When cross-correlations are used to establish relationships between two non-adjacent locations, the maps of correlations are termed teleconnections. The earliest instance of using such a methodology was to establish the correlation structure of the Southern Oscillation (Walker, 1923). Maps of teleconnectivity at widely separated locations at a given geopotential height have been constructed to establish the centers of action of various modes in the mid-troposphere (Wallace and Gutzler, 1981). A catalogue of such teleconnections, based on

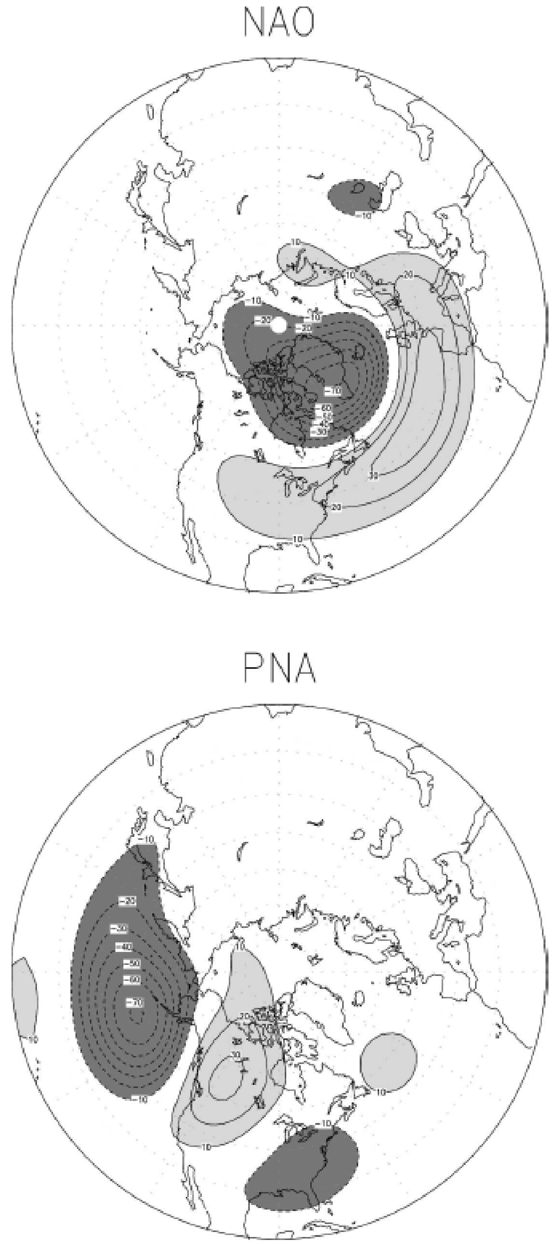

FIGURE 3.4 Example of two teleconnection patterns (North Atlantic Oscillation (NAO); Pacifc-North American pattern (PNA)), shown as anomalies in the 500-hPa geopotential height field. Dark shading indicates negative anomalies, and light shading indicates positive anomalies.The patterns emerge from a rotated Empirical Orthogonal Function analysis of monthly mean 500-hPa geopotential heights. Contour interval is 10 m. SOURCE: Johansson (2007).

principal component loadings (used to summarize the linked regions) was created by Barnston and Livezey (1987).

Empirical Orthogonal Functions (EOF)/Principal Component Analysis (PCA)

The use of eigentechniques was pioneered by Pearson (1902) and formalized by Hotelling (1933). The key ideas behind eigentechniques are to take a high dimensional problem that has structure (often defined as a high degree of correlation) and establish a lower dimensional problem where a new set of variables (e.g., eigenvectors) can form a basis set to reconstruct a large amount of the variation in the original data set. In terms of information theory, the goal is to capture as much signal as possible and omit as much noise as possible. While that is not always realized, the low dimensional representation of a problem often leads to useful results. Assuming that the correlation or covariance matrix is positive semidefinite in the real domain, the eigenvalues of that matrix can be ordered in descending value to establish the relative importance of the associated eigenvectors. Sometimes the leading eigenvector is related to some important aspect of the system. However, modes beyond the first are rarely related to specific physical phenomena owing to the orthogonality imposed on the EOFs/PCs. One possibility is to transform the leading PCs to an alternate basis set. This process is known as PC rotation (Horel, 1981; Richman, 1986; Barnston and Livezey, 1987) and has been shown to offer increased stability and isolation of patterns that match more closely to their parent correlation (or covariance) matrix. For example, in Figure 3.4, the rotated EOFs that are derived from monthly 500-hPa geopotential height data define two teleconnection patterns, the North Atlantic Oscillation (NAO) and the Pacific North American (PNA) pattern. These patterns explain a relatively large portion of the variance in the 500-hPa geopotential height data and can be related to the large-scale dynamics of the atmosphere as well as incidences of extreme weather in certain locations. In cases where the data lie in a complex domain, eigenvectors can be extracted in "complex EOFs." Such EOFs can give information on travelling waves, under certain circumstances, as can alternative EOF techniques that incorporate times lags to calculate the correlation matrix (Branstator, 1987). As was the situation for correlation analysis, EOF/PCA are data compression methods. They do not relate predictors to response variables.

Canonical Correlation Analysis(CCA)/Singular Value Decomposition (SVD)/Redundancy Analysis (RA)

A multivariate extension of linear regression is canonical correlation analysis (CCA). It can be thought of as multiple regression where there is more than one predictor ($x_1, x_2, \ldots, x_m$) and multiple response variables ($y_1, y_2, \ldots, y_p$). Consequently, CCA is useful for prediction of multiple modes of variability associated with climate forcing (Barnston and Ropelewski, 1992). The goal of CCA is to isolate important coupled modes between two geophysical fields. Singular value decomposition (SVD) is analogous to CCA when applied to an augmented covariance matrix. Despite the similarity, Cherry (1996) argues that the techniques have different goals. Both techniques can lead to spurious patterns (Newman and Sardeshmukh, 1995), particularly when the observations are not independent and the cross-correlations/cross-covariances are weak relative to the correlations within the x's and y's. In such cases, pre-

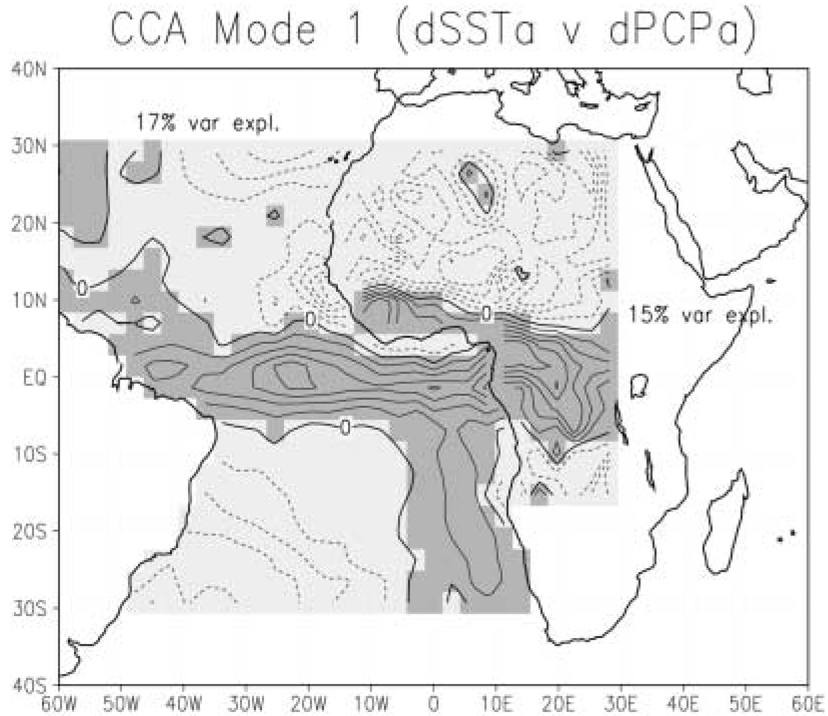

FIGURE 3.5: Patterns relating errors in SST to errors in precipitation. Dark gray areas in the ocean indicate areas of "warm" errors; lighter shading indicates "cold" errors. Over land, dark gray indicates "wet" errors; lighter shading indicates "dry" errors that accompany the pattern of errors in SST. Contours are in normalized units, with a 0.1 contour interval. SOURCE: Goddard and Mason (2002).

filtering the predictors and response variables with EOF or PCA may have benefits (Livezey and Smith, 1999). Given the oversampling in time and space of most climate applications, PCA is often used as the initial step to establish a low dimensional set of uncorrelated basis vectors subject to CCA (Barnett and Preisendorfer, 1987). Despite the potential pitfalls, CCA has been shown to exhibit considerable skill for long-range climate forecasting (Barnston and He, 1996) and is one of the favored techniques for relating teleconnections to climate anomalies. Figure 3.5 shows an example of how CCA has been used to relate errors in SST in the tropical Atlantic Ocean to errors in model-produced estimates of precipitation in parts of Africa. Recently, redundancy analysis (RA), a more formal modeling approach based on regression and CCA, has been applied successfully to find coupled climate patterns useful in statistical downscaling (Tippett et al., 2008).

Constructed Analogues

Natural analogues are unlikely to occur in high degree-of-freedom processes (see Table 2.1 regarding historical use of analogues in prediction). In reaction to this, van den Dool (1994) created the idea of constructing an analogue having greater similarity than the best natural analogue. The construction is a linear combination of past observed anomaly patterns in the predictor fields such that the combination is as close as desired to the initial state. Often, the predictor (the analogue selection criterion) is based on a reconstruction from the leading eigenmodes of the data field at a number of periods prior to forecast time. The constructed analogue approach has been used successfully to forecast at lead times of up to a year (van den Dool et al., 2003) and usually outperforms natural analogues forecasting one meteorological variable from another contemporaneously. A constructed analogue yields a single linear operator derived from data by which the system can be propagated forward in time.

Nonlinear Models

Most of the linear tools have nonlinear counterparts. Careful analysis of the data will reveal the degree of linearity. Additionally, comparison of the skill for linear versus nonlinear counterparts will reveal the degree of additional information to be gained by nonlinear methods. Specific recommendations on techniques to apply are given in Haupt et al. (2009).

Logistic Regression

Logistic regression is a nonlinear extension of linear regression for predicting dichotomous events as the response variable. The function that maps the predictor to the response variable is called the logistic response function, which is a monotonic function ranging from zero to one[10]. This involves minimizing the loss function using a nonlinear procedure. Logistic regression has been applied successfully to problems such as precipitation forecasting (Applequist et al., 2002), medium range ensemble forecasts (Hamill et al., 2004), and blocking beyond two weeks (Watson and Colucci, 2002).

Artificial Neural Networks (ANN)

Artificial Neural networks (ANNs) have been applied successfully to numerous prediction problems, including ENSO (Tangang et al., 1998) and precipitation forecasts from teleconnection patterns (Silverman and Dracup, 2000). An ANN is composed of an input layer of neurons, one or more hidden layers, and an output layer. Each layer comprises multiple units connected completely with the next layer, with an independent weight attached to each connection. The number of nodes in the hidden layer(s) is dependent on the process being modeled and is determined by trial and error. Such models require considerable investigator supervision to train as nonlinear techniques are prone to overfitting noise (finding solutions at local minima). During the training process the error between the desired output and the calculated output is propagated back through the network. The goal is to find the network architecture that generalizes best.

[10] The logistic response function is not simply an arbitrary monotonic function. It also follows a sigmoidal surface, which has an "S" shape when compared to a linear monotonic function.

Support Vector Machines (SVM)

Support vector machines (SVM) are a form of supervised learning techniques that use kernels to arrive at solutions at the global minimum or generalization error. For data existing in high dimensional space, SVM will separate the data into several subsets, attempting to achieve an optimal linear separation. This can be useful for noisy data sets.

SVM have been applied to cloud classification problems (Lee et al, 2004), wind prediction (Mercer at al., 2008) and severe weather outbreaks (Mercer et al., 2009). Comparison of SVM to standard logistic regression in Mercer et al. (2009) suggests that SVM is equal or superior to the more traditional techniques in minimizing misclassification of forecasts. On the ISI time scale, Lima et al. (2009) have shown that kernelized methods lead to additional skill in ENSO forecasts over traditional PCA techniques

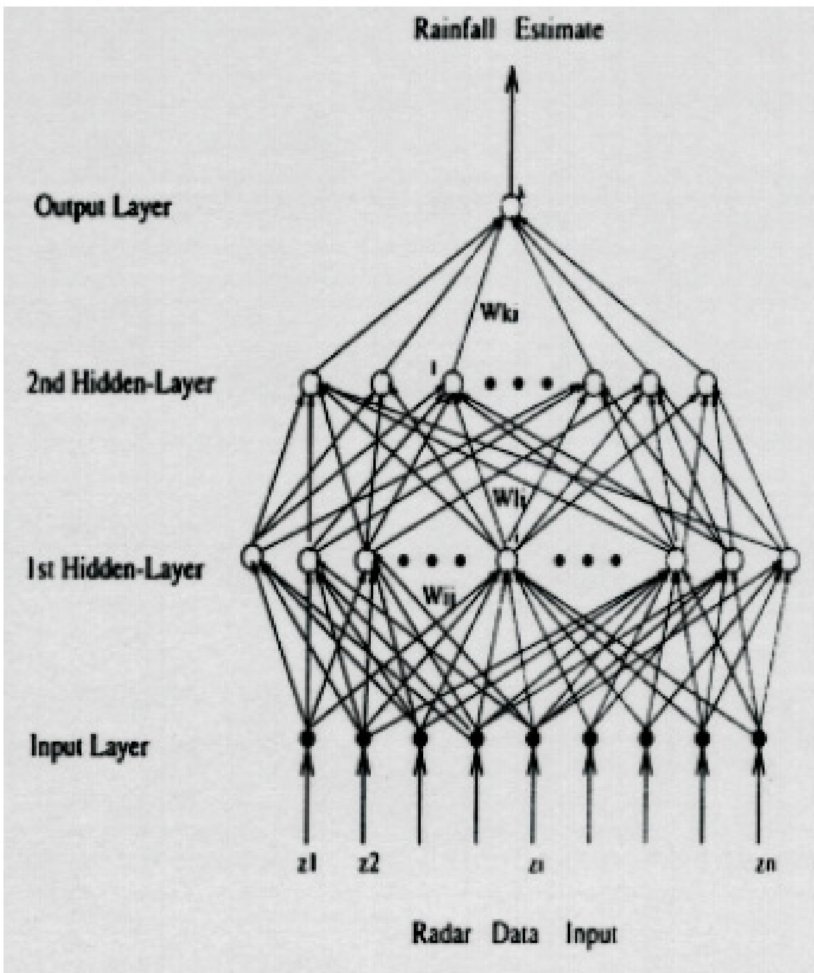

FIGURE 3.6 Schematic of an Artificial Neural Network (ANN) for forecasting rainfall from radar data. The network is composed of "nodes" (circles) that are linked by "weights" (arrows). The number of hidden layers and number of nodes in each layer can be specified by the user, or determined through experimentation. Weights are determined by training the network on a subset of the data. SOURCE: Figure 2 from Trafalis et al. (2005).

Composites

Determination of climate system processes often involves univariate or bivariate displays of the slowly varying forcing (e.g., ENSO, MJO). Investigating how such signal propagates through the climate system can be accomplished through correlation or regression. These approaches are unconditional as all the data are used to establish the pertinent relationships. Another possibility is to condition the relationships on subsets of time when the climate system is in a given state. Such states are termed composites. A key aspect of creating these states is to measure the intra-state variability to ensure that all cases assigned to a given state have commonality. The quality of a composite is often tested by calculating the means of each group to insure adequate separation. Relating climate linkages to such composites is commonly performed to relate forcing to effects in climate studies (e.g., Ferranti et al. 1990; Hendon and Salby 1994; Myers and Waliser, 2003; Tian et al. 2007 for the MJO). Most often, correlations are used to establish the linkages, although comparisons can be based on linear or nonlinear statistics.

Figure 3.7 provides an example of how composites can distinguish, better than linear methods, anomalous precipitation patterns in the continental United States associated with the SST anomalies in the tropical Pacific Ocean. The figure shows that the patterns of precipitation associated with warmer-than-average SST are not necessarily “mirror images” of the patterns of precipitation associated with colder-than-average SST. For example, for the warmer-than-average SST composite, areas of the Midwest exhibit significantly drier-than-average conditions during March, while the Great Plains experience wetter-than-average conditions (third row, left column). By comparison, the composite representing precipitation associated with colder-than-average SST shows near-average conditions for much of the Midwest, with significantly drier-than-average conditions across the Great Plains (third row, right column).

DYNAMICAL MODELS

With the advent of computers, it became feasible to solve the fluid dynamical equations representing the atmosphere and the ocean using a three-dimensional gridded representation. Physical processes that could not be resolved by this representation, such as turbulence, were parameterized using additional equations. The computer software that solves this set of equations is referred to as a dynamical or numerical model. The earliest dynamical models were developed for the atmosphere for the purposes of weather forecasting, and dynamical (or physical) models for other components of the climate system (land, ocean, etc.) followed thereafter.

Evolution of Dynamical ISI Prediction

Some of the earliest attempts at making ISI predictions with dynamical models were performed to essentially extend the range of weather forecasts. Miyakoda et al. (1969) described two-week predictions made with a hemispheric general circulation model (GCM). Miyakoda et al. (1983) used an atmospheric general circulation model (AGCM) with a horizontal resolution of 3–4 degrees and 9 levels in the vertical, which was the state of the art for weather forecasting at the time. Their study used a 30-day prediction to forecast a blocking event that occurred

FIGURE 3.7. Composites of monthly-mean precipitation anomalies associated with months experiencing warm SST (left column) or cold SST (right column) in the tropical Pacific for January through April. These maps illustrate that opposite signed SST anomalies do not necessarily produce opposite signed precipitation anomalies (e.g., the areas experiencing wetter-than-average conditions when SST is warm are not necessarily drier-than-average when SST is cold, and vice versa). Contour interval is 10 mm of precipitation; solid contours are for wet anomalies; dashed contours are for dry anomalies. Shading indicates areas of statistical significance. SOURCE: Adapted from Livezey et al. (1997).

during January 1977. The success of the prediction was attributed to improved spatial resolution and better representation of subgrid-scale processes.

Extended range numerical predictions of this sort were referred to as dynamical extended range forecasting (DERF) to distinguish them from short and medium range weather forecasts. The numerical models used for extended range forecasts were the same AGCMs that were then being used for weather forecasting. These AGCMs solved the basic three-dimensional fluid dynamical equations numerically, using either finite-differencing or spectral decomposition. The AGCMs also incorporated physical parameterizations of shortwave and longwave radiation, moist convection, boundary layer processes, and subgrid-scale turbulent mixing. The complexity and resolution of these AGCMs increased throughout the 1990s in concert with computing power and understanding.

Although the DERF activities had some limited successes, they were barely scratching the surface of short-term climate prediction, in part because some processes were still poorly represented (e.g., MJO, see Chen and Alpert, 1990; Lau and Chang, 1992; Jones et al., 2000; Hendon et al., 2000) and there was no coupling to the ocean. One of the premises of the DERF approach was that there was enough information in the atmospheric initial conditions to make useful extended range predictions. In the terminology of Lorenz (1975), this would be predictability derived from knowledge of the initial condition. Because of the rapid decay of quality with lead time, one would not expect useful predictions on seasonal or longer time scales to arise solely from atmospheric initial conditions. To obtain forecast quality on longer timescales, one has to consider predictability arising from the knowledge of the evolution of boundary conditions or external forcing[11] (Lorenz, 1975; Charney and Shukla, 1981).

One of the most important boundary conditions for an atmospheric model is sea surface temperature (SST). Variations in SST can heat or cool the atmosphere, influence the rainfall patterns, and thus change the atmospheric circulation. This is especially obvious in the tropical Pacific, where strong SST anomalies associated with the El Niño -Southern Oscillation (ENSO) phenomenon significantly alter atmospheric convection patterns. Although more subtle, and secondary to the initial conditions of the diabatic heating and circulation structure, SST as a boundary condition for properly initiating the MJO is also expected to be important (e.g., Krishnamurti et al., 1988; Zheng et al., 2004; Fu et al., 2006). The evolution in diabatic heating associated with ENSO and MJO events affects not only the local atmospheric circulation over the tropics, but also affects atmospheric circulation in extratropical regions such as North America through teleconnections (Wallace and Gutzler, 1981; Hoskins and Karoly, 1981; Weickmann et al., 1985; Ferranti et al. 1990).

The link between tropical Pacific SST and atmospheric anomalies elsewhere makes prediction of ENSO valuable for climate predictions in many remote regions. It is a continuing challenge to characterize this link (i.e., how a particular SST anomaly or evolution of anomalies may affect a given, remote location), especially given the complex interactions among local and remote processes that can contribute to predictability in a particular location. Better characterization of the link between ENSO (and other processes that affect boundary conditions for large-scale circulation) and the climate of remote locations is an important component for translating ISI forecasts into quantities useful for decision-makers (see “Use of Forecasts” section in this chapter).

In order to exploit atmospheric predictability associated with ENSO, one has to predict the SST in the tropical Pacific. The quasi-periodic nature of ENSO, with enhanced spectral power in the 4–7 year band, suggested that useful predictions might be possible months or seasons in advance. The next major step in short-term climate prediction came about when Cane et al. (1986) used a simple model of ENSO, a one-layer ocean representing the thermocline and a simple Gill-type model for the atmosphere, to make numerical predictions of ENSO events.

---

[11] It is possible that initial conditions of long-lived stratospheric phenomena could lead to long lead time skill, but this has yet to be demonstrated.

Successful predictions with the Cane-Zebiak model shifted the focus of short-term climate prediction to ENSO forecasting. ENSO is associated with much of the forecast quality at global scales in current forecast systems on seasonal to interannual timescales, although some other phenomena may dominate in specific regions. The type of model used by Cane and Zebiak is referred to as an Intermediate Coupled Model (ICM), because the atmospheric and the oceanic model are highly simplified. Following the success of the ICM approach, more sophisticated techniques were developed for ENSO prediction. One was the Hybrid Coupled Model (HCM) approach, where the atmospheric model remained simple but the one-layer ocean model was replaced by a comprehensive ocean general circulation model (OGCM). Neither the ICM nor the HCM approaches produced useful predictions of atmospheric quantities over continents. Therefore, a two-tier approach was used to produce climate forecasts over land. The SSTs predicted by the ICM/HCM (Tier 1) were used as the boundary condition for AGCM predictions (Tier 2).

Another approach to ENSO prediction was the use of a comprehensive coupled GCM (CGCM), where an AGCM is coupled to an ocean GCM, with the two models exchanging fluxes of momentum, heat, and freshwater. CGCMs were originally developed for studying long term (centennial) climate change associated with increasing greenhouse gas concentrations. CGCMs used for climate change used coarse spatial resolution to facilitate multi-century integrations. The shorter integrations required for ENSO prediction allowed finer spatial resolution, especially in the ocean, which could better resolve the processes important for ENSO. Finer resolution in the atmosphere improved forecast quality over the continents without requiring a two-tier approach. The quality of ENSO predictions in a CGCM arises almost exclusively from initial conditions in the upper ocean.

The major modeling/forecasting centers began to use CGCMs for ENSO prediction in the 1990s (Ji and Kousky, 1996; Rosati et al., 1997; Stockdale et al., 1998; Schneider et al., 1999) although the two-tier approach continued to be used operationally to predict the associated terrestrial climate. Atmospheric model resolution was initially about 2–4 degrees in the horizontal and the ocean model resolution was 1–2 degrees, often with substantially finer meridional resolution near the equator. Initial conditions were derived from an ocean data assimilation system.

Early attempts to use CGCMs for ENSO prediction fared poorly when compared to the ICM/HCM approaches or statistical techniques. CGCM predictions for ENSO suffered from "climate drift," where the model prediction evolved from the "realistic" initial condition to its own equilibrium climate state. This led to a rapid loss in quality for ENSO predictions. Statistical corrections applied *a posteriori* (Model Output Statistics, see "Correlation and Regression" section in this chapter) had only limited efficacy in arresting this loss of quality. Anomaly coupling strategies, where the atmospheric and oceanic models exchange only anomalous fluxes, were also used (Kirtman et al., 1997), but did not address the underlying deficiencies of the component models.

Over the last decade, the ENSO forecast quality associated with CGCMs has improved significantly. Reductions in the model bias and improved ocean initial conditions have now enabled CGCMs to be competitive with statistical models. An important development has been the use of multi-model ensembles (MME), where predictions from a number of different CGCMs are combined to produce the final forecasts (Krishnamurti et al., 2000; Rajagopalan et al., 2002; Robertson et al., 2004; Hagedorn et al., 2005). The Development of a European Multi-model Ensemble System for Seasonal to Interannual Prediction (DEMETER) project included seasonal predictions from seven different CGCMs, with atmospheric horizontal resolutions ranging from T42–T63 and oceanic horizontal resolution in the 1–2 degree range (Palmer et al., 2004). The MME forecast quality of the DEMETER ensemble (and other ensembles) beats the quality of any single CGCM that is part of the ensemble (Palmer et al., 2004; Jin et al., 2008). The MME anomaly correlation skill of Nino3.4 at 6 month lead time is 0.86 in the ensemble considered by Jin et al. (2008), with individual models showing lower correlations (some as low as 0.6).

In terms of intraseasonal and MJO prediction, evaluating and incorporating the role of ocean coupling has evolved somewhat independently. Given the shorter time scale relative to ENSO, the interaction with SST has been found to be mostly limited to the ocean mixed-layer (e.g, Lau and Sui, 1997; Zhang, 1996; Hendon and Glick, 1997). A number of model studies have indicated improvement in MJO simulation and prediction by incorporating SST coupling of various levels of sophistication (e.g., Waliser et al., 1999b; Fu et al., 2003; Zheng et al., 2004; Woolnough et al., 2007; Pegion and Kirtman, 2008).

Current Dynamical ISI Forecast Systems

Currently, CGCMs serve as the primary tool for dynamical ISI prediction. Improvements in atmospheric model resolution mean that it is no longer necessary to use a two-tiered approach for ISI prediction. In operational forecasting centers, CGCMs are used in conjunction with sophisticated data assimilation systems and statistical post-processing to produce the final forecasts. Typically, the atmospheric component of a CGCM is a coarse-resolution version of the AGCM used for short-term weather forecasts. A CGCM also includes a land component (as part of the AGCM), an ocean component, and optionally a sea ice component (Figure 3.8). CGCMs also include a comprehensive suite of physical parameterizations to represent processes such as convection, clouds, and turbulent mixing that are not resolved by the component models. In this section, we provide a brief overview of the state-of-the-art in model resolution for CGCMs used for ISI prediction at two of the major operational forecasting centers, NCEP and ECMWF.

The atmospheric component of the NCEP Climate Forecast System (CFS) (Saha et al., 2006), which became operational in August 2004, currently has a horizontal resolution of 200 km (T62) with 64 levels in the vertical. It is scheduled to have a six-fold increase in horizontal resolution in 2010. The oceanic component of the CFS is derived from the GFDL Modular Ocean Model version 3 (MOM3), which is a finite difference version of the ocean primitive equations with Boussinesq and hydrostatic approximations. The ocean domain is quasi-global extending from 74°S to 64°N, with a longitudinal resolution of 1° and a latitudinal resolution that varies smoothly from 1/3° near the equator to 1° poleward of 30°. The model has 50 vertical levels, with spacing between levels (resolution) ranging from 10 m near the surface to over 500 m in the bottom level. The atmospheric and oceanic components exchange fluxes of momentum, heat, and freshwater daily, with no flux correction. Soil hydrology is parameterized using a simple two-layer model. Sea ice extent is prescribed from observed climatology.

At ECMWF, the current generation of the Seasonal Forecasting System (v3) has an atmospheric model with a horizontal resolution of 120 km (T159), with 62 levels in the vertical (*http://www.ecmwf.int/products/changes/system3/*). In contrast, the current operational deterministic weather prediction AGCM used by ECMWF has a resolution of 16 km and 91 vertical levels. The ocean model has a longitudinal resolution of 1.4° and a latitudinal resolution that varies smoothly from 0.3° near the equator to 1.4° poleward of 30°. There are 29 levels in

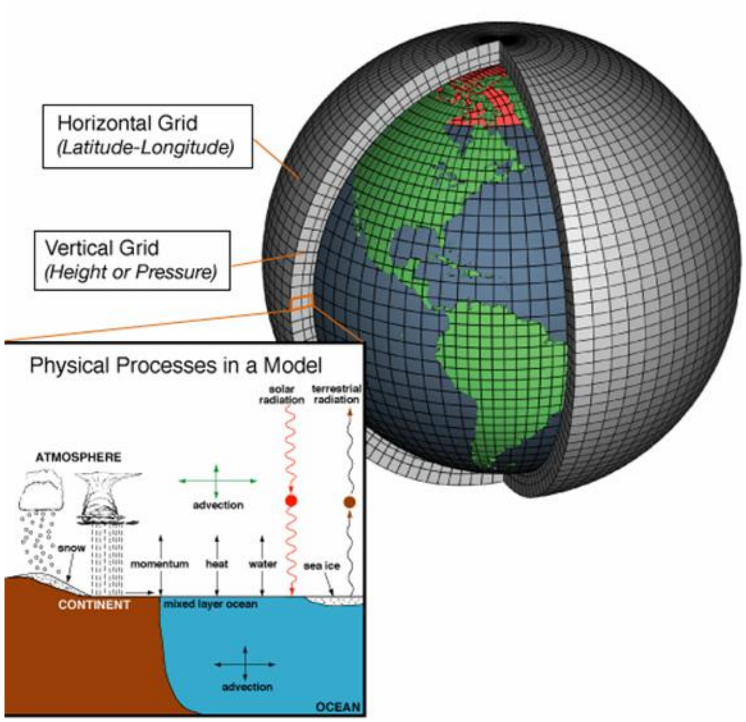

FIGURE 3.8 Schematic of a coupled general circulation model illustrating the horizontal/vertical grid, the different components (atmosphere, land, ocean), important physical processes, and air-sea flux exchange. SOURCE: NOAA.

the vertical. A tiled land surface scheme (HTESSEL) is used to parameterize surface fluxes over land. Sea ice is handled though a combination of persistence and relaxation to climatology.

Systematic errors are found in the mean state, the annual cycle, and ISI variance of climate simulations in the current generation of CGCMs (Gleckler et al., 2008). Model errors in the tropical Pacific, such as a cold SST bias or a ‘double’ Inter-tropical Convergence Zone, are particularly troublesome because they impact phenomena such as ENSO and the MJO that are important for ISI prediction. Indeed, the models often exhibit significant errors in the simulation of spatial structure, frequency, and amplitude of ENSO and the MJO. These errors lead to the degradation of ISI prediction quality in CGCMs. Although some of the systematic errors can be attributed to poor horizontal resolution of the CGCMs, other errors are attributable to deficiencies in the subgrid-scale parameterizations of unresolved atmospheric processes such as moist convection, boundary layers and clouds, as well as poorly resolved oceanic processes such as upwelling in the coastal regions. Improvements in both model resolution and subgrid-scale parameterizations are needed to address these problems.

Multi-Model Ensembles

As mentioned above, one source of error in dynamical seasonal prediction comes from the uncertainties arising from the physical parameterization schemes. Such uncertainties and errors may be, to some extent, uncorrelated among models. A multi-model ensemble (MME) strategy may be the best current approach for adequately resolving this aspect of forecast uncertainty (Palmer et al., 2004; Hagedorn et al., 2005; Doblas-Reyes et al., 2005; Wang et al., 2008; Kirtman and Min, 2009; Jin et al., 2009)[12]. Figure 3.9 demonstrates how a multi-model ensemble can outperform the individual models that are used to form the ensemble. The MME strategy is a practical and relatively simple approach for quantifying forecast uncertainty. In fact, as argued in Palmer et al. (2004), Kirtman and Min (2009) and a number of studies using the DEMETER seasonal prediction archive and the APCC/CliPAS seasonal prediction archive, the multi-model approach appears to outperform any individual model using a standard single model approach (e.g., Jin et al., 2009; Wang et al., 2009). Although the "standard" MME approach applying equal weights to each model is relatively straightforward to implement, it has some shortcomings. For example, the choice of which models to include in the MME strategy is in practice ad-hoc and is limited by the "available" models. It is unknown whether the available models are in any sense optimal. Indeed, it is an open question as to whether more sophisticated single model methods such as perturbed parameters or stochastic physics will out perform MME strategies.

Developing alternative methodologies for combining the models can be challenging since the hindcast records for CGCMs used to assign weightings to the models are limited. Using predictions or simulations from AGCMs allows for longer records. One example is the super ensemble technique proposed by Krishnamurti et al. (1999), where the individual model weights depend upon the statistical fit between the model's hindcasts with observations during a training period. If a model has consistently poor predictions for a variable at a specific location during the training period, the weight could be zero or negative. Another approach is the Bayesian combination approach developed by Rajagopalan et al. (2002) and refined by Robertson et al. (2004), in which the prior probabilities are equal to the climatological odds, and models are optimally weighted based on probabilistic likelihood based on past performance. An outstanding question for MME research involves explaining why some MME statistics, such as the ensemble mean, consistently outperform the individual models. Similarly, it would be valuable to improve our understanding of what the ensemble mean and ensemble spread represent and how differences among these statistics can be best evaluated following MME experiments.

DATA ASSIMILATION

For the purposes of climate system prediction, data assimilation (DA) is the process of creating initial conditions for dynamical models. Since ISI predictions are based on coupled ocean-land-atmosphere models, it seems apparent that data assimilation eventually needs to be carried out in a coupled mode. At the present, however, data assimilation is being done separately for different model components, with exceptions such as the partial coupling carried out in the recent NCEP reanalysis (Saha et al., 2010). In the following sections, the current approaches (non-coupled) for carrying out assimilation on atmospheric, ocean, and land observations are discussed.

[12] Uncertainties in predictions can also arise from the initial conditions. The approach to quantifying the uncertainty due to initial condition uncertainty includes selective sampling procedures such as the bred-vector perturbation (e.g., Toth and Kalnay, 1993) and the singular vector technique (e.g. Molteni et al., 1996), or simply sampling initial states that are separated by several hours (Saha et al., 2006).

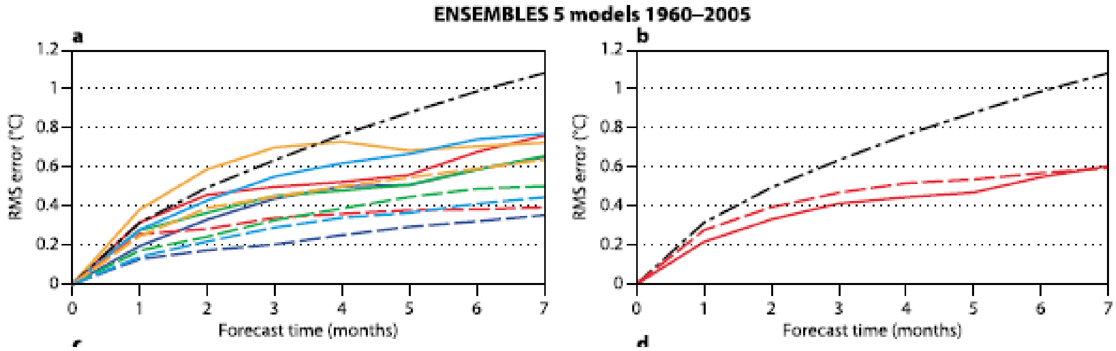

FIGURE 3.9 Comparison of RMSE of individual models to the multi-model mean for for Nino3 SST (solid) and ensemble standard deviation (spread) around the ensemble mean (dashed). The curves on the left represent individual models; the red curve on the right represents the multi-model mean The multi-model mean has a lower RMSE at nearly all forecast lead times, and a larger spread among its members, indicating that it outperforms any of the individual models. The black dash-dotted curve indicates the performance of a persistence forecast. SOURCE: Weisheimer et al. (2009).

Atmospheric Data Assimilation

Early efforts at data assimilation for short-term weather prediction used *a priori* assumptions about the statistical relationship between observed quantities and values at model gridpoints. The most sophisticated of these early methods was referred to as optimal interpolation (OI). Modern data assimilation for short-term numerical weather prediction objectively combines observations, model predictions started at earlier times, and *a priori* statistical information about the observations and the model to create initial conditions for updated model predictions.

The central theme of the evolution of atmospheric DA has been to use more information from the prediction model as both the models and the DA algorithms themselves have improved (Kalnay, 2003). In the earliest systems, the only information used from the model was the relative locations of model gridpoints (Daley, 1993). Later, short-term model predictions were used as a “first-guess” field that was then adjusted to be consistent with available observations (Lorenc, 1986).

A major advance was the development of variational data assimilation methods in which a cost function measuring the fidelity of the model’s estimation of the observed values is minimized using tools from variational calculus (LeDimet and Talagrand, 1986). Three-dimensional variational (3D-Var) techniques were implemented first (Parrish and Derber, 1992), with the most recent state from a model prediction being modified to better fit the observations. Variational techniques require *a priori* specification of a background error covariance, an estimate of the statistical relationship between different model state variables (Courtier et al., 1998). Although in principle OI and 3D-Var are nearly equivalent, (Lorenc, 1986), the ability of 3D-Var to find a global solution using all observations simultaneously resulted in less noisy and more balanced initial conditions for the predictions. More recently, four-dimensional variational (4D-Var) techniques have become the state of the art for operational numerical weather prediction. These techniques adjust the initial state of the model at an earlier time so that the

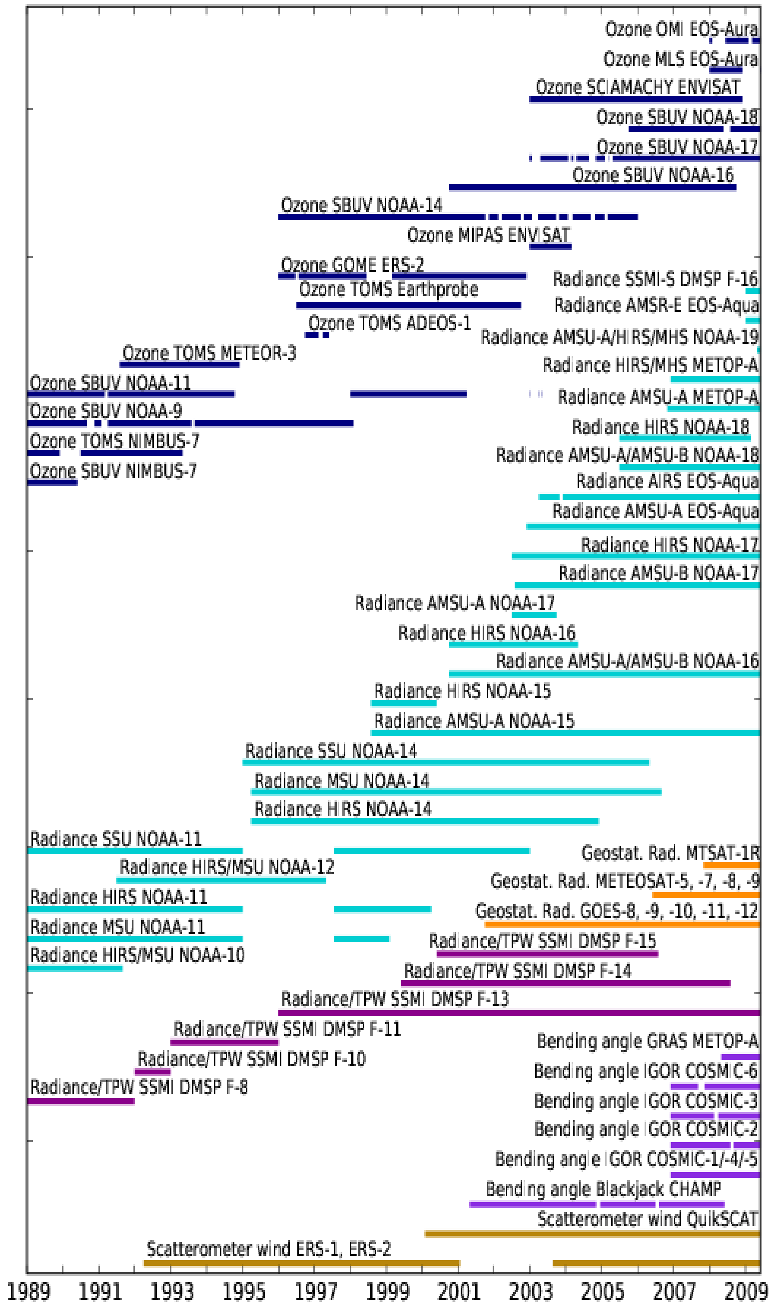

FIGURE 3.10 Satellite observing systems available for data assimilation in the ERA-Interim starting in 1989. SOURCE: ECMWF.

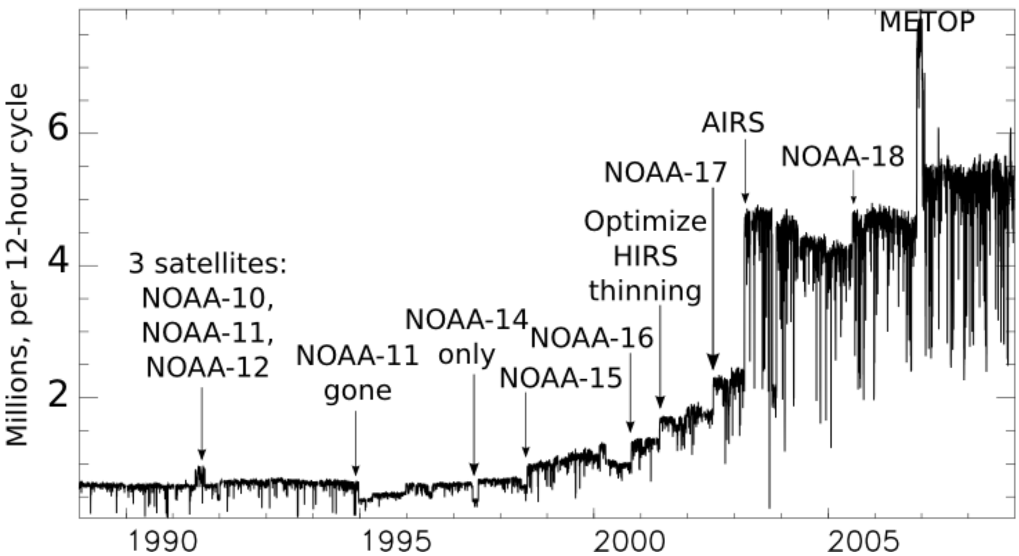

FIGURE 3.11 Number of satellite observations assimilated into the ECMWF ERA-Interim reanalysis (data assimilation system), with arrows indicating the introduction of new systems. SOURCE: ECMWF.

model evolves to fit a time sequence of available observations (Rabier et al., 1999). Predictions are then made by extending this model trajectory into the future.

Assimilation of remotely sensed atmospheric observations has played a large role in the increase in prediction quality over the last two decades (see "Atmospheric Observations" section, Figures 3.10, and 3.11). Figure 3.10 shows the types of remotely sensed satellite observations since 1989, and Figure 3.11 shows the number of satellite observations assimilated in the ECMWF Interim Reanalysis (ERA-Interim): about 1.5 million/day in 1989, jumping to 10 million/day in 2002 with the introduction of AIRS high resolution infrared sounder, and with another large increase from the high resolution infrared interferometer IASI (METOP). Figure 3.12 compares two recent reanalyses performed at ECMWF, the ERA 40, carried out with 3D-Var, and ERA-Interim, an experimental 4D-Var reanalysis. The obvious difference in performance between the two systems (which use the same observations but differ in their data assimilation systems) during the overlapping years quantifies the importance of the methods used for data assimilation, quality control, and advances in the model. It is remarkable that in the "reforecasts" from the ERA-Interim, it is possible to detect the improvement due to the introduction of AIRS in 2003, with a perceptible increase in anomaly correlation values in the five- and seven-day predictions.

The most recently developed methods for atmospheric data assimilation are ensemble Kalman filter (EnKF) techniques that use a set of short-term model predictions to sample the probability distribution of the atmospheric state. The ensemble provides information about both the mean state of the model and the covariance between different model variables. The ensemble members are adjusted using observations to produce initial conditions for a set of predictions. EnKF techniques are now in operational use for ensemble weather prediction (Houtekamer and Mitchell, 2005). Understanding the relative capabilities and advantages of 4D-Var and ensemble methods is an area of active research (e.g., Kalnay et al., 2007; Buehner et al., 2009a and b). There is a developing consensus that a "hybrid" approach combining a variational system (3D-Var or 4D-Var) with EnKF may be optimal.

In concert with using increasing amounts of information from the numerical model, increasingly sophisticated DA techniques facilitated the use of a more diverse set of observations. The earliest techniques were limited to assimilating observations of quantities that were one of the model state variables. Variational methods facilitated the assimilation of any observation that could be functionally related to the model state variables. However, *a priori* estimates of the relationship between errors in estimates of model state variables and the observed quantities were required. Ensemble methods automatically provide estimates of these relationships making it mechanistically trivial to assimilate arbitrary observation types. The types and numbers of observations assimilated for NWP has soared as DA techniques have improved in concert with the development of remote sensing systems that produce ever increasing numbers of observations.

While it can be difficult to separate prediction and forecast improvements due to model enhancements, DA advances, and increased numbers of observations, there is no doubt that all three have played a major role in improvements in NWP during the last decade (Simmons and Hollingsworth, 2002). In the mid-1990s, operational NWP centers seemed to be faced with a saturation in prediction quality (see Figure 2.1 on ECMWF 500-hPa geopotential height anomaly correlation). Since then, however, the rate of quality improvement has accelerated again. This acceleration is generally attributed to the direct assimilation of globally distributed Advanced Microwave Sounding Unity (AMSU) radiances (English et al., 2000), but it is important to remember that these observations could only be used effectively with advanced DA techniques (e.g., 4D-Var) and improved models. Removing bias from the observations was also essential. Similar improvements in quality for intraseasonal to interannual prediction could be expected by implementing improved DA in ocean, land surface, and possibly cryosphere models.

Ocean Data Assimilation

As pointed out in Chapter 2, much of the information required for successful ISI predictions resides in the initial conditions for the ocean, land surface, and cryosphere components of the climate system. There is a much shorter history of prediction and data assimilation for these components. This is partially due to the difficulty of observing these systems. *In situ* observations of the ocean, especially in remote areas or the deep ocean, have been difficult and expensive to obtain. It is also difficult to take *in situ* observations of quantities like land surface temperature and moisture (Reichle et al., 2004), or snow and ice thickness and extent (Barry 1995). In the last two decades, the number of observations of the ocean has soared, as noted earlier in the Observations section. Moored buoy systems have been developed in all of the world's tropical oceans and provide high frequency measurements of temperature, salinity, and currents. Global networks of autonomous drifting surface buoys and ocean sounders have been deployed, and remote sensing measurements of surface temperature, sea surface height, and ocean color are routinely available. Remote sensing observations of the land and cryosphere are also now available.

Since the ocean is considered to be the source of the majority of seasonal predictability, and the ocean is the best observed non-atmospheric climate system component, it has been natural to focus ISI DA efforts here. Sea surface temperature estimates, made using primitive assimilation techniques, have been available since the 1970s (Miyakoda et al., 1979) and

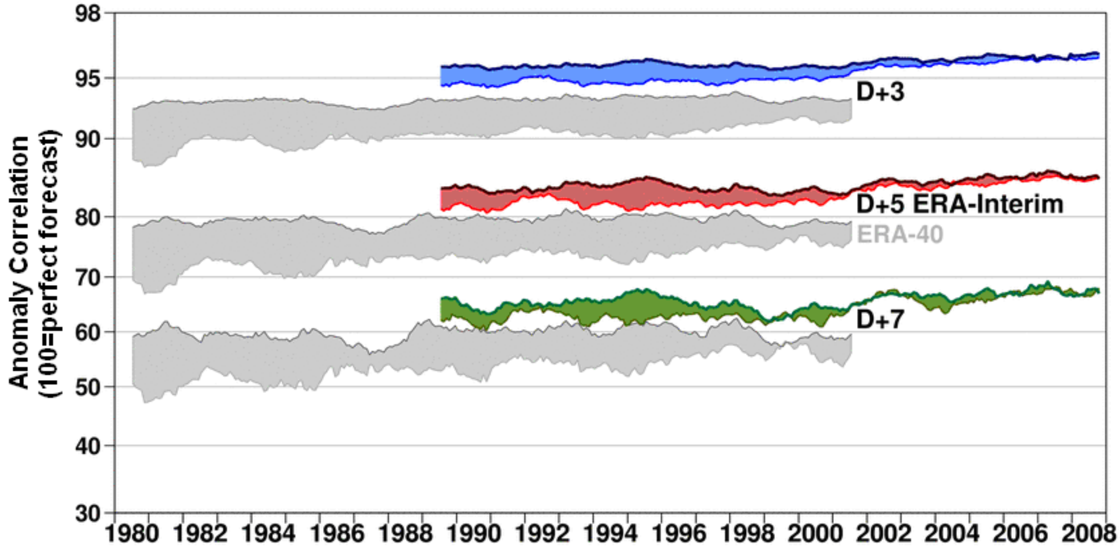

FIGURE 3.12 ECMWF 500-hPa geopotential height anomaly correlations from two different reanalysis systems. Gray: ERA-40 with 3D-Var (ca. 1998); Colors: ERA-Interim which uses 4D-Var (ca. 2005). "D+3" corresponds to the 3-day forecast; "D+5" the 5-day forecast; and, "D+7" the 7-day forecast. In each case the top line is the anomaly correlation of the forecasts started from the reanalysis for the Northern Hemisphere, and the bottom line is the corresponding forecast for the Southern Hemisphere. Note the improvement brought about by the improvement of the data assimilation system, which is especially important in the Southern Hemisphere. SOURCE: ECMWF.

continue to be produced (Reynolds et al., 2002). These methods are used routinely to produce products like analyses of the sea surface temperature. The first use of more modern data assimilation was a 3D-Var/OI system that used the Geophysical Fluid Dynamics Laboratory (GFDL) ocean forecast model and surface fluxes from the GFDL First GARP Global Experiment (FGGE) reanalysis to predict a "first-guess" for the assimilation (Derber and Rosati, 1989). While capable of assimilating observations of any model state variable, the system was primarily used with observations of temperature, and to a lesser extent salinity. Ocean DA systems based on this algorithm continue to be used at operational prediction centers like NCEP, ECMWF and UKMO. Many of them incorporate a number of heuristic enhancements, for instance the use of independently produced sea surface temperature analyses instead of observations of near-surface ocean temperature. Most current ocean DA efforts are using assimilation algorithms that would be regarded as outdated for atmospheric applications.

Ocean data assimilation has several major challenges compared to atmospheric data assimilation: (1) the observing system is sparser, and started later than the global atmospheric observing system, (2) the models are arguably worse in representing the real ocean, and (3) the time scales for forecasting are longer, so both the analyses and the forecasts (and the verifications) are less frequent. However, if advanced assimilation systems (4D-Var or EnKF) for the ocean are developed, tuned, and used operationally, similar improvements in the ocean analysis and predictions would be expected.

Here we briefly review two ocean DA systems that are currently in use: the NCEP Global Ocean Data Assimilation System (GODAS) and the ECMWF System 3 (S3), implemented in 2006[13].

NCEP GODAS

Coupled ocean-atmospheric forecasts and ocean data assimilation (ODAS) were pioneered at NCEP under the direction of Ants Leetmaa (Ji et al., 1998), who created a data assimilation system for the Pacific for the purpose of predicting ENSO. This was significantly improved, and the last version (RA6) of ODAS has been widely used (Behringer et al., 1998). RA6 was replaced by a Global Ocean Data Assimilation System (GODAS), which was coupled with the successful Climate Forecast System (Saha et al., 2006).

The numerical model in GODAS is the GFDL MOM-v3, with a horizontal resolution of 1° x 1°, enhanced to 1/3° in latitude within 10° of the equator. It has 40 levels with 10 m resolution in the upper 200 m, an explicit free surface, a Gent-McWilliams (Gent and McWilliams, 1990) mixing and a KPP (K-Profile Parameterization, Large et al*,* 1994) vertical mixing. The model is forced at the surface by analyzed momentum flux, heat flux, and fresh water flux produced by the NCEP atmospheric Reanalysis 2 (R2).

The GODAS data assimilation is based on the 3D-Var/OI of Derber and Rosati (1989). In addition the model top level is relaxed towards the Reynolds weekly SST analysis, and the surface salinity is relaxed towards the annual salinity climatology (from Levitus, 1982). GODAS assimilates temperature profiles from XBTs, moored buoys including TAO, TRITON and PIRATA, and from Argo profiling floats (see "Ocean Observations" section in this chapter). In addition, for each temperature profile, a synthetic salinity profile is created from a local climatology of the temperature-salinity relationship. These salinity profiles are also assimilated. Although observed salinity is not currently assimilated, experiments using Argo salinities (Huang et al, 2008) showed a clear improvement with a reduction of errors not only in salinity but also in currents.

One of the challenges for 3D-Var is defining appropriate multivariate background covariances that allow observations of one quantity to impact state variables of another type, for instance, having assimilation of salinity observations directly impact temperature state variables. The current GODAS is univariate in this sense. Multivariate GODAS was developed and tested but performs worse than univariate in assimilating salinity, possibly due to the use of synthetic salinity profiles.

Remote sensing measurements of sea surface height, which is a model state variable, from the TOPEX/Jason-1 altimetry have been available since 1992. Behringer (2007) indicates that assimilating surface heights (SSH) directly is not effective in this system, and instead it is used as a constraint on the baroclinic (temperature and salinity) analysis. When assimilated, it improves the anomalous SSH with respect to the observations but other aspects of resulting forecasts may be degraded.

[13] The committee is aware that other more sophisticated ocean DA systems exist (e.g., MERCATOR, HYCOM, NCOM, OPAVAR, MIT ECCO). However, this section focuses on systems currently used in operational settings and the prospects for their improvement.

ECMWF S3

The ECMWF S3 and NCEP GODAS systems are quite similar. The S3 (Balmaseda et al., 2008) is based on the HOPE-OI scheme. The ocean model (HOPE, Wolff et al., 1997) has the same resolution as MOM-v3 (1° x 1° horizontally) but only 29 vertical levels. It uses Optimal Interpolation (OI), which is nearly equivalent to the 3D-Var of GODAS. The operational system assimilates subsurface temperature and salinity, along with altimeter sea surface height anomalies. All observations in the upper 2000 m are assimilated. The sea surface temperatures are strongly relaxed to the Reynolds SST analysis. Surface forcing is similar to that in GODAS with fluxes from the ERA40 reanalysis (1959–June 2002) and the operational prediction system thereafter. Since the precipitation-evaporation flux is inaccurate in ERA40, a correction for precipitation (Troccoli and Kallberg, 2004) is used. An online additive bias correction has recently been added to the system (Balmaseda et al., 2008) that allows a reduction of the relative weight given to observations and reduces the strength of the relaxation to climatology. The main differences between the S3 and GODAS systems are that the S3 assimilates Argo salinities and altimeter data, and it includes a relatively sophisticated bias correction.

Improvements to Operational Ocean Data Assimilation Systems

There has been research with both 4D-Var (Weaver et al., 2003: Stammer et al., 2002) and EnKF (Keppenne and Rienecker, 2002) assimilation for the ocean. To date, the operational ISI community has been hesitant to adopt these more advanced assimilation methodologies because tests indicate that they may result in comparable and even worse forecasts, and in data sparse conditions they may not offer improvement. Since more advanced methods rely increasingly on the fidelity of the dynamical model, this may indicate that ocean models are not yet sufficiently accurate for these methods. Analogy to the atmosphere suggests that a program of model improvement combined with the incorporation of more sophisticated assimilation techniques that can make better use of all available observations is likely to lead to improvement for the ocean component of operational ISI predictions.

Land Data Assimilation

Land data assimilation systems (LDAS), as described in the "Land Observations" section, use surface meteorological forcing inputs that are based on observations as much as possible. Here we describe progress on another aspect of land assimilation, i.e., the merging of land surface state observations with estimates from the corresponding land model prognostic variables using mathematically optimal techniques.

A popular approach involves adjusting the land model's soil moisture reservoirs in response to screen-level (2 m) observations of atmospheric temperature and humidity using Optimal Interpolation (OI). If simulated relative humidity is too low compared to observations, soil moisture is increased so that evaporation increases, thereby increasing the simulated humidity. While this approach for initialization has been used with success in many operational centers (with success measured as improved weather forecasts), errors in simulated relative humidity and temperature need not stem from errors in soil moisture; they could stem from errors in parameterization, so that the modified soil moisture contents may not be more accurate than the original estimates. Drusch and Viterbo (2007) note that soil moisture profiles obtained through the OI approach are not necessarily sufficient for hydrological or agricultural applications; presumably, they may not be optimized for seasonal forecast initialization, either.

More recent land assimilation efforts have focused on the Kalman filter. ECMWF and Meteo-France, for example, are poised to implement in their operational NWP systems an extended Kalman filter (EKF) for assimilating soil moisture information derived from active and passive sensors (and potentially other variables as well). Currently, operational soil moisture retrievals are generated from Advanced Scatterometer (ASCAT) observations by the EUMETSAT satellite. The recently launched SMOS and the planned SMAP missions will provide L-band soil moisture information down to 5 cm in many areas. NASA/GSFC has pioneered the development of the ensemble Kalman filter (EnKF) for assimilating soil moisture retrievals or associated radiances into a land model; analyses with SMMR and AMSR-E soil moisture data show that EnKF assimilation produces soil moisture products with increased accuracy over model products or satellite retrievals alone (Reichle et al. 2007). Further plans for land data assimilation in various institutions include the use of a Kalman filter (EKF or EnKF) for snow data assimilation, the development of multivariate land data assimilation methods, and the assimilation of land variables as part of a coupled land data assimilation system.

USE OF FORECASTS

ISI forecasts can be valuable tools for decision makers. However, the use of a forecast is predicated on its quality. A forecast that is not sufficiently accurate or reliable is unlikely to be used to make decisions. The quality of a forecast can be determined through a variety of metrics (see “Forecast Verification” section in Chapter 2 for more details on forecast verification metrics). Typically, multiple metrics are used to provide an overall sense of forecast quality.

Understanding or improving forecast quality requires information about the predictions that go into a forecast. Increasingly, forecasts are generated from multiple prediction inputs, which can be objective (e.g., predictions from statistical or dynamical models) or subjective (e.g., expert opinion of forecasters). Detecting how these inputs or changes to these inputs affect forecasts is critical for improving forecast quality.

Finally, forecasts can be used by decision makers if they are provided in the appropriate format. A forecast regarding SST in the tropical Pacific may not be easily translatable to the local climate conditions of a particular user. In addition, different variables will have varying levels of value for different users. Some decision makers may require information about the seasonal mean values for particular meteorological variables, such as precipitation or temperature, while others may be more interested in certain extreme events, such as heat waves or incidents of heavy precipitation. Scale can also be important—some decision makers may require national or regional information, while others might be more focused on a city or even a piece of infrastructure. All of these factors can contribute to the utility of an ISI forecast with respect to societally-relevant decisions.

FIGURE 3.13 Seasonal precipitation forecasts are skillful over a larger portion of land during ENSO events (solid black line) than ENSO-neutral conditions (dashed black line). The purple line indicates the percentage increase in area between ENSO events and ENSO-neutral conditions. Top panel corresponds to the entire globe; bottom panel the area within 30° latitude of the equator. SOURCE: Goddard and Dilley (2005).

Measuring Forecast Skill and Assessing Forecast Quality

Evolution over time in skill according to various metrics can occur for multiple reasons, the main two factors being changes in the sources of predictability within the climate system and changes in the forecast system. Given a constant forecast system, changes in, for example, correlation between the predictions and observations of the climate over time come primarily from changes in the sources of predictability. It is a matter of signal versus noise; the signal due to forcing from SSTs or soil moisture changes noticeably from year to year, whereas the noise due to the internal dynamics of the atmosphere remains largely constant (Kumar et al., 2000). Overall, the seasonal climate predictions are more confident, and many skill metrics are higher, during ENSO events (Figure 3.13: Goddard and Dilley, 2005), as demonstrated by comparing forecast verifications during El Niño and La Niña conditions against those during ENSO-neutral conditions. The influence of ENSO, and other drivers of teleconnection patterns, on predictability remains incompletely understood.

For the United States, as with many other regions, much of currently realized quality in official forecasts is due to ENSO (Livezey and Timofeyeva, 2008), and the rest is attributed to climate trends, at least for temperature. For the official precipitation forecasts most of the skill

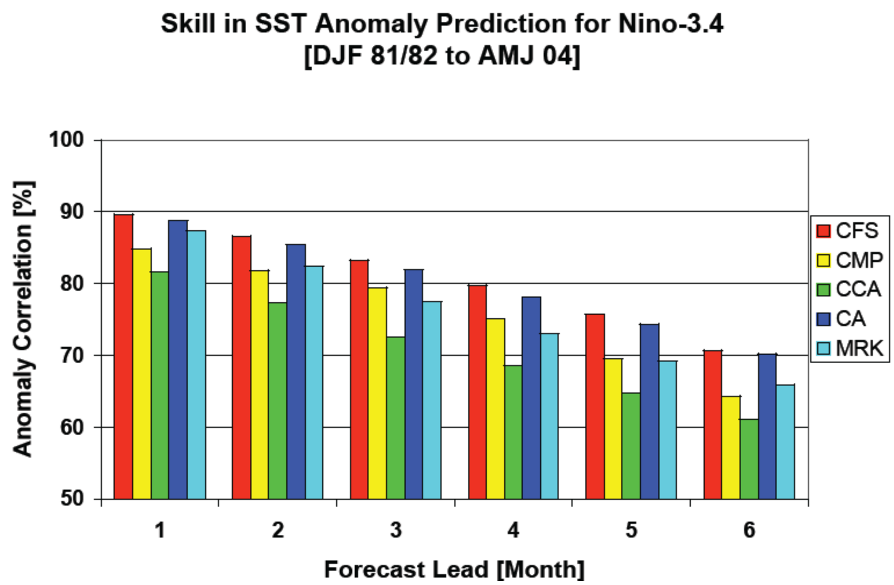

FIGURE 3.14 Nino3.4 correlation coefficient (predictions versus observations) for retrospective forecasts plotted as a function of lead time. The red and yellow bars correspond to dynamical models: the Climate Forecast System (CFS) is a state-of-the-art model developed in the mid-2000s (Saha et al., 2006), while the Coupled Model Project (CMP) prediction is older and was developed in the mid 1990s (Ji et al., 1995). CCA, CA, and MRK correspond to statistical models (Canonical Correlation Analysis, Constructed Analogues, Markov; several of these methods are discussed in "Statistical Models" section in this chapter and Appendix A). The figure highlights two points: (1) comparing the red and yellow bars indicates how coupled dynamical models have improved for this particular metric over the last two decades and (2) the statistical methods and the dynamical model methods are quite competitive with each other. SOURCE: Adapted from Saha et al. (2006)

assessed by the modified Heidke skill score derives from ENSO. Outside of ENSO, the skill assessment (Livezey and Timofeyeva, 2008) suggests that where trends are observed but skill is low, the information from trends has been underutilized by CPC. In the cases where skill is significant but the trends are negligible, it is suggested that the source of skill is the decadal scale variability captured through their statistical tool of Optimal Climate Normals. For temperature, the analysis does show that a few areas of positive Heidke skill in the official forecasts are found additionally during non-ENSO conditions in regions where temperature trends are weak (Livezey and Timofeyeva, 2008). Whether this indicates potential sources of predictability beyond ENSO and trend or is a lucky draw from a limited set of subjectively derived predictions is unclear.

In the case of a changing forecast system, one can demonstrate improvements in prediction quality by comparing the different forecast systems, such as the predictions of the newer system to those of older system over a common period. For example, when the correlation skill of Nino3.4 in the NCEP Climate Forecast System (CFS) dynamical model reached parity with that of statistical approaches (Figure 3.14), this was seen to be a significant

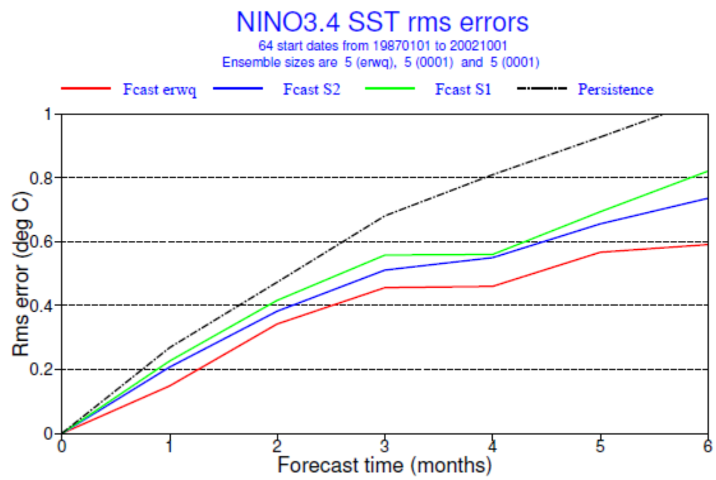

FIGURE 3.15 Improvement in forecasts from the latest ECMWF forecast system (red line) compared to earlier versions (blue and green lines). The black line corresponds to a Persistence forecast. Metric shown is RMS error for Nino 3.4 SST for 64 forecasts in the period 1987–2002. SOURCE: ECMWF, Anderson et al (2007).

accomplishment. Incremental improvements in the ECMWF dynamical model are illustrated through successive reduction in the RMSE of Nino3.4 predictions (Figure 3.15). Comparative assessments of new versus old forecast systems can really only be quantified for fully objective forecast systems, although one could demonstrate improvements of a newer objective system over a previous non-objective system (i.e., one involving subjective intervention). It should also be noted that these assessments of improvement potentially suffer from sampling issues, since there are typically not more than 20 years of retrospective forecasts for comparison. However, there are coordinated international efforts (e.g., the Climate Historical Forecast Project; CHFP) to extend the retrospective forecast period further back in time.

Currently, forecast quality is often difficult to compare across systems because of differences in forecast format, verification data, the choice of skill metrics, or even differences in graphical appearance. A mechanism to provide a consistent view of prediction quality across models was established in 2006 by the World Meteorological Organization. The charge was taken up by the lead center for the Standard Verification System of Long Range Forecasts (LC-LRFMME: *http://www.bom.gov.au/wmo/lrfvs/*) co-hosted by the National Meteorological Services of Australia and Canada. The LC-SVSLRF responsibilities include maintaining an associated website displaying verification information in a consistent and similar way. It allows forecasting centers to document prediction quality measured according to a common standard. The SVS is defined in Attachment II.8 (p. 122) of the WMO Manual on the Global Data-Processing and Forecasting System (WMO No. 485). Unfortunately, the goal of the comparative assessment envisioned by the WMO has not been achieved because it depends on the cooperation of the global producing centers (*http://www.wmolc.org*) to contribute consistent verification data, preferably in a common graphical format, which has not yet happened.

Comparative estimates of quality can be similarly difficult to quantify, even for the U.S. forecasts. One of the few studies to date compares the official subjective forecasts since 1995 with a newly implemented objective methodology that combines three statistical and one dynamical tool (Figure 3.16; O'Lenic et al., 2008). As stated above, the objective combination outperforms the subjective forecasts. The skill metric used in that study is the Heidke score, which was discussed in the previous section; it is not advocated by the WMO-SVSLRF. Assessment of forecast quality from the NCEP CFS model does not use Heidke skill scores, but rather correlation, Brier skill scores (BSS), and reliability diagrams (Saha et al., 2006). In Saha et al. (2006), a widely used statistical tool is compared to the dynamical model, with the result that the two methods have comparable but complementary BSS; regions of highest skill rarely overlap. Their result strongly suggests that additional predictability that is seen by the statistical tool but not currently captured in the dynamical model could result from improvements to that model. It also suggests the benefit of using both statistical and dynamical modeling approaches for seasonal climate prediction.

Combined Forecast Systems

A growing body of literature touts the benefit of multiple prediction inputs in climate forecasts. Many national centers that produce real-time forecasts include one or more dynamical models, one or more statistical models, and perhaps also the subjective interpretation or experience of the forecasters involved. As this practice continues, and as more prediction inputs become openly available, it is possible to assess the relative benefits of each type of prediction input to the quality of the forecast. In addition, as more prediction data becomes openly available, new methods for making the best use of that information can be tested and documented.

Subjective Combination

Since the early 1970s, weather forecasts in the United States and elsewhere were subjectively derived using the objective input as guidance (Glahn, 1984). In the mid-1980s, comparison of the skill of these objective and subjective forecasts according to several metrics indicated that the subjective weather forecasts were generally more skillful than the objective ones for shorter lead times (e.g. 12–24 hours), whereas the two types of forecasts exhibited approximately equal quality for longer lead times (e.g. 36–48 hours; Murphy and Brown, 1984). The same study further showed that both types of forecasts had positive trends in correlation skill over the decade, with improvements in objective forecasts equaling or exceeding improvements in subjective forecasts.

The use of subjective guidance has continued to this day for weather and now climate forecasts. Many seasonal-to-interannual forecasting centers, particularly those that use multiple prediction inputs, maintain a subjective element in their forecasts. At CPC and UKMO, for example, inputs from both statistical and dynamical prediction tools are considered and discussed, prior to "creating" a forecast (Graham et al., 2006; O'Lenic et al., 2008). In some instances, a subset of the tools will be objectively combined prior to their consideration next to other tools. Starting around 2006, CPC began objectively combining its main prediction tools, which consist of the Climate Forecast System (CFS) dynamical model and three statistical prediction tools, using an adaptive regression technique. This consolidation serves as a "first guess" but then is discussed with a number of other inputs, which include other consolidations as

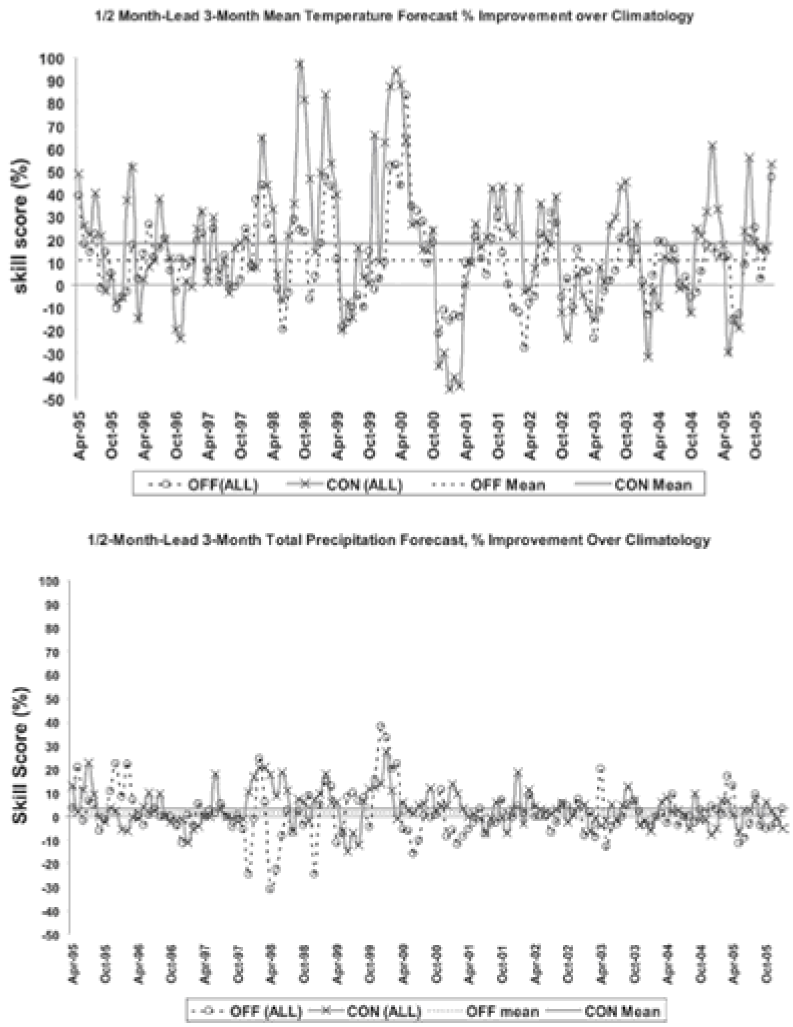

FIGURE 3.16 Objective forecasts tend to perform better than forecasts with a subjective element. Lines represent the percentage improvement of forecasts relative to climatology for a 3-month forecast issued with a 1/2 –month lead time. The mean skill of objective forecasts for the entire period (solid horizontal line, “CON” for “objective consolidation”) is above the mean skill of the forecasts with a subjective element for the entire period (dashed horizontal line, “OFF” for “official forecasts”). The individual forecasts (non-horizontal lines) throughout the period often indicate a similar relationship. Top panel: temperature; bottom panel: precipitation. SOURCE: Adapted from O’Lenic et al. (2008).

well as individual tools and may or may not incorporate historical forecast quality. Comparison of the official forecasts, which include the subjective intervention, against the purely objective consolidation indicates that the subjective element reduces forecast quality (O’Lenic et al., 2008), particularly during winter in the absence of a strong ENSO signal (Livezey and Timofeyeva, 2008).

A notable subjective element also exists in regional climate outlook forums (RCOFs[14]). These forums, initiated in the late 1990s by the WMO, National Meteorological and Hydrological Services (NMHSs) and other international organizations, bring together countries within a region, such as Southeastern South America or the Greater Horn of Africa, to develop a consensus outlook for the climate of the upcoming season. Seasonal climate predictions from the participating NMHSs are discussed in conjunction with those from international centers. The inputs are not combined objectively or systematically, although they do often consider past forecast quality in the discussions. Some analyses suggest that the subjective element of the process causes the forecasts to be quantitatively less skillful than if the input predictions were combined more objectively (Berri et al., 2005).

Efforts are underway to produce more consolidated inputs and other objective input tools that can be used in the RCOFs to encourage the reduction of the subjective element. One of these efforts includes the establishment of a Lead Centre of Long-Range Forecast Multi-Model Ensembles (LC-LRFMME[15]), which objectively combines the predictions contributed by the current nine Global Producing Centres (GPCs). However, associated skill information, which would presumably be provided by the WMO Lead Centre for the Long Range Forecast Verification System (LRFVS[16]), does not accompany these forecasts primarily because this model performance information is not provided by the GPCs. The GPCs also do not readily provide access to the historical model data that would allow users to evaluate the performance for themselves. In terms of other objective input tools, one that has been increasingly used in the RCOFs is the Climate Predictability Tool[17], which allows forecasters to develop statistical predictions that use as input either observed precursors or dynamical model output.

Objective Combination of Predictions

As discussed in the previous sections, the few studies that have compared objective forecasts and the subjective forecasts to which they contribute indicate that the subjective element degrades the quality of the objective "first guess" (Berri et al., 2005; O'Lenic et al., 2008). But beyond that, objective methods allow a forecaster to demonstrate how the forecast would have performed in the past given new prediction inputs, which is not possible if the inputs are subjectively combined.

Statistical and dynamical predictions each have their own merits and should not necessarily be viewed as competitors. It is nonetheless desirable to compare the performance of statistical and dynamical tools when both exist for a given prediction target. This comparison can serve two purposes if there is a clear difference in performance: first, it may indicate that an important process is missing from one of the prediction approaches, and second, it may indicate that one of the predictions be given greater weight in the final forecast. Several studies have shown that statistical and dynamical methods have comparable quantitative skill for specific forecast targets such as ENSO (e.g. Saha et al., 2006) or precipitation in some parts of the world (e.g. Moura and Hastenrath, 2004). In other parts of the world, such as the United States, statistical and dynamical information bring complementary information (e.g. Saha et al., 2006).

[14] *http://www.wmo.int/pages/themes/climate/consensus_driven_predictions.php*
[15] *http://www.wmolc.org/*
[16] *http://www.bom.gov.au/wmo/lrfvs/*
[17] *http://iri.columbia.edu/climate/tools/cpt*

The same can be said for different statistical predictions, such as those capturing ENSO teleconnections compared to those isolating recent trends (e.g. Livezey and Timofeyeva, 2008). In addition, considerable value can be gained by employing the two approaches together, such as model output statistics (MOS, see "Correlation and Regression" section in this chapter), which refers broadly to the statistical correction of dynamical models. MOS techniques can be used to correct systematic biases of dynamical models by translating the aspects of the observed variability that the model captures correctly into something that more closely resembles the observations (e.g. Feddersen et al., 1999; Landman and Goddard, 2002; Tippett et al., 2005).

By far, the greatest boost to objective combination of prediction inputs has come through advances in multi-model ensembles (MME). These advances within the climate community have been particularly rapid since the advent of publically available archives of model data, such as the DEMETER dataset for seasonal-to-interannual predictions (Palmer et al., 2004), with many decades of hindcasts, and the Coupled Model Intercomparison Project v3 (CMIP3) that provided the data of the climate change simulations of the 20th century and projections of the 21st century summarized in the 4th Assessment Report of the IPCC (IPCC, 2007). Although in each case the databases contain coupled ocean-atmosphere models with similar external forcing and/or initial conditions, the dynamical cores of the models and their physical parameterizations differ. The premise holds that although models have deficiencies, they do not all have the same deficiencies. Thus, combining models brings out the robust information they have in common and reduces the individual or random biases that they do not share, which can provide more reliable forecast information. By allowing scientists from all over the world to access a common set of models from different modeling centers, results are easier to compare and possible to replicate.

One result derived from these archives is that there is no single best model; one model may be best in some aspect, but turning to another aspect will highlight a different model (Gleckler et al., 2008; Reichler and Kim, 2008). Furthermore, it has been generally found that the multi-model mean outperforms the individual models (Hagedorn et al., 2005; Gleckler et al., 2008). Assigning weights to the individual models according to their historical performance (Rajagopalan et al,. 2002) can further improve upon the skill of MME relative to the simple model mean, provided that a sufficient number of hindcasts exist to distinguish the relative performance between models, i.e., about 40–50 years. Due to the need to fully cross-validate the weights assigned to models in the combination, it becomes difficult to improve upon the simple multi-model mean for MME with shorter hindcast histories (DelSole, 2007). The degree to which performance can be improved, both in terms of mean error reduction and probabilistic reliability, depends on the number of models involved, with more models yielding a higher quality MME (Robertson et al., 2004). However, it is not clear at what number the incremental benefit from adding more models begins to plateau. The magnitude of the benefit varies with the forecast target, including variable, region, and season, and with the quality of the individual models that contribute.

The wide community involvement in MME has shown that:

- All models do have their deficiencies; the one weak point in the premise of MME is that models often do contain some common biases (Gleckler et al., 2008). It therefore makes good sense to calibrate models in terms of both their mean and variability to the greatest extent possible, prior to combination (e.g. Hagedorn et al., 2005).
- Hindcast records are necessary to assess model performance prior to its inclusion in an MME. The hindcast may not be long enough for the purposes of weighting models, but it

needs to be long enough to vet the realism of the model's mean state and variability relative to other models in the MME suite because poor models will degrade forecast quality.

- Forecasts that objectively combine a number of prediction inputs allow information with different strengths and weakness to be distilled and yield more robust and reliable results. The prediction inputs can include statistical models, dynamical models, and the combination of the two. The main weakness of MME is a lack of design behind the specific models included; MME usually draws on whatever respectable models are available, and thus does not necessarily span all uncertainties in model physics.

Consideration of End User

Other approaches exist for going beyond the quality of a given forecast or model prediction to determine its value to a potential user. The provision of quantitative, probabilistic outlooks of societally-relevant variables can increase the use of climate forecasts even if the underlying quality were unchanged. Although seasonal climate forecasts are now commonly issued as probabilities for pre-defined categories (Barnston et al., 1999; Mason et al., 1999), those categories may not align with the risks and benefits of many decision makers. Additionally, users of the climate forecasts, from sectoral experts to the media, are often interested in relatively high resolution information that can be relevant to local concerns, even if it means reduced accuracy of the information. This information mismatch is one of the most commonly cited reasons for not using seasonal forecasts (e.g. CCSP, 2008). Good quality intraseasonal-to-interannual forecasts are only a starting point. In order for forecast information to be incorporated into climate risk management and decision making, it has to be in an appropriate format, at an appropriate space and time scale, and of the right variables to mesh with the decision models it is to inform.

One way to address the information mismatch between the coarse spatial resolution of global seasonal climate forecasts and the high-resolution needs of the end user is to use downscaling techniques. In statistical downscaling, the global climate forecast provides the input parameters for an empirical model with high spatial resolution. In dynamical downscaling, the global forecast is used to provide lateral boundary conditions to a high-resolution nested regional atmospheric model. Although downscaling has been used extensively in climate change research, its use on ISI timescales has been more on an exploratory basis. With increases in computing power, global climate models are starting to close the gap with the fine spatial resolution needs of the end user. However, there is still a window of a decade or so during which downscaling techniques will continue to add significant value to the dissemination of ISI forecasts.

Recent research has opened other possibilities of providing richer seasonal climate information. For example, the provision of the seasonal forecast as the full probability distribution as opposed to fixed, relative categories permits the determination of probabilistic risk of some decision-specific threshold (e.g. Barnston et al., 2000). Or, one may desire the characterization of the weather within the climate, such as the likely number of dry spells of a given duration. In some cases, certain weather characteristics of the seasonal climate may even be more predictable than the seasonal totals (e.g. Sun et al., 2005; Robertson et al., 2009). Similarly, Higgins et al. (2002, 2007) have documented how the character of daily weather changes over the United States during ENSO events. This information could complement forecasts of the seasonal mean in ENSO years, particularly for the winter season, and provide true forecasts of opportunity (Livezey and Timofeyeva, 2008) if it were packaged and communicated in that manner.

Users can improve the application of forecast information if they are made aware of instances of conditional forecast skill (Frias et al., 2010) or forecasts with no skill. As shown in Figure 3.13, forecasts are often more skillful during ENSO events, which could guide decision makers to selectively use forecast information as part of their planning. Likewise, there may be certain regions or situations for which forecasts, or specific improvements to the building blocks of forecasts, offer little or no skill. For example, information on soil moisture can contribute to predictions of air temperature (see the soil moisture case study in Chapter 4; Figure 4.11), but the improvements are limited to certain key regions and seasons. In regions and seasons for which there is no forecast skill, or in situations where there is no forecast signal, operational centers can still provide a useful service through the issuance of information on the historical range of possible climate outcomes (i.e., climatology).

The difficulty for forecast centers in producing tailored forecasts is that what is needed is often specific to a particular problem, which in turn depends on the sector and location. This can be difficult for national or even regional forecast centers to provide on an operational basis. If the forecast data and the associated history are openly available, the tailoring of the information to the specific uses may be possible. The actual tailoring may be conducted by local forecast centers, intermediaries, or directly by the end-users. The national and international forecast centers could provide sufficient information through data archives, such that forecasts can be tailored to more specific decisions. This is not a trivial activity, however. Financial and computing resources would be required to maintain such a service.

Given the investments that have already contributed to the development of intraseasonal to interannual prediction information, such an infrastructure would be a very economical extension that could dramatically increase the use of climate forecasts. For example, users would be able to evaluate past performance in terms of their own relevant metrics, or even in terms of their own local or regional observational data. Forecast centers regularly assess the quality of their prediction models or forecast systems (O'Lenic et al., 2008; Barnston et al., 2009), which is necessary for their own feedback and interaction with the climate community. However, the value of access to data for verification, tailoring, or even just formatting should not be underestimated.

EXAMPLE OF AN ISI FORECAST SYSTEM

The building blocks of ISI forecasts systems have been described in detail above. Here we provide a specific example of how these basic building blocks ultimately culminate in an ISI prediction. This example is based on current operational forecasts at NCEP. The intent here is to highlight the complexity of the problem, the multitude of inputs to the process, and where and when subjective input is used. A flow chart for the forecast production procedure is given in Figure 3.17.

The forecast production process is described in detail in O'Lenic et al. (2008) and is summarized as follows. Climate Prediction Center operational seasonal forecasts are issued on the 3rd Thursday of each month at 8:30 AM, and a team of 7 forecasters at CPC rotates throughout the course of the year in preparing these forecasts. The process begins with a

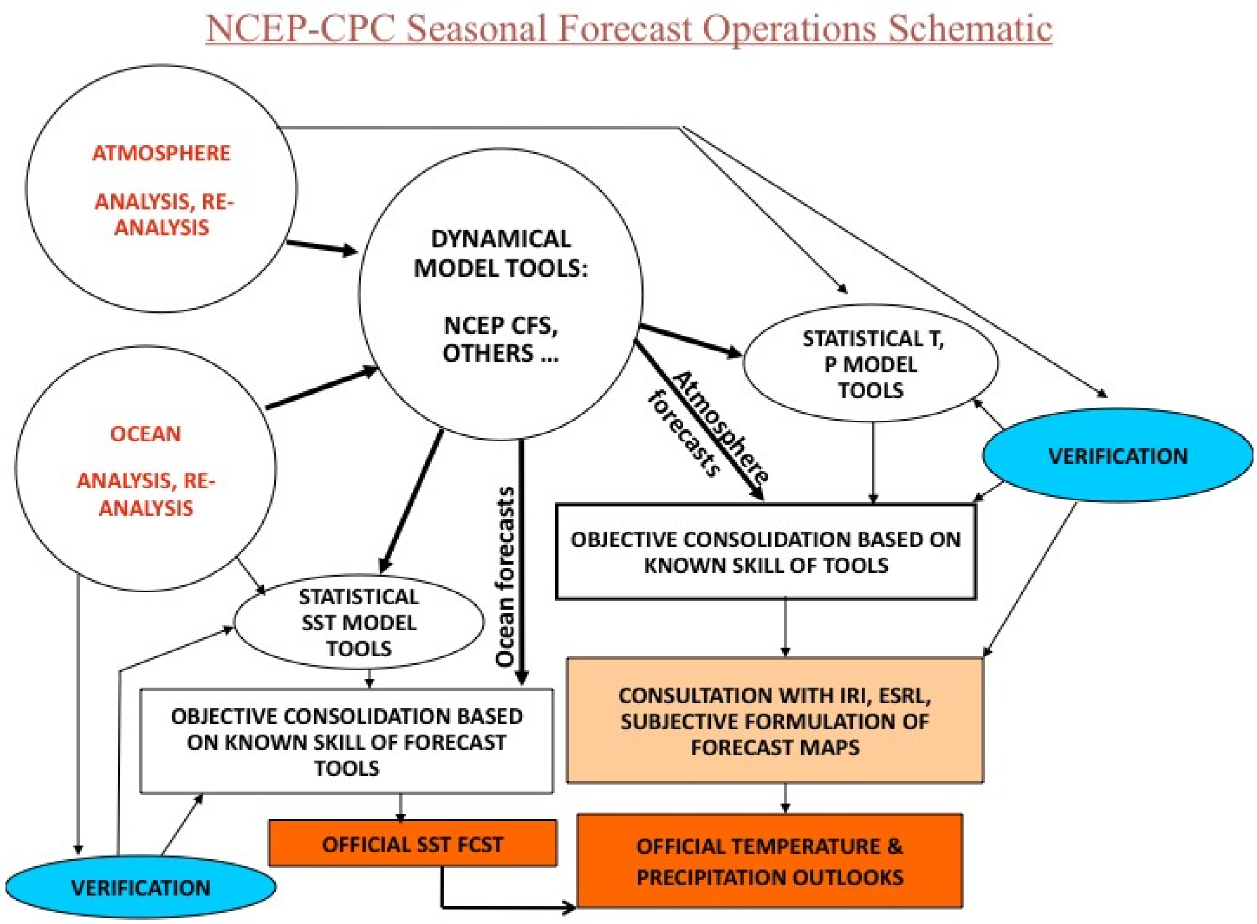

FIGURE 3.17 Graphical representation of the NCEP forecast system, showing the relationship among observations, climate system models, and data assimilation schemes as well as the steps where subjective judgment and verification are used. SOURCE: John Gottschalck, NCEP, personal communication.

comprehensive analysis of the state of the global oceans and atmosphere. This is largely based on best estimates of the current state of the climate system. Forecast tools, both CFS and statistical, are then consolidated into an objective first-guess forecast for U.S. temperature and precipitation. A telephone conference call is conducted the preceding Friday to discuss the current status of the climate system and the content of the available tools with partners in the broad climate community. Based on these discussions and the forecaster's own interpretation of the forecast tools, the forecaster manually draws draft forecast maps for all thirteen forecast leads for both temperature and precipitation. A second conference call is then used to review the draft forecast maps with governmental climate partners only. Forecast maps are finalized and processed to produce images, raw data files, and files for the National Digital Forecast Database (NDFD) for a large range of users. Finally, the lead forecaster writes a "Prognostic Map Discussion" that includes a review of the climate system, rationale for the forecasts, and an overview of the forecast maps.

POTENTIAL IMPROVEMENTS TO ISI FORECAST SYSTEMS

In examining the components of existing ISI forecast systems and current practices, a number of opportunities for improvement have been identified. These opportunities are summarized here in a structure that parallels the previous discussion by component of the ISI forecast system. In Chapter 4, there is more detail about three specific ISI forecast topics: ENSO, MJO, and soil moisture. The illustrative nature of the three case studies, together with the opportunities identified here, provide the foundation for the recommendations presented in Chapter 6.

Observations

- **Many observations that could potentially contribute to ISI predictions are not being assimilated into ISI forecast systems (see DA bullet below).**

- **The increase in the number of observations assimilated by ISI forecast systems has led to improvements in prediction. However, the attribution of these improvements to specific observations can be difficult to confirm. Also, study is required to determine the potential benefit for adopting new research observations as ongoing, operational climate observations to support ISI prediction.**

- **Targeted observations for specific climate processes that are poorly understood could improve dynamical models by providing more realistic initial conditions, improved parameterizations of sub-grid scale processes, and/or data to be used in validation.**

- **Sustained observations of the fluxes of heat and moisture between the atmosphere and ocean or between the land and atmosphere are useful for identifying biases and errors in dynamical models.** Many processes that act to couple earth system components are poorly understood and undersampled, and observations of the coupling are needed.

Statistical and Numerical Models

- **Nonlinear statistical methods can augment linear statistics.** While linear methods have been used in forecasting with moderate success in the past, positive skill is geographically dependent and primarily related to the presence of strong forcing, such as El Niño. Nonlinear techniques (e.g., nonlinear regression, neural networks, kernel methods) have been shown to be valuable in providing additional skill, especially at ISI timescales.

- **Present statistical models are not in competition with dynamical models and can be combined usefully with dynamical models.** They offer quality in certain areas where dynamical models fail and may point to areas where dynamical models can be improved.

- **Proper cross-validation is an essential tool to estimate the true forecast skill.** The use of repeated cross-validation on the same data, however, can inflate the estimated skill when models are tuned after each iteration. Such a process can result in overfitting. Data need to be divided into training and testing sets where the testing data are set aside for an unbiased estimation of true skill. It is acceptable to use subsets of the training data for model selection. However, the testing data have to be kept out of the tuning process and used for the final assessment of skill.

- **Most statistical tests assume stationarity, but the climate system is not stationary on ISI timescales.** Statistical tests exist that can address such non-stationarity (e.g., variance stabilization techniques, Huang et al., 2004). Non-stationarity can also be exploited to improve predictions.

- **Dynamical models exhibit systematic errors in their representation of the mean climate, the annual cycle, and climate variability.** While many of these shortcomings highlight opportunities for model improvement, they also contribute to forecast error. The physical processes associated with several sources of predictability (such as ENSO or the MJO) are not adequately simulated in numerical models.

- **Use of multi-model ensembles in an operational setting is still in its early stages.** MMEs need to be developed further and research on proper methods of selection, bias correction, and weighting will likely help improve the forecasts.

Data Assimilation

- **The most advanced data assimilation algorithms are predominantly focused on atmospheric observations, while the DA schemes tend to be less advanced for the ocean than for the atmosphere**. Ideally, data assimilation would be performed for the coupled Earth system. Specifically, more work is required to identify biases in the observational data and improve the ocean models so that advanced DA techniques can be applied to ocean observations.

- **Observations of many components of the Earth system are not part of DA algorithms.** Estimates of prognostic states at the land surface (e.g. soil moisture) and cryosphere (e.g., snow, sea ice extent) are generally not assimilated with operational DA schemes. Some ocean observations are assimilated as part of operational forecasts but some are not (e.g., SSH).

Forecast Verification and Provision

- **Forecast quality assessment needs to be made and communicated through multiple metrics**. Forecast quality has often been expressed through a single method (e.g., Heidke skill). Multiple metrics and graphical techniques, including ones that assess the quality of the probabilistic information, will provide a better assessment of the fidelity of the forecast system.

- **Access to archived hindcasts and real-time forecasts is required to tailor climate information to the needs of decision makers.** Information regarding forecast quality and skill varies widely among forecast systems. Comparison among systems is critical for identifying opportunities for model improvement, as well as novel combinations of forecast models that may improve quality.

- **Subjective intervention into forecasts needs to be minimized and documented.** The subjective component can limit reproducibility, restricting retrospective comparison of forecast systems. Although there are time constraints around issuing forecasts, it is helpful to have written documentation of the subjectivity of forecast preprocessing and post-processing to assess the relative performance of the inputs and outputs.

4

Case Studies

The discussions in Chapter 3 above are general in nature, comprehensively addressing the building blocks of a full ISI forecast system. To provide a more concrete flavor for many of the issues involved in ISI prediction, the present chapter focuses on three cross-cutting examples—El Niño-Southern Oscillation (ENSO), the Madden-Julian Oscillation (MJO), and soil moisture. Knowledge of the state of each of these phenomena or quantities is known to contribute to ISI forecast quality. For each process, we address the scientific basis for expecting the process to contribute to the quality of ISI predictions, the manner in which forecast systems try to realize this potential (using the elements outlined in Chapter 3), their ability (or lack thereof) to do so, and the gaps in understanding and treatment of the process that have yet to be overcome.

EL NIÑO-SOUTHERN OSCILLATION (ENSO)

On interannual timescales, the variability in the tropical Pacific is dominated by the ENSO phenomenon (Rasmusson and Carpenter 1982; among others). During the transition to the warm phase of this oscillation, there is a dynamic adjustment of heat and mass between the western and eastern tropical Pacific, producing a positive sea surface temperature anomaly (SSTA) in the eastern Pacific. Associated with this dynamic adjustment, precipitation is displaced eastward from the climatological warm pool region toward the date line, and the normally easterly trade winds weaken or even become temporarily westerly. During the cold phase, the eastern tropical Pacific SSTA is negative, the trade winds are anomalously strong, and the precipitation is tightly confined to the warm pool region of the western Pacific. Typically, the time between warm events is around 2 to 7 years; however, there is also considerable modulation of the ENSO cycle on decadal timescales. Despite the quasi-periodicity, the predictability of ENSO is largely determined by the life cycle of individual events, which depends on the memory or inertia associated with upper ocean heat content and coupled ocean-atmosphere interactions. As the event evolves there are large spatial shifts in tropical Pacific rainfall leading to large-scale changes in global circulation and precipitation (Figure 4.1). It is these changes in circulation and precipitation that make ENSO a primary source for ISI predictability in remote regions and offer the potential for decision support and risk management.

Scientific Basis for Prediction

The advent of dynamic ENSO prediction can be traced back to Bjerknes's conceptual model. While the phenomenon of the Southern Oscillation was discovered and applied to

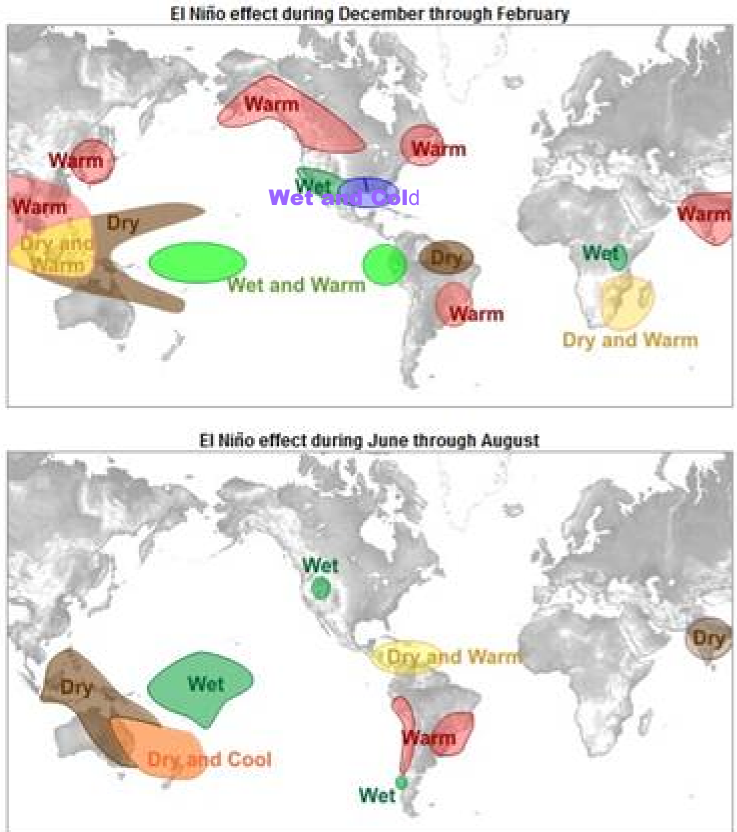

FIGURE 4.1 ENSO impacts on seasonal climate. SOURCE: Adapted from CPC/NCEP/NOAA.

seasonal prediction of the Indian monsoon in the early 1930s (Walker, 1932), the modern concept of ENSO was established three decades later by the pioneering work of Bjerknes (1966, 1969) who visualized the linkage between the atmospheric Southern Oscillation and oceanic warming in the eastern Pacific (El Niño) via tight coupling among the sea surface temperature, surface winds, and precipitation. Bjerknes's work laid a physical basis and a conceptual framework for ENSO prediction (Neelin et al., 1998; Burgers et al., 2005). The coupled instability mechanism described by Bjerknes (1969) works as follows. If a positive eastern equatorial SSTA exists, the temperature gradient between the eastern Pacific and the western Pacific is reduced, which then produces a weakening of the easterly trade winds, augmenting the warming in the eastern Pacific. The additional warming in the east further weakens the trade winds, constituting a coupled ocean–atmosphere positive feedback. A reversal of the argument explains the growth of a cold event. This positive feedback between the SSTA and the atmospheric wind anomaly leads to a continually growing anomaly, although it does not provide an explanation for what causes the transition from one extreme state to the other.

About the same time, independent but related theoretical findings coincidently emerged, i.e., the equatorial wave theory in an unbounded equatorial atmosphere (Matsuno, 1966) and in a finite ocean basin (Moore, 1968). These theoretical findings stimulated rapid development of the equatorial oceanography. In the 1970s and 1980s, Bjerknes's conceptual model was transformed and expanded into theoretical and simple dynamical models, which advanced understanding of the basic ENSO dynamics, including coupled ocean-atmospheric instability, thermocline adjustment by oceanic waves, and mechanisms possibly sustaining ENSO and causing irregularity, non-stationarity, and skewness. The theoretical work has led to the understanding of destabilized equatorial waves (Philander et al., 1984) and a theory for the transition between warm and cold states (e.g., the so-called delayed or recharge oscillators, Suarez and Schopf, 1988; Battisti and Hirst, 1989; Jin, 1997; Kirtman, 1997). Moreover, this theoretical understanding clearly delineates the scientific basis for ENSO prediction, namely that sub-surface ocean thermal structure or thermocline displacements are pre-cursors for ENSO events and can be used to make forecasts.

Forecast System Methodologies

The pioneering El Niño forecast was made with an intermediate-complexity coupled ocean-atmosphere model (Cane et al., 1986). Since then, a variety of methods for ENSO prediction have flourished, including purely statistical techniques (e.g., Graham et al., 1987), combinations of dynamical and statistical models, and purely dynamic models (e.g., Ji et al., 1994; Kleeman et al., 1995; Rosati et al., 1997; Berhinger et al., 1998; Stockdale et al., 1998; Schneider et al.,1999; Kirtman, 2003; Keenlyside et al., 2005; DeWitt, 2005; Gualdi et al., 2005; Jin et al., 2008). All of these forecasting strategies rely on four well-defined basic ingredients of a state-of-the-art prediction system: (1) a dynamical model (i.e., coupled GCM) that consists of a series of mathematical expressions or statistical relationships that represent the physical laws that govern how the ocean and atmosphere behave and interact; (2) an observing system to provide input for initializing, developing, and verifying both dynamic and statistical forecast systems; (3) initial conditions or an estimate of the current state of the climate system, usually based on sophisticated data assimilation systems; and (4) a series of retrospective forecasts for calibration and assessing quality. The models used in these forecasting strategies have varying degrees of sophistication and diverse initialization schemes.

The importance of the observing system for the development of the statistical and dynamical prediction capability cannot be overstated. NOAA's Equatorial Pacific Climate Study program deployed the earliest buoys along 110°W and then along 140°W, starting with the prototype Atlas moorings in 1984–85. This pioneering observational effort eventually led to the establishment of the mooring array along the entire equatorial Pacific basin. Development of the Tropical Atmosphere Ocean (TAO) array (Hayes et al., 1991) was further motivated by the 1982–1983 El Niño event, the strongest of the century up to that time, which was neither predicted nor detected until nearly at its peak. The event highlighted the need for real-time data from the tropical Pacific for monitoring, prediction, and improved understanding of El Niño. Eventually, the TOGA-TAO array (now renamed TAO/TRITON, see "Ocean Observations" section in Chapter 3) was completed in 1994 to include 70 moorings. The moorings provide winds, sea surface temperature, relative humidity, air temperature, and subsurface temperature at 10 depths in the upper 500 m. Five moorings along the equator also measure ocean velocity. These data remain key components of operational ENSO prediction and continue to provide critical data for understanding ENSO-related physical processes.

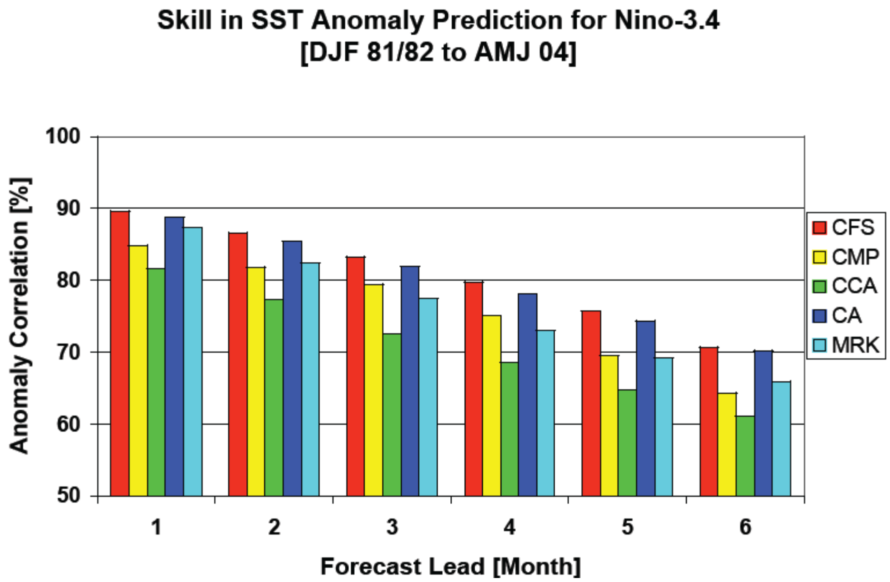

FIGURE 4.2 Nino3.4 correlation coefficient (predictions versus observations) for retrospective forecasts plotted as a function of lead time. The red and yellow bars correspond to dynamical models: the Climate Forecast System (CFS) is a state-of-the-art model developed in the mid-2000s (Saha et al. 2006), while the Coupled Model Project (CMP) prediction is older and was developed in the mid 1990s (Ji et al., 1995). CCA, CA, and MRK correspond to statistical models (Canonical Correlation Analysis, Constructed Analogues, Markov; see Appendix A). The figure highlights two points: (1) comparing the red and yellow bars indicates how coupled dynamical models have improved for this particular metric over the last two decades and (2) the statistical methods and the dynamical model methods are quite competitive with each other. Identical to Figure 3.14. SOURCE: Adapted from Saha et al. (2006)

Forecast Quality

ENSO forecast quality has clearly improved from initial attempts in the late 1980s. Undoubtedly, this improvement is due to better models, enhanced observing systems, and better use of observational estimates through improved data assimilation techniques. In the late 1980s and early 1990s the state-of-the-art dynamical model forecast quality for ENSO markedly lagged behind most statistical techniques (Anderson et al. 1999). During the mid to late 1990s dynamical models began to make much better use of observational estimates of the state of the sub-surface ocean in the tropical Pacific, which led to notable improvements in forecast quality. Most recently, the dynamical models have become quite competitive with the statistical models in predicting the state of ENSO in the tropical Pacific (see Figure 4.2) as seen by comparing the red and yellow bars (dynamic models) with the remaining bars (statistical models). The unprecedented 1997–1998 El Niño was perhaps the first real-time test of modern ENSO prediction systems and was fairly well predicted three to six months in advance using a number of different state-of-the-art coupled GCMs (e.g., Anderson et al., 2003), although the models failed to capture its large amplitude and its rapid onset (Barnston et al., 1999).
Figure 4.3 underscores the point that multi-model combinations can improve ENSO prediction skill as measured by, for example Nino3.4 correlation skill. It indicates that combining

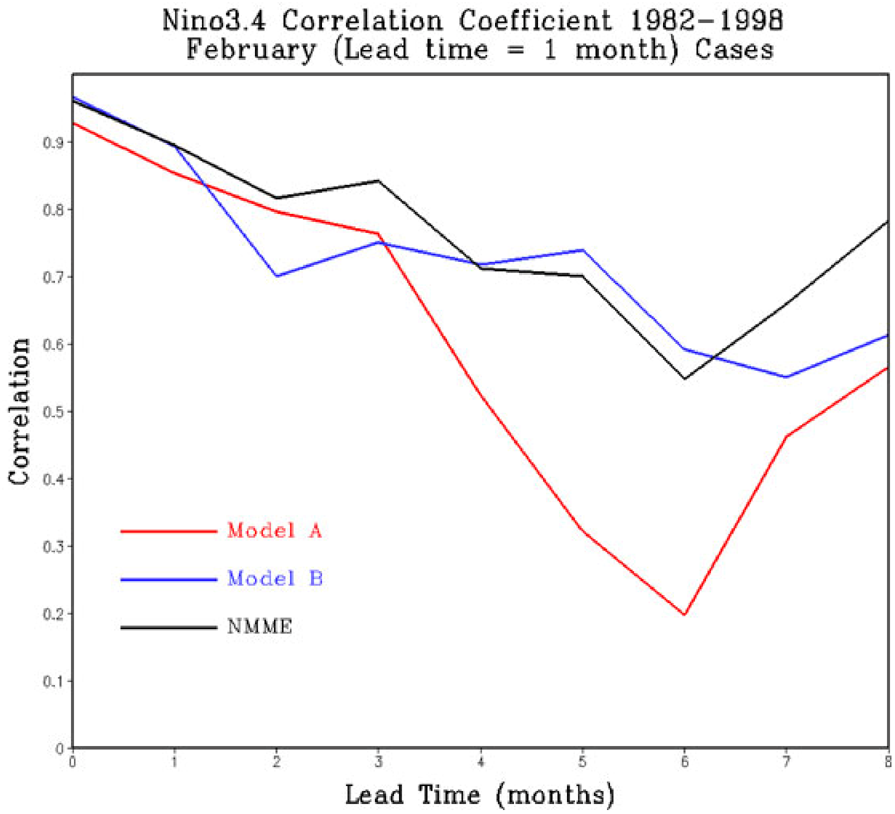

FIGURE 4.3 Nino3.4 correlation coefficient (predictions vs. observations) for two U.S. models (red and blue lines) and a multi-model ensemble (black line). The figure indicates that multi-model combinations improve this particular skill metric (black line is above the colored lines), that rapid progress on multi-model prediction in the United States can be made quickly (i.e., multi-model U.S. prediction systems are available today), and in comparison with Figures 2.11 and 3.13, that significant seasonal dependence exists in the correlation. The correlation shown here shows a rapid drop in boreal spring. This is commonly referred to as the "spring prediction barrier," but it is unknown if it is due to model errors or is a fundamental property of the climate system. SOURCE: Kirtman and Min (2009).

two U.S. national models improves this particular skill metric and suggests that the United States is indeed well positioned to make multi-model operational ENSO predictions. Finally, this figure, consistent with Figures 2.11 and 3.13, indicates that there is significant seasonality in the correlation; this seasonality is not fully understood.

Most recently, the advances in ENSO prediction include: (1) the recognition that forecasts need to include quantitative information regarding uncertainty (i.e., probabilistic prediction) and that verification needs to include skill measures for probability forecasts (Kirtman, 2003); and (2) that a multi-model ensemble strategy (see Figures 2.9, 3.9, and 4.3) may be the best current approach for adequately quantifying forecast uncertainty. Although an MME strategy represents the "best current approach" for estimating uncertainty, it should be noted an MME forecast does not quantify the relationship between individual model errors and forecast uncertainty. The spread among the ensemble members can only provide a relative estimate of the forecast uncertainty.

How individual model errors contribute to forecast uncertainty is an active area of research. In terms of quantifying how model error contributes to forecast uncertainty, a number of international projects have been organized, among which the most comprehensive projects are the European Union-sponsored Development of a European Multi-model Ensemble System for Seasonal to Inter-Annual Prediction (DEMETER; Palmer, 2004), the Climate Prediction and its Application to Society (CliPAS) project, sponsored by the Asian-Pacific Economic Cooperation (APEC) Climate Center (APCC), and the Climate System Historical Forecast Project (CHFP, Kirtman and Pirani, 2009). These hindcast datasets provide a test-bed for assessing forecast quality and forecast uncertainty based upon uncertainty in model formulation. The multi-model approach has proven to be more skillful than any single model (Krishnamurti et al., 1999, 2000; Doblas-Reyes et al., 2000; Shukla et al. 2000; Palmer et al., 2000; Jin et al., 2008; Kirtman and Min, 2009). For example, the Mean Square Skill Score (MSSS) against climatology is a good way to assess the deterministic skill of ENSO forecasts. In particular, it gives a measure of error relative to the signal strength; in this case, the signal is the forcing on the atmosphere by the ocean. State-of-the-art MSSS for Nino 3.4 SST at 5 months lead-time is about 0.7 for the best single models, and 0.75 for multi-model combinations. Despite the improvements associated with the multi-model approach, the ad hoc nature of the strategy is troubling (Blanke et al., 1997; Kessler and Kleeman, 2000; Boulanger et al., 2001; Lengaigne et al., 2003; Eisenman et al., 2005; Vecchi et al., 2006; Gebbie et al., 2007; Jin et al., 2007; Kug et al., 2008; Kug et al., 2009; see "Multi-Model Ensembles section in Chapter 3).

Comparing the correlation skill (i.e., correlation between forecast and observations) of different forecasting systems on common sets of forecasts shows that there has been slow but steady progress over the last 10 years. Again, this is highlighted by comparing the red and yellow bars in Figure 4.2. Improvement in ENSO SST forecasts is expected to continue in the years ahead—the errors in today's forecasting systems are still substantial and statistical post-processing and calibration improve forecast quality.

Consideration of a set of individual forecasts shows that today's models provide good guidance as to the future evolution of SST, but relatively large errors can still occur. Some of the failures in the past might be related to an inadequately observed initial state, and certainly some of these errors are related to model fidelity.

The use of multi-model ensembles can give a definite boost to the quality compared to that obtained by a single model (e.g., Fedorov and Philander, 2001; Hagedorn et al., 2005; Codron et al., 2001; Guilyardi, 2006; Zhang et al., 2008; Jin et al., 2008; Kirtman and Min, 2009), and multi-model approaches to ENSO prediction are encouraged. Nonetheless, improvement of the individual models is strongly needed to improve the quality of future forecasts (single or multi-model). Typically, models are used to produce ensemble forecasts in order to quantify uncertainty and estimate higher moments. Forecast spread does vary according to season and ENSO phase in the models, but the relationship of forecast error to model spread is weak. For real applications, any model forecast needs to be post-processed in some way. Probabilistic verification of calibrated and post-processed forecasts is to be encouraged, but at the moment the information content of the forecasts is thought to be very largely dominated by the first moment, i.e. the ensemble mean.

Gaps in Understanding

Our ability to predict ENSO with dynamical models has dramatically improved from the mid-1980s to the beginning of the 21st century. Past improvements were due to the convergence of many factors, including a theoretical understanding of coupled ocean-atmosphere dynamics, improvement of the coupled model forecast systems, and international efforts to observe and monitor conditions in the tropical Pacific. However, basic questions regarding our knowledge of physical processes in the tropical Pacific remain open challenges in the forecast community. For instance, it is unclear how the MJO, westerly wind bursts (WWBs), intra-seasonal variability or atmospheric weather noise influence predictability of ENSO (e.g., Thompson and Battisti, 2001; Kleeman et al., 2003; Flugel et al., 2004; Kirtman et al., 2005). It has been suggested that enhanced MJO and WWB activity was related to the rapid onset and the relatively large amplitude of the 1997–1998 event (e.g., Vecchi and Harrison, 2000; Eisenman et al., 2005). However, more research is needed to fully understand the scale interactions between ENSO and the MJO and the degree that MJO/WWB representation is needed in ENSO prediction models to better resolve the range of possibilities for the evolution of ENSO (Wittenberg, 2004). Typically, prediction systems do not adequately capture the differences among different ENSO events (Goddard and DeWitt, 2005). In essence, the prediction systems do not have a sufficient number of degrees of freedom for ENSO as compared to nature. There are also apparent decadal variations in ENSO forecast quality (Balmaseda et al., 1995; Ji and Kousky, 1996; Kirtman and Schopf, 1998; Barnston and Tippett, 2009), and the sources of these variations are the subject of some debate. It is unclear whether these variations are just sampling issues or are due to some lower frequency changes in the background state (see Kirtman et al. 2005 for a detailed discussion). Chronic biases in the coupled models in their mean states and intrinsic ENSO modes remain, and it is believed that these biases have a deleterious effect on SSTA forecast quality and the associated teleconnections. Some of these errors are extremely well known throughout the coupled modeling community. Three classic examples, which are likely interdependent, are (1) the so-called double ITCZ problem, (2) the excessively strong equatorial cold tongue typical to most models, and (3) the eastern Pacific and Atlantic warm biases endemic to all models. Such biases may limit our ability to predict seasonal-to-interannual climate fluctuations and could be indicative of errors in the model formulations. Finally, it remains unclear how changes in the mean climate will ultimately affect ENSO predictability (Collins, 2000).

In addition, procedural issues remain when initializing, making, and verifying ENSO forecasts. Quantifying the relationship between model uncertainty and forecast uncertainty is an area of active research. For instance, in an individual model, it is possible to introduce stochastic physics schemes in order to approximate the uncertainty arising from the model parameterizations of unresolved sub-grid scale processes (Buizza et al., 1999; Shutts, 2005; Bowler et al., 2008). These approaches are operationally used by ECMWF and the United Kingdom Meteorological Office for medium-range forecasts and are being tested on the seasonal-to-interannual prediction problem.

Other advancements include using novel data sets to initialize forecasts, particularly involving ocean data (Alves et al., 2004); moreover, other research indicates that forecast initialization strategies that are implemented within the framework of the coupled system as opposed to the individual component models may also lead to substantial improvements in skill (Chen et al., 1995).

MADDEN-JULIAN OSCILLATION (MJO)

Scientific Basis for Prediction

The dominant form of intraseasonal atmospheric variability, particularly in terms of rainfall generation and global reach of influence, is most often referred to as the Madden-Julian Oscillation (MJO; also known as the 30–60 day, 40–50 day, and intraseasonal oscillation (ISO), after its discoverers, Madden and Julian, (1971, 1994, 2005)). The left panels of Figure 4.4 (see also Figure 2.6) illustrate the space-time structure of rainfall and low-level winds in the tropics associated with an MJO "event" during boreal winter, with the interval between maps being 12.5 days. These maps illustrate the eastward propagation of the MJO's large-scale tropical rainfall anomalies. In conjunction with these rainfall anomalies are baroclinic wind anomalies, with upper tropospheric divergence (convergence) occurring in conjunction with positive (negative) rainfall anomalies and vice versa for the lower troposphere (see Waliser, 2006). In addition, it can be seen that there is a significant modulation by the relatively warmer (cooler) eastern (western) hemisphere background state, with the large rainfall anomalies developing and propagating (~5 m s^{-1}) over the warm waters of the Indian and west Pacific Oceans. Once the disturbances reach the vicinity of the International Date Line and the cooler eastern Pacific Ocean equatorial waters, the convection tends to subside and propagate southeastward into the South Pacific Convergence Zone. Beyond the Date Line, the disturbance continues to propagate eastward (~15–20 m s^{-1}) and tends to be evident only in the near-equatorial wind field. (Hendon and Salby, 1994).

The off-equatorial structure of the MJO is also important, especially in relation to its connections to mid-latitudes. For example, associated with the positive near-equatorial rainfall anomalies are upper-level cyclonic (anticyclonic) gyres to the northeast and southeast (northwest and southwest) centered at latitudes of about 20° (Rui and Wang, 1990; Hendon and Salby, 1994). These tropical heating and subtropical circulation anomalies act as Rossby wave sources for mid-latitude variability (e.g., Weickmann, 1983; Liebmann and Hartmann, 1984; Weickmann et al., 1985; Lau and Phillips, 1986; Sardeshmukh and Hoskins, 1988; Berbery and Nogues-Paegle, 1993). Such connections with the extra-tropics have important ramifications for mid-latitude weather variability, regime changes, and forecasting capabilities (e.g., Ferranti et al., 1990; Higgins et al., 2000; Jones et al., 2004b). For example, Figure 4.5 illustrates the MJO influence on the mid-latitude circulation and its relationship with rainfall anomalies in the Pacific Northwest.

The characteristics of the intraseasonal variability driven by the MJO tend to be most strongly exhibited during the boreal winter and spring when the Indo-Pacific warm pool is centered at or near the equator. In the boreal summer, the MJO is still present, although its spatial variability and propagation characteristics are modified by the changes associated with the annual cycle. The right panels of Figure 4.4 illustrate the space-time structure of the MJO in boreal summer (for more in-depth observational descriptions see recent reviews by Goswami, 2005; Hsu, 2005). Note that the summertime manifestation of the MJO is often referred to as the Intraseasonal Oscillation (ISO), the boreal summer ISO, or monsoon ISO (MISO). Examination of the boreal summer rainfall maps shows that positive rainfall anomalies in the western and central Indian Ocean occur in conjunction with negative rainfall anomalies over a region extending between India and the western equatorial Pacific. This system then appears to

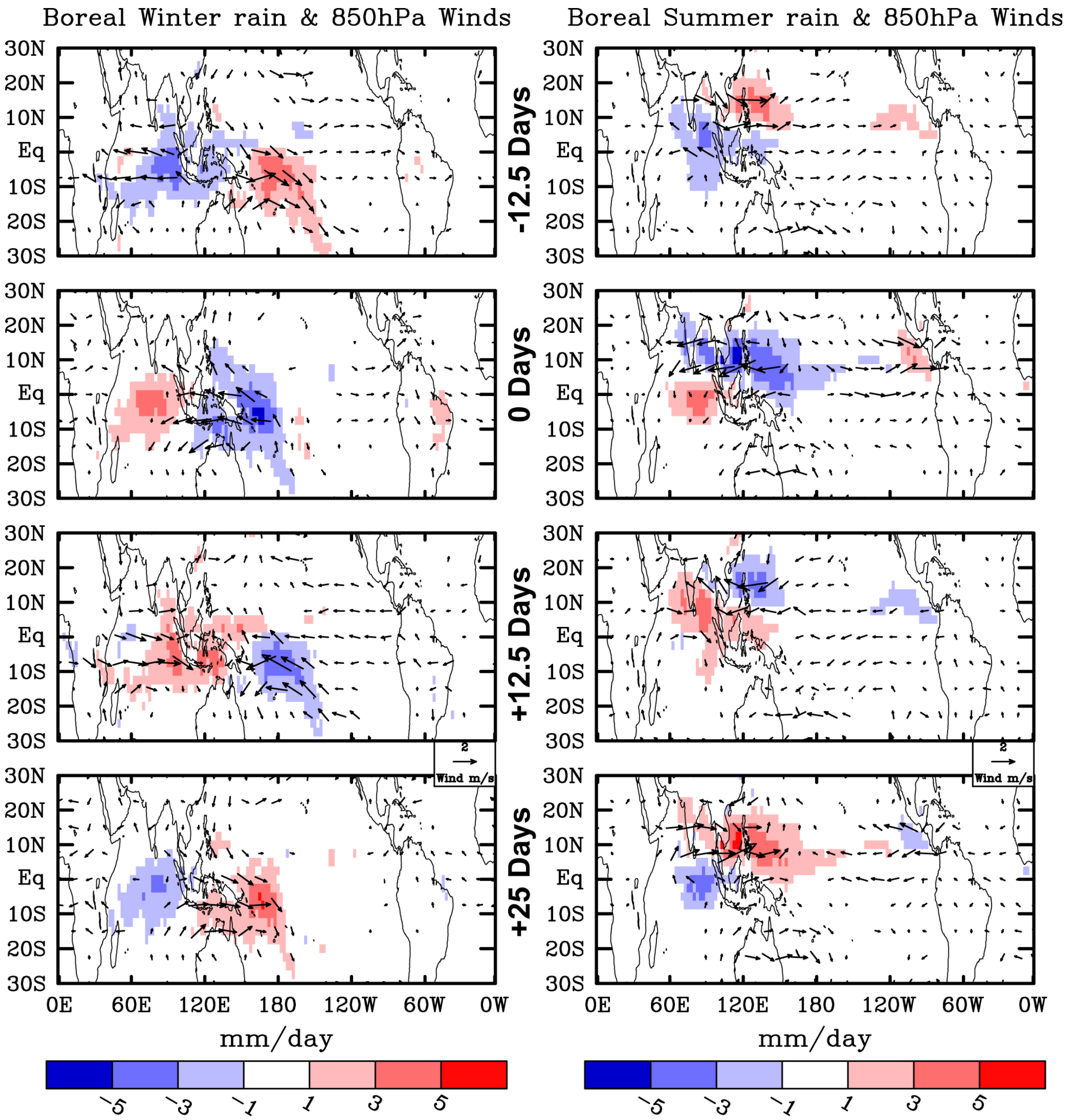

FIGURE 4.4 Characteristic circulation and precipitation patterns associated with an MJO event. Anomalous winds at 850 hPa (vectors) and precipitation anomalies (red: wet anomaly; blue: dry anomaly) are shown for the periods prior (top panels, 12.5 days preceding the event), during (second from top panels), and after an event (bottom two panels, 12.5 and 25 days following the event). Left panels show an event in the boreal winter (November–April); right panels are for the boreal summer (May–October). SOURCE: Waliser (2006).

propagate both eastward and northward (Yasunari, 1979; Lau and Chan, 1986; Lawrence and Webster, 2002; Hsu, 2005), similar to the boreal winter case. These large-scale rainfall variations have important implications for Asian monsoon onset and breaks.

While the diagrams in Figure 4.4 illustrate what might be considered typical winter and summer MJO events, it is important to recognize that these events have considerably more complexity in reality and exhibit significant interannual variability. For example, the study by Wang and Rui (1990), and later by Jones et al. (2003), have further diagnosed the "synoptic climatology" of tropical MJO events, including their seasonal modulation. These studies show that boreal winter events display considerable variation in the longitudes at which the convection develops and dissipates. Moreover, it is well known that the convection associated with MJO events typically propagates further east during El Nino events (e.g., Kessler, 2001). Some of these features are illustrated in Figure 4.6. For the boreal summer case, Kemball-Cook and Wang (2001) show that there is a systematic intraseasonal change in the spatial structure and propagation characteristics of the MJO. In the early part of the summer (e.g., May-June), the

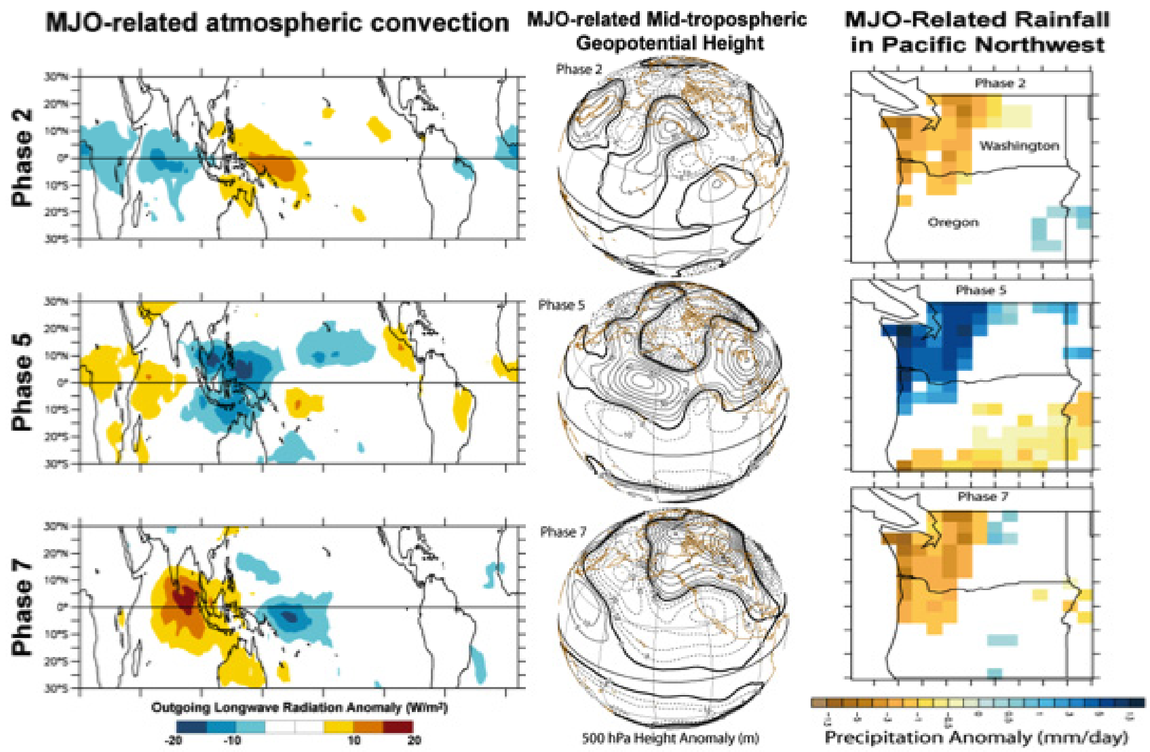

FIGURE 4.5 Example of the relationship among tropical outgoing long-wave radiation (OLR, left column), which is used to define the phase of the MJO, wintertime (JFM) 500-hPa geopotential height anomalies (middle column), and precipitation anomalies (right column). For example, Phase 5 (the middle row) of the MJO exhibits enhanced convection over the Maritime Continent that is accompanied by deep-troughing in the mid-troposphere over the North Pacific and enhanced precipitation in the Pacific Northwest. SOURCE: Adapted from Bond and Vecchi (2003).

off-equatorial variability is generally found west of Southeast Asia and the Maritime Continent, while in the later part of the summer, it expands to include much of the northwestern Tropical Pacific.

By the early 1990s, many physical characteristics of the MJO were documented and a number of reproducible features were recognized as occurring from one event to another as well as in events from one year to the next. In addition, theoretical and modeling studies suggested that the coupling observed between organized convection and low-frequency equatorial waves (e.g., Kelvin, Rossby) were responsible for the slow, eastward propagation of the MJO and thus suggested an unexploited form of intraseasonal predictability (e.g., Wang et al., 2005). Given this and the emerging knowledge of the interactions of the MJO with other features of our weather and climate (e.g., Lau and Waliser, 2005; Waliser et al., 2005; Tian et al., 2007; 2008; Wong and Dessler, 2007), it was an obvious step to seriously consider MJO forecasting.

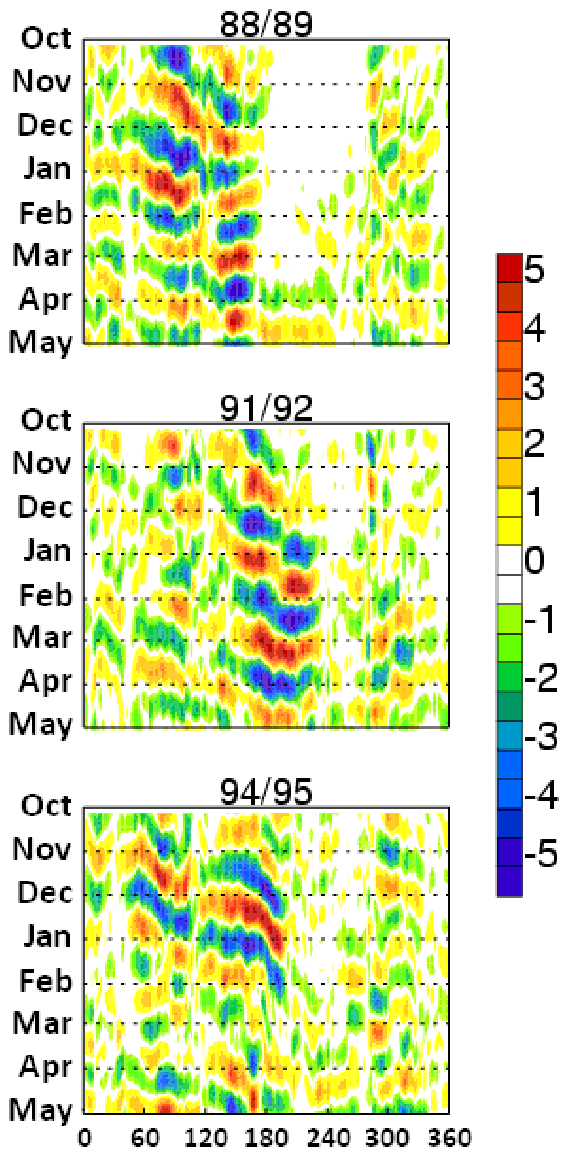

FIGURE 4.6 Precipitation anomalies (mm/day) observed for three winter periods (1988–1989, 1991–1992, and 1994–1995) illustrating the year-to-year differences in the location, magnitude, and rate of propagation of anomalies in convection near the Equator. The colors indicate OLR anomalies, with red corresponding to suppressed convection and blue corresponding to enhanced convection. Essentially, large differences can exist among MJO events in different years. SOURCE: Courtesy of D. Waliser and B. Tian.

Predictability Estimates and Forecast System Developments

Given that numerical weather and climate models typically have had, and still for the most part have, a relatively poor representation of the MJO (e.g., Slingo et al., 1996; Waliser et al., 2003a; Allen et al., 2005; Slingo et al., 2005; Kim et al., 2009), a natural avenue to consider for establishing an MJO forecasting capability was the development of empirical models. There were a number of different approaches and data sets used in these empirical studies, with most relying on linear approaches and quantities such as outgoing longwave radiation (OLR) and/or upper-level circulation quantities (von Storch and Xu, 1990; Waliser et al., 1999a; Jones et al., 2004a; Mo, 2001; Wheeler and Weickmann, 2001; Goswami and Xavier, 2003; Webster and Hoyas, 2004). The upshot of these studies is that empirical models demonstrate some accuracy in the prediction of the MJO on the order of 15–25 days or more, depending on the spatial scale and quantity being predicted. However, as with any empirical model, these models are limited in the totality of the weather and climate system they can predict, their ability to adapt to arbitrary conditions, and their ability to take advantage of known physical constraints.

To date, the majority of dynamical models still exhibit significant shortcomings in terms of their MJO simulation, particularly if pressed to do operational prediction. There have been a few models or versions of models that have demonstrated success at representing a number of the principal features of the MJO (Slingo et al., 1996; Sperber et al., 1997; Waliser et al., 1999b; Kemball-Cook et al., 2002; Maloney, 2002; Fu et al., 2003; Zheng et al., 2004; Kim et al., 2009). This level of model success was suggestive that better representation of the MJO in operational prediction models was likely to lead to improved quality of intraseasonal predictions (Waliser et al., 2003b; 2003d; Liess et al., 2004) with the additional indication that ocean coupling may yield further enhancements (Fu et al., 2006; Zheng et al., 2004; Zhang et al., 2006; Woolnough et al., 2007; Pegion and Kirtman, 2008).

Forecast Performance

Based on the results from empirical and dynamical modeling studies discussed above, there has been ample reason to push towards an operational MJO prediction capability. However, indications from early studies carried out in the context of operations showed little prediction quality due to the poor representation of the MJO in the models (Chen and Alpert, 1990; Lau and Chang, 1992; Jones et al., 2000; Hendon et al., 2000). In general these studies found reasonable accuracy only out to about 7–10 days for MJO-related variability, and were mainly hampered by MJO variability that was too weak and/or that propagated too fast. Probably the most optimistic set of demonstrative hindcast skill experiments for the MJO were a set of Asian monsoon MJO case studies performed by Krishnamurti et al. (1990; 1992; 1995). The novel approach in these cases was that an attempt was made to filter out the "weather" time and space scales from the initial conditions and leave only the "low-frequency modes." The hindcast results demonstrated reasonable accuracy out to 3–4 weeks; however, there are some uncertainties associated with making such a technique operational and with the handling of the boundary-layer forcing (i.e. SST). More recent studies on this topic by Agudelo et al (2008) and Fu et al. (2009) raise questions as to the manner and degree that the fast time scales associated with processes such as organized convection contribute to the overall forecast error, and suggest the need for further research in this area.

Given the need for forecast capability at the intraseasonal time scale, along with the poor representation of the MJO in dynamical models, a number of real-time efforts have been developed based on empirical methods. These include some of the schemes referenced above (Wheeler and Weickmann, 2001; Jones et al., 2004a; Wheeler and Hendon, 2004; Webster and Hoyas, 2004) as well as other novel techniques such as empirical wave propagation (EWP, van den Dool and Saha, 2002) and the use of a Linear Inverse Model (LIM, Winkler et al., 2001). Some of these techniques are being utilized in operational contexts, such as in forecast products at the Australian Bureau of Meteorology (ABOM; personal communication Matthew Wheeler) and the Global Tropics Benefits/Hazards Assessment at the National Centers for Environmental Prediction (NCEP, personal communication John Gottschalck). Given the developing reliance on the empirical forecast products mentioned above, it is important to note that many of them lie outside formal prediction centers and thus are only quasi-operational.

While dynamical models have not simulated the MJO in a particularly faithful manner, there have been improvements in their performance. An excellent example of improvement in forecast model quality is given in Figure 4.7, which shows the evolution of the ECMWF hindcast skill for the MJO during the Tropical Ocean Global Atmosphere Coupled Ocean Atmosphere Response Experiment (TOGA COARE) period (1992–1993). Notable improvements made to the model relative to the MJO had mostly to do with model parameterization changes including convection, clouds, radiation, and turbulent diffusion (Tompkins et al., 2007; Bechtold et al., 2008). These model improvements have resulted in a present-day ECMWF forecast skill, as measured by anomaly correlation, of about 0.6 at lead times extending out to about 20 days for the large-scale spatial patterns of the MJO (e.g., Figure 4.4). These values are competitive with if not better than the empirical models developed to date.

Since the MJO is largely an atmospheric phenomenon, most of the predictability is thought to come from the atmospheric structure of the mode itself that tends to be relatively well captured by conventional synoptic and satellite observations available today. However, there have been few studies that have addressed which aspects of the initial conditions of the atmospheric state are important to MJO forecasts. The study by Vintzileos and Pan (2007) is an exception, as it illustrated that the prediction quality of the MJO in the NCEP CFS model was significantly more sensitive to the choice of initial conditions (in this case Reanalysis-2 versus GDAS operational) than to model resolution (in this case T62, T126, T254). These types of issues for future work are discussed in the following section.

Gaps in Understanding

Based on the sorts of forecast activities and developments described above, there have been a number of efforts to take a more systematic and cross-community/center approach to MJO forecasting and model diagnosis. This has included a number of workshops held in recent years (Zhang et al., 2001; Schubert et al., 2002; Waliser et al., 2003c; ECMWF, 2004; ICTP, 2006; Sperber and Waliser, 2008) and efforts to develop cross-center/model experimental predictions (Waliser et al., 2005). These activities in turn helped to support and guide the formation of an MJO Working Group (MJOWG) under the auspices of U.S. CLIVAR. The objectives of this working group included the development of model diagnostics (CLIVAR

Madden-Julian Oscillation Working Group, 2009) to facilitate model development, assessments and comparison, and application of these diagnostics to a number of contemporary weather and climate models (Kim et al., 2009). Figure 4.8 shows wavenumber-frequency spectra for eight models (3 coupled and 5 uncoupled) and for observations/reanalysis. Although a couple of the models exhibit sizeable intraseasonal variability, none of these models provides a robust representation of the dominance of the variability at wavenumbers 1–3 and 30–80 days. Note the 30– and 80–day designations as vertical lines. In addition, common shortcomings among the

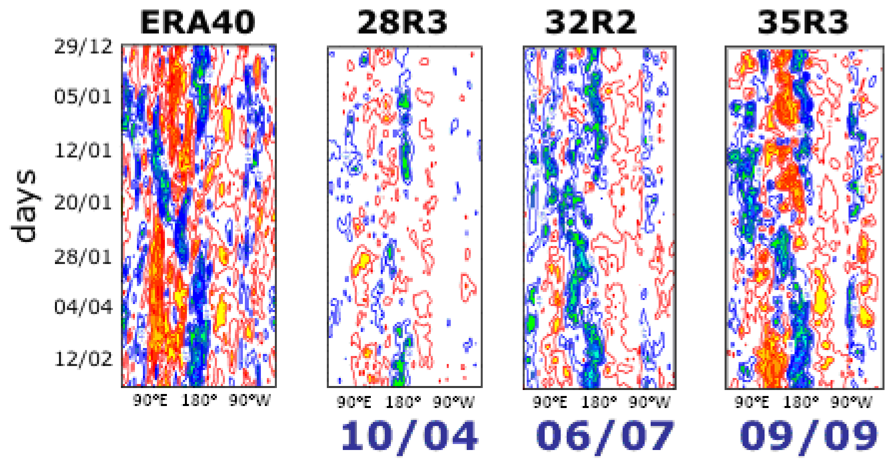

FIGURE 4.7 Improvements in forecasts of tropical outgoing long-wave radiation (OLR) from October 2004 to September 2009 over time. The time-longitude diagrams show the averaged outgoing long-wave radiation (OLR) near the Equator from observations and a series of forecasts for the period December 29, 1992 (top of panels) to February 15, 1993 (bottom of panels) with a lead time of 15-days from various versions of the ECWMF operational forecast model. The earlier version of the model performs poorly (labeled “10/04”), while the later version replicates (labeled “09/09”) more features of the observations. Cycle identification and the date it became operational are given at the top and bottom of each panel, respectively. Red shading represents positive OLR anomalies and blue shading represents negative OLR anomalies. SOURCE: ECMWF; Bechtold et al. (2008); updates courtesy of Frederic Vitart (ECMWF).

models include weak and/or incoherent eastward propagation, particularly across the Maritime Continent, and difficulties representing the vertical structure associated with water vapor, clouds, and convective processes.

To facilitate more substantial gains in the operational context, the MJOWG developed an MJO forecast metric (adapted from Wheeler and Hendon, 2004), and along with the Working Group on Numerical Experimentation (WGNE) facilitated its adoption by a number of operational forecast centers. In addition, MJOWG has worked with CPC/NOAA to aggregate, display, and disseminate its real-time forecasts in a uniform manner (Gottschalck et al., 2010). The motivation for having such a metric is that it allows for a more quantitative assessment of forecast quality, the ability to track model forecast quality over time, the capability to measure potential model improvements relative to the MJO, and the means to develop a multi-model ensemble forecast of the MJO. The motivation for improving MJO forecasts and the means to evaluate them involve not only the MJO itself but extend to other weather patterns it influences and interacts with such as extratropical weather patterns (e.g., Ferranti et al. 1990; Bond and Vecchi 2003; Jones et al. 2004a, b; Vecchi and Bond, 2004) and tropical cyclones (e.g., Maloney and Hartmann, 2000; Vitart et al. 2010). Given the importance of the MJO in terms of its contribution to ISI predictability, the above programmatic research work started by the US CLIVAR Working Group has been extended through the recent formation of a WCRP-WWPR/THORPEX MJO Task Force (see *http://www.ucar.edu/yotc/mjo.html*).

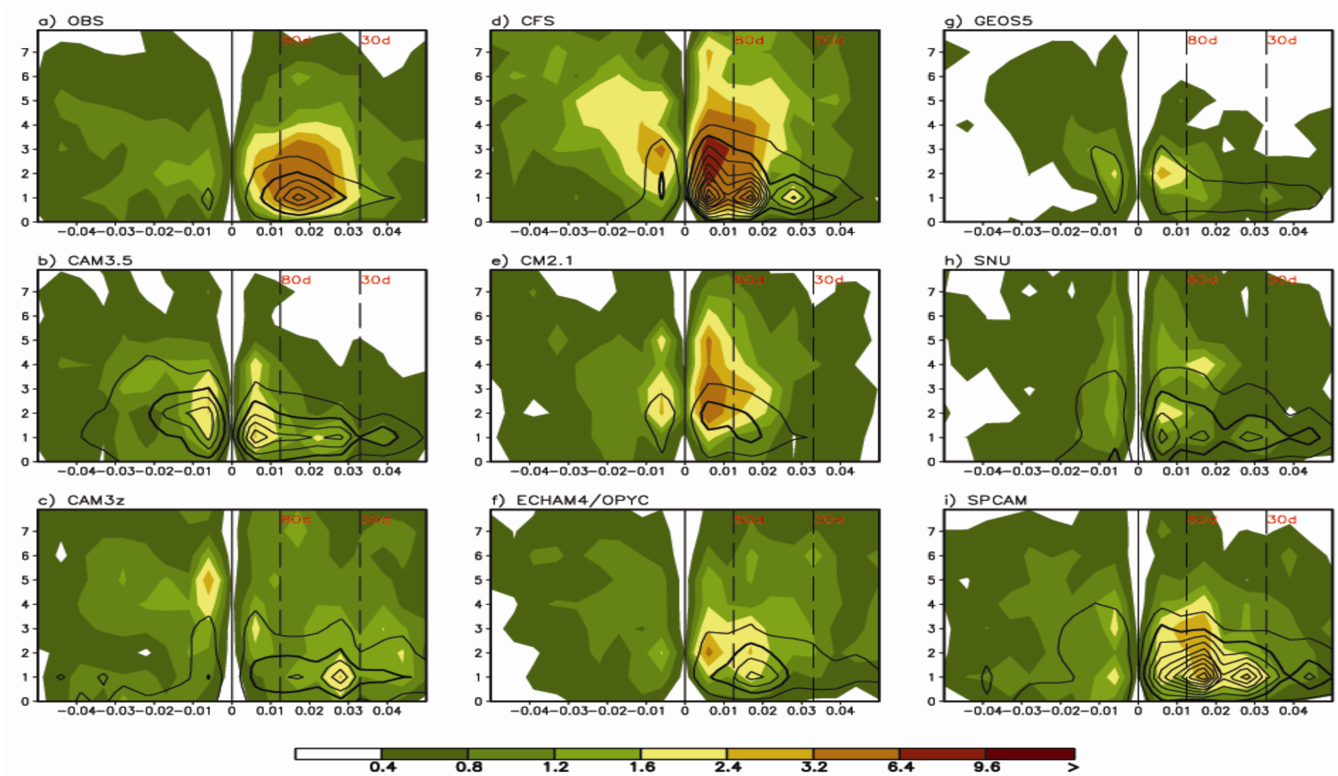

FIGURE 4.8 Precipitation and wind variability generated by dynamical models serve as a poor match to observations. The first panel (a) shows the observations of November–April wavenumber-frequency spectra (wavenumbers on the y-axis; temporal frequency on the x-axis) of averaged precipitation (shaded) and 850-hPa zonal wind (contoured), near the Equator. Very few of the models (b–i) exhibit rainfall or wind variability similar to that of the observations. The vertical lines indicate the spectral region where variability "should" be occurring. Units for the precipitation (zonal wind) spectrum are mm^2 day^{-2} (m^2 s^{-2}) per frequency interval per wavenumber interval. SOURCE: Kim et al. (2009).

While the above efforts will facilitate continued development and improvements in operational MJO predictions, it is certain that capturing the forecast quality associated with the MJO is most strongly hampered by the systematic biases in the models, such as those illustrated in Figure 4.8. Rectifying these will likely require more examination and understanding of the vertical structure of clouds and associated diabatic heating (Lin et al., 2004; Jiang et al., 2009; Fu and Wang, 2009), cloud-radiative interactions (Lee et al., 2001; Lin and Mapes, 2004), microphysical processes (Tompkins et al., 2007; Waliser et al., 2009), interactions with the surface (Maloney and Sobel, 2004; Sobel et al., 2008) and the fine-scale structure embedded in the MJO (Nakazawa, 1988; Lau et al., 1991; Chen et al., 1996; Moncrieff, 2004; Moncrieff and Liu, 2006; Majda and Stechmann, 2009).

In addition, there are a number of operational and implementation issues that need to be explored. For example, given that most of the predictability of the MJO is thought to reside in its large-scale atmospheric structure suggests that its state can be well captured by the conventional synoptic and satellite observations available today. However there have been few studies that identify which aspects of the initial conditions of the atmospheric state are important to MJO forecasts, as mentioned above (see Vintzileos and Pan, 2007). Apart from the prominence of the atmospheric state to the MJO, there are a host of theoretical and modeling studies that suggest the SST and mixed-layer heat content are important in the evolution of the MJO and thus are quantities that need to be represented accurately in the initial conditions and subsequent prediction model evolution (Fu et al., 2003; Zheng et al., 2004; Fu et al., 2006; Zhang et al., 2006; Woolnough et al., 2007; Pegion and Kirtman, 2008). Also, the initiation mechanism(s) associated with the MJO in the eastern/central Indian Ocean represents a significant outstanding question that impacts the development of MJO forecasting. A number of these issues will be explored through the analysis of the multi-model hindcast data set that was specifically designed for examination of the predictability and prediction sensitivities of the MJO and the development of multi-model ensemble strategies (see *http://www.ucar.edu/yotc/iso.html*). Moreover, field studies such as the upcoming DYNAMO/CINDY campaign in the Indian Ocean in late 2011 through early 2012 will be helpful to address the questions concerning MJO initiation. Additional discussion on the above issues and recommendations for ways forward can be found in numerous studies (Zhang, 2005; Waliser, 2006; Moncrieff et al., 2007; Sperber and Waliser, 2008; Waliser and Moncrieff, 2008; Gottschalck et al., 2010).

SOIL MOISTURE

The use of realistic soil moisture initialization can potentially improve precipitation and air temperature forecasts on ISI time scales. The mechanistic pathway can be described through a hypothetical situation. Consider a soil that is anomalously wet at the beginning of a forecast period. Due to the inherent memory associated with soil moisture, the anomaly would likely persist for several weeks, and during this time period, the evaporation rate from the land surface would likely be anomalously high. High evaporation, in turn, could lead to anomalously cool surface temperatures as a consequence of enhanced evaporative cooling. In addition, the anomalously high rate of evaporation could affect the atmosphere (e.g., changes to the boundary layer structure, humidity), which could increase the chance or amount of precipitation.

While easy to describe, many of the mechanisms involved are profoundly complex and are not yet fully understood, particularly those associated with the impacts of evaporation on rain generation. Understanding and quantifying the contribution of soil moisture initialization to climate forecast quality is the focus of many recent studies.

Scientific Basis for Prediction

Energy and water are naturally conserved at the land surface, and the energy balance and water balance equations share a common term: evaporation. Through the energy balance, higher evaporation means a greater evaporative cooling of the land surface, which can translate to cooler air temperatures, as noted above. It can also, through both balances, affect precipitation—the lower sensible heat flux associated with higher evaporation can lead to shallower boundary layers and thus an easier build-up of the conditions that trigger convective rainfall (Betts et al. 1994), and the evaporation itself can serve as a moisture source. However, under certain (and probably rarer) conditions, higher evaporation rates may have the opposite effect—they may act to inhibit precipitation (Findell and Eltahir, 2003; Cook et al., 2006).

One of the chief controls over evaporation in many parts of the world is soil moisture. If the soil is not too wet to begin with, higher soil moisture tends to lead to higher evaporation; evaporation in many parts of the world is water-limited. (Soil moisture has an additional, secondary effect on the surface energy balance, and thus evaporation, through its impact on the surface albedo.) Thus, in a forecast system, if soil moisture could be accurately predicted, then evaporation could be better predicted, with potentially positive impacts on the prediction of precipitation and air temperature (see Figure 4.9) Soil moisture can in fact be predicted reasonably well at subseasonal timescales due to its inherent memory; observational studies (Vinnikov and Yeserkepova, 1991; Vinnikov et al., 1996; Entin et al., 2000) show that this memory is of the order of weeks to months, much longer than that of tropospheric variables. Note that the role of soil moisture in prediction is in some ways complementary to that of SSTs. While the timescale of soil moisture memory pales relative to that of ocean processes, the effects of soil moisture tend to be local and can help make up for a lack of teleconnection between forecasted SSTs and midlatitude continental weather during summer (Koster et al., 2000).

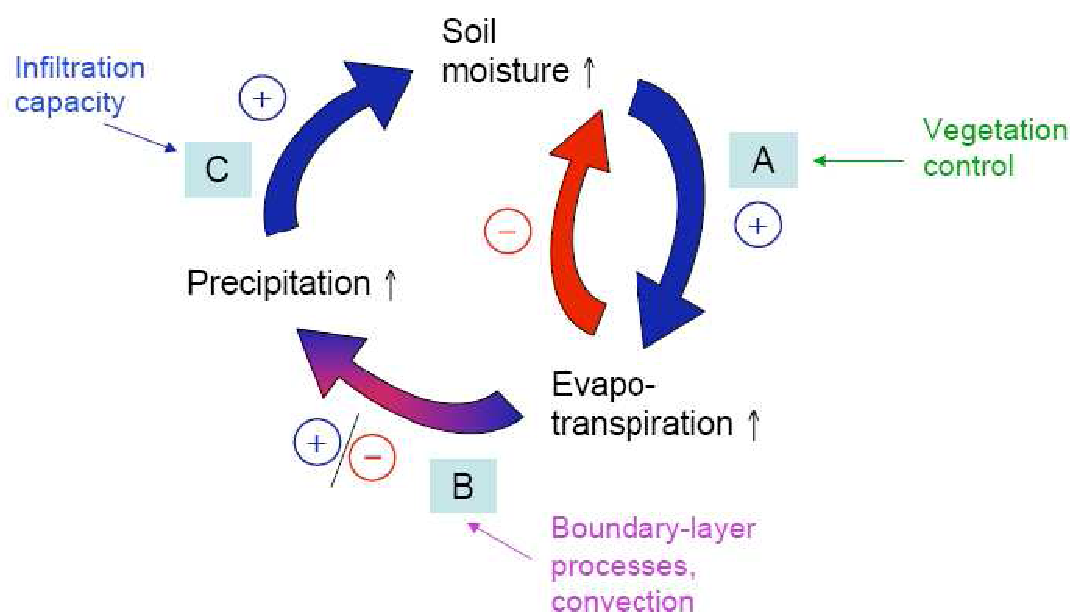

FIGURE 4.9 Simplified schematic showing how soil moisture anomalies can feed back on precipitation. In segment A of the cycle, a positive soil moisture anomaly leads to a positive evaporation anomaly (assuming evaporation is in a soil moisture-limited regime), and this in turn reduces the positive soil moisture anomaly. In segment B, a positive evaporation anomaly can lead to a positive precipitation anomaly, though in certain situations, it can instead lead to a negative precipitation anomaly. In segment C, a positive precipitation anomaly leads to a positive soil moisture anomaly. SOURCE: Seneviratne et al. (2010).

Naturally, given its potential, much research has focused on assessing the usefulness of soil moisture initialization in a forecast system. The first step is to examine how soil moisture variations help guide the evolution of atmospheric variables. The atmospheric general circulation model (AGCM) environment is a natural place to examine this, given the limitations of purely observation-based studies. Using observations, for example, to garner evidence of soil moisture impact on precipitation is difficult because the reverse direction of causality is overwhelmingly dominant—when rainfall is high, the soil gets wet. Some research has looked at lagged correlations between observed soil moisture and precipitation (does a wet soil tend to precede an anomalously high rainfall?), but even such lagged statistics cannot prove causality, given that the rainfall itself may have some long memory associated with it, as induced by remote SSTs, for example. With an AGCM, dominant directions of causality can be artificially disabled, allowing the lesser and potentially interesting ones to be isolated and quantified, at least for the biased AGCM climate.

Such AGCM studies can be segregated into two categories. The first involves assessing the response of the atmospheric variables to the prescription of soil moisture throughout the simulation period. In these studies, no thought is given to soil moisture prediction itself; soil moisture is in essence assumed to be "perfectly forecasted." Shukla and Mintz (1982) demonstrated with an AGCM that a very dry land surface produces a substantially different climate response than a very wet surface. Numerous further studies have since shown that: (1) a suitably wet or dry soil moisture boundary condition can generate precipitation and runoff extremes (droughts and floods; e.g., Atlas et al., 1993; Hong and Kalnay, 2000), and (2) prescribed, more subtle variations in soil moisture lead to correspondingly subtle yet still measurable variations in precipitation and air temperature (e.g., Delworth and Manabe, 1989; Koster et al., 2000; Dirmeyer, 2000; Douville et al., 2001), if only in certain regions.

Such studies were formalized recently in GLACE (the Global Land-Atmosphere Coupling Experiment), a Global Energy and Water Cycle Experiment (GEWEX) and CLIVAR-sponsored international project (Koster et al., 2006; Guo et al., 2006). The GLACE project involved a dozen state-of-the-art AGCMs, each performing precisely the same numerical experiment, one designed specifically to quantify the degree to which the time evolution of simulated precipitation and air temperature can be guided through the specification of the (time and space-varying) soil moisture state. GLACE produced two main results: (1) AGCMs differ significantly in their estimates of how soil moisture variations affect precipitation and air temperature, and (2) they do nevertheless tend to agree that certain regions during certain times of year are more prone to soil moisture impacts on the atmosphere than others.

The regions for which the GLACE models show some consensus regarding impact are indicated in Figure 4.10. These areas are generally transition zones between wet and dry areas—regions like the U.S. Great Plains, which lies between the arid west and the humid east. Mechanistically, the relative importance of these transition zones for land-atmosphere coupling makes perfect sense; in regions that are too arid, evaporation variations are too small to affect the atmosphere, and in regions that are too wet, soil moisture availability does not limit evaporation (energy availability, through radiation, does instead), so that soil moisture variations are not translated into evaporation variations. These mechanistic controls have indeed been demonstrated outside of GLACE through various specially designed AGCM experiments. Fortuitously, satellite retrievals of soil moisture are generally reliable in these transition zones, where the vegetation is not too dense.

The second category of purely model-based studies focuses on the degree to which soil moisture initialization can affect the subsequent simulation of atmospheric variables. These studies do address the prediction of soil moisture since it is not prescribed throughout the forecast period—soil moisture and atmospheric variables are free to evolve together. Large initial soil moisture anomalies have been found to have a significant impact on subsequent precipitation (e.g., Rind, 1982; Oglesby and Erickson, 1999). Less extreme anomalies produce correspondingly more subtle impacts on the atmosphere, with some studies showing an almost

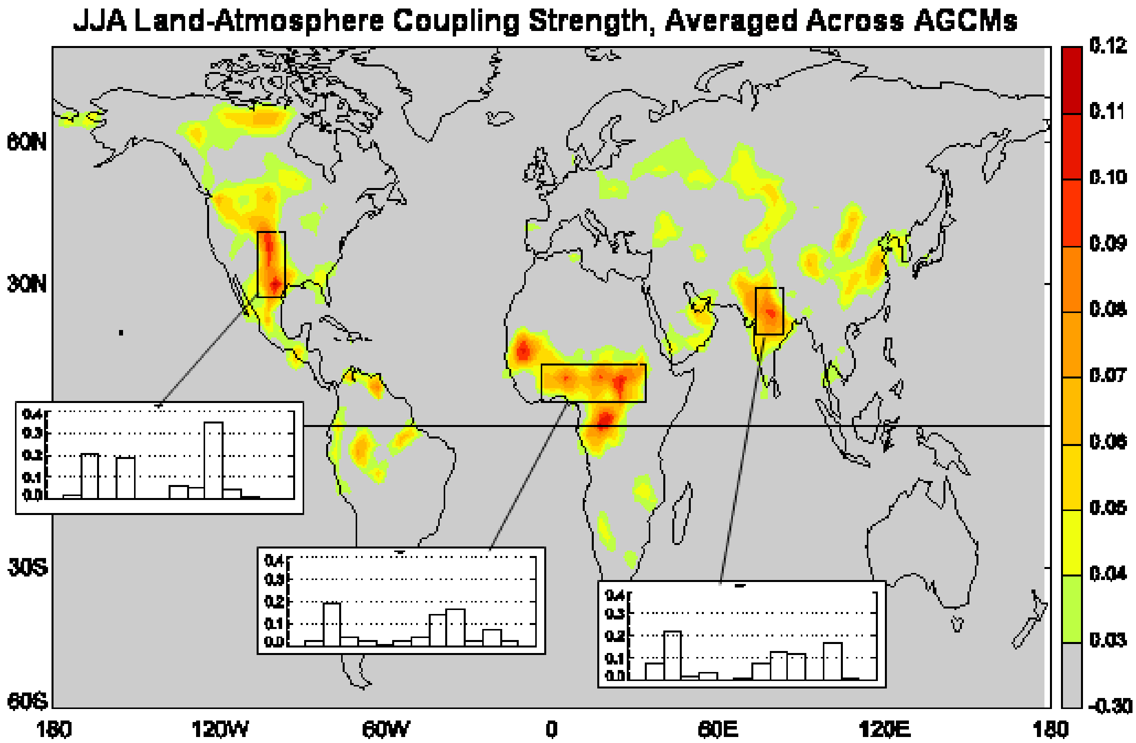

FIGURE 4.10. Areas for which the numerical models participating in the GLACE study tend to agree that variations in soil moisture exert some control on variations in precipitation. The variable plotted is the average across models of a land-atmosphere coupling strength diagnostic; the insets show how the magnitude of this diagnostic differs amongst the participating models. SOURCE: Koster et al. (2004).

negligible impact on (model-specific) precipitation predictions (though significant impact on air temperature predictions) associated with soil moisture initialization (e.g., Schlosser and Milly, 2002).

Forecast System Methodologies

Dynamical models used in forecasting rely on the initial state of soil moisture, which is provided by observations. As explained in Chapter 3 ("Land Observations" and "Data Assimilation" sections), analysis products provide the most useable data for forecasts, as they can provide information about soil moisture at broad spatial scales. In particular, the application of Land Data Assimilation Systems (LDAS) to raw meteorological observations (e.g., temperature, humidity, precipitation, winds) can greatly improve the initialization of soil moisture in dynamical models. The expected state-of-the-art for soil moisture initialization in coming years involves extending LDAS to include assimilated satellite information (soil moisture retrievals or the raw radiance measurements).

Statistical forecast models can also make use of such soil moisture information. For example, the Constructed Analogue technique (see "Constructed Analogs" section in Chapter 3) of van den Dool et al. (2003) makes use of historical soil moisture patterns to infer the evolution of soil moisture, temperature, and precipitation from present soil moisture patterns.

Forecast Performance

The obvious extensions to the idealized, model-based studies discussed in "Scientific Basis for Prediction" section are studies that use true forecast systems and observations of forecasted variables to examine the degree to which realistic soil moisture initialization improves forecast quality. Fennessy and Shukla (1999), for example, examined increases in accuracy derived from a proxy analysis-based soil moisture dataset. Douville and Chauvin (2000) initialized their model with soil moisture estimates derived from the Global Soil Wetness Project, and Viterbo and Betts (1999) examined the impact of soil moistures derived from the ERA-15 reanalysis on the simulation of the 1993 Mississippi flood. Koster et al. (2004) examined 15 years worth of forecasts (5 independent forecasts per year) to quantify forecast accuracy from a relatively sizeable set of forecasts. These studies, and others like them, strongly suggest that realistic soil moisture initialization can provide some increase in the quality of precipitation and air temperature prediction out to a month or more.

In an attempt to generate a multi-model "consensus" view of how realistic land initialization affects forecast quality, several modeling groups have recently embarked on the GLACE-2 project. In this project, the participants perform two parallel sets of forecasts: one in which land surface states, particularly soil moisture, are initialized realistically and one in which they are not. A comparison of the skill (square of correlation relative to observations) derived from these two sets allows a direct quantification of the impact of land initialization on forecast quality. The experiment is ongoing; first results show that across the models, land initialization does improve the correlation skill of temperature forecasts out to 60 days. The correlation skill increases substantially when conditioned on the size of the initial (local) soil moisture anomaly (see Figure 4.11).

The patterns in Figure 4.11 differ somewhat from those in Figure 4.10. These apparent discrepancies might be explained in several ways. First, the original GLACE examined the ability of imposed soil moisture variations to influence the atmosphere, whereas GLACE-2 examined the full prediction question, which also involves the ability of a model to retain an initial soil moisture anomaly through a forecast period. Analyses of model-generated soil moisture memory (e.g., Seneviratne et al. 2006a) suggest that the south central United States has a reduced soil moisture memory relative to the north central region, perhaps hindering the generation of skill there. Sampling error may also be a factor. In addition, results from the two GLACE experiments almost certainly differ because the first was purely synthetic—the patterns produced in the original GLACE experiment necessarily reflect the biased climatologies of the models, possibly shifting, for example, the area of high land-atmosphere coupling in the north central U.S. toward the west relative to that implied by the GLACE-2 results. The GLACE-2 experiment was not synthetic, and thus the patterns produced are controlled by both the climatology inherent in nature (through the realistic initialization) and by the ability of the models to perform realistically during the forecast.

Regarding precipitation, the consensus from GLACE-2 is less robust. The results do show, however, some small land-related increases in accuracy for precipitation out to at least 45 days, again especially when conditioned on the initial anomaly (Koster et al., 2010), and again especially in the north central United States.

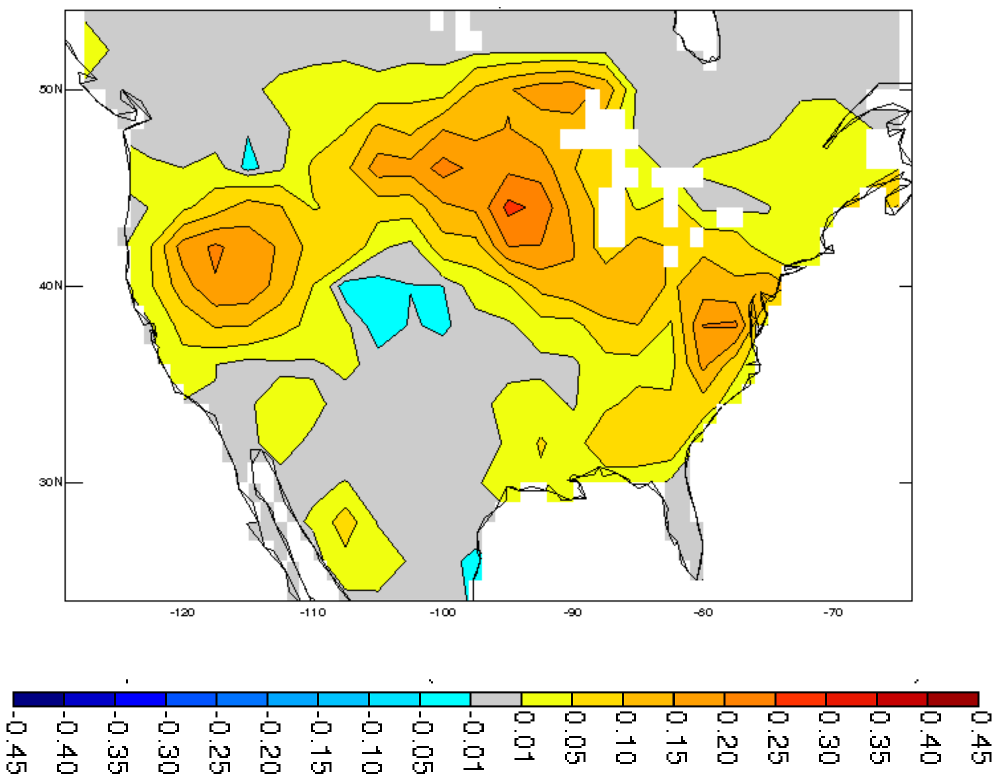

FIGURE 4.11. Conditional improvement in forecast accuracy for air temperature related to improved land initialization. "Warm" colors indicate areas where accuracy is gained when realistic land initialization is used. The metric plotted represents the square of the correlation coefficient (R^2) obtained with realistic land initialization minus the R^2 value obtained without it. The quantity predicted and compared to observations is the average air temperature over days 31–45 of the forecast. Forecasts at a given location are conditioned on the initial soil moisture at that location—only those start dates for which the initial soil moisture lies in the lowest or highest quintile of all realized values are used in the calculation. Adapted from Koster et al. (2010).

Some statistical analyses have addressed the potential impacts of soil moisture on skill (e.g., Karl, 1986; Huang et al., 1996b). These also point to the interior of North America as the place to find positive correlations between soil moisture and future temperature, though not always in the same locations.

Gaps in Understanding

While improvements to satellite retrievals and LDAS can improve the initialization of soil moisture, improving our understanding of the physical processes linking soil moisture to precipitation remains critical to exploiting the benefit of soil moisture to forecast quality, especially with respect to its impact on precipitation. The insets in Figure 4.10 above highlight the fact that models differ widely in their assessment of soil moisture's ability to affect precipitation. This ability, the "land-atmosphere coupling strength", underlies any contribution made by soil moisture initialization to precipitation prediction. (Similar inter-model variability is seen for the soil moisture—temperature connection.) As suggested above, we cannot even say which model performs best because the effective coupling strength operating in nature cannot be directly measured. Only indirect estimates are available, so only these can be used for model evaluation. The best way to do this is still unclear.

Presumably, an accurate simulation of land-atmosphere coupling relies on an accurate simulation of the complex and interacting model formulations that contribute to it. Work is needed to evaluate and improve model formulations of soil-moisture limited evaporation, turbulent transport from the land surface, atmospheric boundary layer generation, and moist convection, among other processes. The global distributions of the parameters (e.g., effective soil texture at large scale) that control some of these processes are often crudely estimated; work (possibly involving calibration) is needed to improve these fields. Additional work is needed to ensure that the simulated processes interact with each other in realistic ways.

Of course, even if all model formulations, and thus the simulated coupling strength, were perfect, prediction quality would still be limited by deficiencies in our ability to assign realistic global distributions of soil moisture initial conditions. As noted above, the expected state-of-the-art in the coming years is the production of soil moisture fields through data assimilation systems utilizing satellite data; even so, these estimates will be limited by errors in the meteorological forcing fields and retrieval estimates. The incremental improvement of satellite soil moisture retrievals (or the corresponding raw radiances) on the effectiveness of the soil moistures used in a forecast system is yet to be established.

Several facets of soil moisture's impact on precipitation and temperature forecast quality are still largely unexplored. For example, soil moisture's contribution to the accuracy of climate predictions may be larger for certain background climatic conditions. For a given region, can we expect the impact of soil moisture to be greater under certain SST and atmospheric circulation regimes, or for modified conditions associated with global climate change (Seneviratne et al., 2006b)? The time is ripe, and modeling systems are adequately advanced to address such questions.

Finally, in regard to ISI forecasting, we note that the value of soil moisture initialization is not necessarily limited to the prediction of meteorological variables. For example, water resource estimates in regions with significant snowpack are often made months in advance by estimating snow amounts—a greater amount of snow implies a greater amount of meltwater during the spring season and thus greater streamflow. However, the future streamflow can also depend on soil moisture. If the soil is dry below a melting snowpack, much of the meltwater may infiltrate the soil and then evaporate, and would not then be available to contribute to streamflow and thereby reservoir storage. On the other hand, if the soil is wet, a greater amount of the meltwater may flow into streams. This potential use of wintertime soil moisture to add skill to springtime water resources prediction is not yet fully tapped.

5

Best Practices

The scientific and programmatic challenges identified in our discussion of the basic building blocks of ISI prediction systems (i.e., Chapter 3) and the case studies of Chapter 4 have common themes that naturally lead to a discussion of "Best Practices." Essentially, Best Practices aim to answer the following:

- How can we improve prediction systems and the provision of forecasts?
- Given that both qualitative and quantitative improvements in seasonal forecasts are possible, how do we begin to map a path forward?

"Best Practices" is an optimum process, assessing forecast quality and enabling productive interactions among the various ISI forecasting communities (e.g., users, developers, providers, researchers). The discussion of Best Practices necessarily cuts across the details of how forecasts are evaluated and shared to more programmatic issues of how operational centers collaborate with the outside community. Specifically, four important aspects of Best Practices for the production, reproduction, evaluation, and dissemination of ISI forecasts are presented: public archives of forecast information, forecast metrics, more useful forecast products, and an improved synergy between operational and academic communities.

PUBLIC ARCHIVES

Transparency and reproducibility are essential for assessing and improving ISI forecast quality and enhancing communication among operational centers, researchers, and users. Currently, it can be difficult to determine or access the inputs (e.g., observations) and methods (e.g., models, data assimilation schemes, subjective input) that underlie a particular forecast. Likewise, many forecast products (e.g., hindcasts, analyses, forecasts, re-analyses, re-forecasts, verifications, outlooks) may not be archived or the existing archives may not be publically accessible. As noted in several instances by this report, valuable research has been preformed when such data sets are available (e.g. CMIP, ENSEMBLES, DEMETER). For many of these projects, international collaborations were enabled by the availability of forecast data. Ensuring that forecast centers establish and maintain archives is important to bolster these collaborations among forecast centers and to continue the large-scale assessment and comparison of prediction systems.

Assessing current ISI forecast performance, making comparisons among forecast systems, devising strategies for improving forecasts, and understanding the impact of a change to a forecast system (e.g., incorporation of new observations, updates to model parameters) all require the existence of publically accessible and comprehensive archives. Documentation and archiving of data, models, methods, and products by operational centers would serve as an integral step to assessing and ultimately improving ISI forecast quality. This is especially the case if attribution of forecast improvements to specific proposed/implemented changes to the system is desired. The observing systems will evolve and studies are needed to assess and guide that evolution from the perspective of the role of observations in ISI forecasting.

Given that subjective intervention is a component of many forecast systems, it is important that the objective inputs can be easily separated from the subjective final product for independent analysis and appraisal. This separation is necessary for assessing whether the objective elements are improving and whether improvements in observations, understanding, or models, or some combination, are having a positive impact. Similarly, for forecast systems that combine statistical and dynamical prediction techniques, it is important to be able to separate the contributions from each component.

METRICS

Evaluating ISI forecast quality requires a set of well-defined model performance and forecast metrics that can be applied to current and future prediction systems. Forecast metrics need to include both deterministic and probabilistic measures. Model performance metrics, which in this case are generally associated with dynamical models, need to include measures of model success in representing the mean climate, forced variability (e.g., diurnal and annual cycles), unforced variability (e.g., ENSO, MJO, PNA) and key physical processes (e.g., convection, fluxes, tropical waves). Multiple metrics are recommended since no single variable or metric is sufficient to fully characterize model and forecast quality for multiple user communities. These aspects include, but are not limited to, measures of bias (correspondence between the mean forecast and the mean observation), accuracy (the level of agreement between the forecast and the observation), reliability (the average agreement between the forecast values and the observed values when the forecasts are stratified into different categories, e.g., conditional bias), resolution (the ability of the forecast to sort or resolve the set of events into subsets with different frequency distributions), sharpness (the tendency of a forecast to predict extreme values), and discrimination (the ability of a forecast to discriminate between observations to have a higher prediction frequency for an outcome when an outcome occurs).

Regardless of which metrics are used, the following properties are necessary for a set of metrics:

- Provide the ability to track forecast quality to determine if models are improving. This implies that the uncertainty in the skill statistics needs to be quantified.
- Provide some feedback on model strengths and weaknesses in providing an accurate forecast.
- Allow forecasts from different systems to be compared to identify which system is superior.
- Provide information on metric uncertainty. This allows for forecast consistency to be evaluated.
- Include a justifiable baseline of forecast quality for comparison.

The WMO Standard Verification System (SVS) for Long Range Forecasts (LRF) and the text by Jolliffe and Stephenson (2003) are excellent starting points for developing these metrics, but additional metrics will need to be developed as forecasts and their use evolve. For example, until recently there was no well-accepted and documented forecast metric for the MJO (Gottschalck et al. 2010). Examples of model performance metrics for the MJO include those developed by U.S. CLIVAR (see section MJO case study in Chapter 4; CLIVAR MJO Working Group, 2009). A number of programmatic activities have interest in developing model performance metrics that would be applicable to models used in ISI forecasting (e.g. GEWEX Cloud System Study (GCSS), Climate Historical Forecast Project (CHFP), and Climate Process Teams (CPT)). Such consideration of metrics and their development only reinforces the need for open and easy access to forecast information as discussed above.

MORE USEFUL FORECAST PRODUCTS

The promise that climate forecast information may benefit society through improved decisions and climate risk management motivates much of the human and fiscal investments in the research and production of these forecasts. Although it is often assumed that climate forecasts would be used more if they were of better quality, other factors are often cited as equally important, including the retrospective forecast performance, the societal and scientific relevance of the forecast variables and their specificity, and the manner in which the forecast is communicated.

The forecasts have to be probabilistic, as estimates of the state of the climate system are inherently probabilistic. Decision makers are accustomed to using uncertain information; risk is by definition probabilistic. But, use of probabilistic climate forecasts requires information regarding the reliability of the probabilities. Thus, additional information is needed on past performance of the prediction inputs and the overall forecast system, or the data have to be available for users to assess the forecast system based on their own requirements. If the reliability of the forecast probabilities is unknown, users often subjectively fold in additional uncertainty to the probabilities, which reduces the usefulness of the forecast. Even if there is a no-skill forecast, users need to be made aware of it either in a raw data or graphical format. It is important to document no-skill forecasts, recognizing that conditional skill may exist for a particular variable and region (i.e., areas with no-skill forecasts during one season may have useful forecasts during a different season). In addition, for the purposes of tracking forecasts over time, users may potentially find such information helpful. However, even if a forecast is labeled as “no-skill,” information on the historical climatology can still be provided to indicate the range of possibilities seen in past years.

Beyond providing information on the forecast quality, the forecast variables have to be relevant for the decision maker. Much of the ISI forecast verification to date has involved variables such as the Nino3.4 index, and even more recently the phase of the MJO. These variables are arguably distant from the needs of end users, who might instead require information on precipitation or air temperature over populated regions. Information on seasonal rainfall totals or average seasonal temperature may in turn be less important than the frequency and duration of dry spells or heat waves, or favorable conditions for tropical cyclone formation. Moreover, given that many decisions are triggered by risk of threshold crossings (e.g., not enough rain, overly high temperatures), the events or categories for which probabilities are provided need to be determined by these conditions. Other decisions might be strongly tied to spatial considerations; for example, in some cases a large basin-scale pattern might be relevant and in others the evolution of a given variable at a specific model grid-box location might be needed. Some studies have demonstrated significant associations between ENSO events and societally relevant weather variability, such as mid-Atlantic winter storms (Hirsch et al., 2001) and variability in peak wind gusts (Enloe et al., 2004), but these have not been translated into operational products. Operational forecast centers may address some of these needs in collaboration with certain decision or policy makers, but because forecast needs are typically sector-specific and even region-specific, they cannot anticipate every decision setting. If the forecast data and tools are made available as discussed above it will be possible for users to tailor their own forecasts.

While not all end-users of the forecasts will be willing or able to tailor forecast information themselves, it is necessary to remember that users of climate prediction information encompass more than end-users or decision makers. Sectoral scientists that develop system analyses and decision models use climate prediction information as input to their models. Also, climate scientists who conduct research in areas including process studies, multi-model ensembles, and downscaling use climate prediction data. These scientists add considerable value to the development and improvement of the ISI prediction process. By working together with the end-users or decision makers, such intermediaries can help society realize the value of ISI forecasts.

ACCELERATED SYNERGY WITH THE RESEARCH COMMUNITY

There are too many major science directions for possible improvements of operational systems to be examined and implemented by either the operational centers or the research community alone. Each community has its own strengths and purpose: operational centers excel in creating robust and reliable forecasting systems using state-of-the-art models, observations and data assimilation systems; academic researchers excel in developing new ideas and approaches. Large-scale scientific challenges can often exceed the capacity of each group acting alone with respect to extreme infrastructure demands (e.g., computational resources), ancillary but important expertise (e.g, satellite or *in situ* measurement development) or the need for interdisciplinary approaches and expertise. This necessitates collaboration. Efforts to improve ISI forecasting should enhance communication and interaction between these communities while drawing on their complementary strengths and differentiated roles.

In terms of accelerating this synergy, the committee noted two positive examples. First, ECMWF has an annual targeted workshop with a specific focus of improving some particular element of the operational prediction system (e.g., data assimilation system, estimates of forecast uncertainty). External visitors are invited to not only make presentations, but more importantly to chair breakout groups that provide detailed and specific recommendations to the center. This activity is effective because the outside community welcomes the opportunity to affect and improve operational prediction, and because the center is committed to be responsive to the breakout group recommendation. The second example is the ongoing NOAA/NCEP Climate Test-bed seminar series. This particular seminar series has speakers from both the operational centers and the outside research community. The location of the seminars rotates through COLA (Center for Ocean-Land-Atmosphere Studies, Calverton, MD), ESSIC (Earth System Science

Interdisciplinary Center, University of Maryland) and NCEP, with the intention of having the operational scientists speak at the research centers and the external scientists speak at NCEP, thus fostering cross fertilization between operations and research.

6

Recommendations and Remarks on Implementation

In this report, the assessment of prediction capabilities for intraseasonal to interannual (ISI) timescales has been made by focusing on the variables and processes that act as sources of predictability for the climate system. The assessment also describes the building blocks of ISI forecast systems, how forecasts are verified and disseminated, and the relationships among the building blocks, forecasting procedures, and improvements in ISI forecast quality. The committee focuses on qualitative estimates of potential improvements in ISI prediction systems, since a quantitative upper bound of predictability for the climate system cannot be made at this time (see Chapter 1).

In this chapter, the committee's recommendations are presented. Following the recommendations, the committee presents its thoughts regarding the expanded use of observations, the prospects for seamless forecasting, and the magnitude and rate of expected forecast improvements.

RECOMMENDATIONS

The committee identified three general categories of actions to advance ISI predictions: Best Practices, Improvements to the Building Blocks of ISI Forecast Systems, and Research for Sources of Predictability. The Best Practices are largely focused on the activities of the operational forecast centers and aim to improve the delivery and dissemination of forecast information for both decision-makers and researchers. The Improvements to the Building Blocks of ISI Forecast Systems pertain to both the operational and research communities and focus on the continued development of observations, statistical and dynamical models, and data assimilations systems. Research for Sources of Predictability is aimed primarily toward the research community.

These three categories indicate the relative time horizons associated with the recommendations. Many of the Best Practices could be adopted relatively quickly. Although some of the suggestions for archiving forecasts and convening meetings may require increased resources and planning, most of the recommendations involve modifications to the routine activities of operational centers rather than new initiatives or programs. Improvements to the Building Blocks of ISI Forecast Systems will likely require more time and effort to pursue and will necessitate significant collaboration between operational forecasters and research scientists. Incorporating and validating some of these changes would likely occur over several years. The

Research for Sources of Predictability provides a set of longer-term research goals. Although many experiments can be designed and run now, results may take several years to emerge, given the pace of scientific publication and discourse. Once some of this research has been completed, translation of the results into an operational setting will require subsequent efforts by both operational centers and research scientists. Thus, the research goals constitute a longer-term vision.

Best Practices

Lack of access to forecast and verification information and issues with communication between research and operational forecasting communities are major barriers to the improvement of ISI forecast systems. The committee recommends several steps to foster expanded collaboration, create an archive of key ISI forecast data, set standards for verification techniques, and minimize subjective components of ISI forecast.

(1) The synergy between operational ISI forecasting centers and the research community should be enhanced.

A number of important activities would contribute greatly to the goal of accelerated synergy and progress. The committee recommends the following.

- Targeted workshops focused on specific areas relevant to model and forecast improvement should be held at least annually at the operational centers. The workshops should produce actionable recommendations that result in specific plans for developing and testing new ideas for operational forecasting.
- Scientists in the operational centers should participate actively in scientific meetings, especially in the areas of modeling and use of observations.
- Short term positions in operational centers should be granted to academic researchers. These positions could focus on a particular scientific issue that is both in the researcher's field of expertise and offers opportunities for improving operational forecast quality.
- New data sets from both the academic researchers and scientists in operational centers should be made available to the scrutiny of the broader academic community at an early stage to help identify early strengths and weaknesses.
- The development of new observations to support ISI forecasting should be carried out with the engagement of the operational centers through an ongoing dialog about the efficacy of the observing system and the need for further observational campaigns by the research community.

(2) Operational ISI forecasting centers should establish public archives of all data used in forecasts, including observations, model code, hindcasts, analyses, forecasts, re-analyses, re-forecasts, verifications, and official forecast outlooks.

Archives of forecast information are needed by national and international operational centers, researchers, and the private sector in their efforts to quantify and identify sources of forecast error, provide the baseline for forecast assessment and model fidelity, develop metrics and diagnostics for model assessment, calibrate model forecasts, quantify and document model and forecast improvement, such as those that results from changing resolution or parameterizations, and develop tailored forecast products for decision systems and climate risk management.

Archives serve as an important mechanism for bridging the gap between operational centers and forecast users, whether they are involved in making climate-related management decisions or conducting societally relevant research. Since it is not possible for operational centers to foresee or address all possible needs of these users, archives of forecast information will permit users to access the information that is most important to them and, in some cases, develop their own derivative products. Feedback from forecast users can also offer pathways to improving ISI forecast quality.

(3) Operational ISI forecasting centers should broaden and make available the collection of metrics used to assess forecast quality.

No perfect metric exists that conveys all the information about a forecast. Quantitative skill assessment of forecast quality should be determined and made available through multiple metrics and graphical techniques, including ones that assess the quality of the probabilistic information and multi-model ensembles. Some of these metrics should include information on the distribution of skill in space and time.

(4) The subjective components of operational ISI forecasts should be minimized.

Recent research suggests that the subjective component of many present-day forecasts can reduce forecast quality (e.g., O'Lenic et al. 2008). The subjective component generally comes from qualitative discussion and interpretation by forecasters regarding the state of the climate system and forecasting tools. The subjective component also limits reproducibility, restricting retrospective comparison of forecast systems.

Model "modularization" was an aspect of Best Practices that was considered by the committee but not recommended. Frameworks for modularizing climate modeling codes have been suggested as a way to provide the software infrastructure to enhance collaboration, research, and ultimately the transition of research into operational activities. One such example is the Earth System Modeling Framework[18] (ESMF). In principle, activities such as ESMF have universal appeal; however, in practice they are more complicated than anticipated and their adaptation is often uneven. These efforts should be encouraged in the long view, but their utility in facilitating the transition of research to operations, making operational models more accessible to researchers or enhancing seasonal forecast quality has yet to be demonstrated.

Improvements to the Building Blocks of Forecast Systems

ISI forecast quality will improve with better building blocks. New observations and new uses of existing observations can provide better initializations of various components of the climate system (atmosphere, ocean, land, and ice) and information regarding poorly understood processes that operate within and among components of the climate system. New uses of

[18] The basic idea behind ESMF is that complicated applications should be broken up into smaller pieces, or components. A component may be a physical domain, or a function such as a coupler or I/O system. ESMF also includes toolkits for building components and applications, such as regridding software, calendar management, logging and error handling, and parallel communications

statistical methods, especially nonlinear techniques, can serve as useful forecast models. They can also be used to diagnose problems in dynamical models or to identify and characterize novel sources of variability. Dynamical models can be improved since they exhibit several well-known biases. More advanced data assimilation methods can be used in operational settings, and more observations can be assimilated into forecasts.

(5) Statistical techniques, especially nonlinear methods, should be pursued in order to better characterize processes that contribute to ISI forecasts.

Comparisons between statistical models and dynamical models provide information on the deficiencies within dynamical models. Historically, linear statistical analyses of observational data have provided an awareness of important spatial patterns and teleconnections. Recent research (e.g., Lima et al., 2009) demonstrates that nonlinear methods can yield statistically significant increases in prediction skill on ISI time scales when compared to traditional linear techniques. However, these techniques have not been incorporated operationally. The committee finds that there is a value in expanding such analyses to nonlinear counterparts.

(6) Systematic errors in dynamical models should be identified.

Current state-of-the-art ISI prediction models have relatively large biases that reduce prediction quality. Some classic examples include: (1) the so-called double ITCZ problem, (2) the excessively strong equatorial cold tongue, (3) weak or incoherent intraseasonal variability, (4) failure to represent the multi-scale organization of tropical convection, and (5) poorly represented cloud processes, particularly low level stratus. These errors have both regional and global impacts and could be indicative of errors in the model formulations that are limiting predictability.

Sustained observations are needed to quantify model systematic errors. Examples of sustained observations include those related to describing the properties of or fluxes among the atmosphere, ocean, and land surface (e.g., boundary layer humidity, exchange of heat between the atmosphere and ocean).

(7) To reduce errors produced by dynamical models, the representation of physical processes should be improved.

Systematic errors in dynamical models should be reduced by expanding the understanding of underlying physical processes, with the goal of transferring improvements into operational ISI forecasts. This includes systematic errors in the mean and the variability and their interaction. Process studies that are closely tied to operational ISI model improvement should be carried out on specific components of the climate system (e.g., sea ice, aerosols, snow cover), specific processes and variability (e.g., triggering the onset of an MJO), and the interactions among components of the climate system (e.g., air-land coupling strength, stratosphere-troposphere interactions).

The research community should be engaged in the conduct of process studies that address the physical processes governing ISI variability. Specific physical phenomena are poorly represented in most ISI prediction systems (e.g. the MJO). A strategy to expand knowledge for use in model development is exemplified by the CLIVAR climate process teams (CPTs). The CPTs focus modelers and process scientists on poorly represented or unrepresented physical process in models. Similarly, the WCRP-WWRP/THORPEX's Year of Tropical Convection (YOTC) is another approach that focuses community effort via a virtual Intensive Observation

Period (IOP), combining already existing observational resources with incremental programmatic efforts to target model improvements of a few specific phenomena (e.g., MJO, easterly waves, tropical cyclones).

Process studies and phenomenological focus aside, enough cannot be said for seeking brute force improvements in computing capabilities for better resolving subgrid scale processes and removing as much reliance on parameterization as possible. This could be in the context of limited domain models (e.g., cloud-resolving models) used to develop and evaluate subgrid scale parameterizations as well as cloud or cloud-system resolving/permitting global models. In addition, the impact of increasing the resolution of ISI models should be further investigated[19].

(8) Statistical and dynamical models should continue to be used in a complementary fashion by operational ISI forecasting centers.

Using multiple prediction tools leads to improved and more complete ISI forecasts. Forecasting centers should continue to use statistical and dynamical models in a complementary fashion. Examples of statistical techniques include stochastic physics, interactive ensembles, empirical corrections or empirically-based parameterizations and process models.

The use of statistical and dynamical downscaling methods is another application that should be explored to address the information mismatch between the coarse spatial resolution of operational climate forecasts and the fine resolution needs of some end users.

(9) Multi-model ensemble (MME) forecast strategies should be pursued, but standards and metrics for model selection should be developed.

Multi-model ensembles (MME) have been shown to outperform individual models for forecasting. The committee encourages continued exploration of MME experiments. Understanding why the statistics associated with them consistently outperform the predictions from individual models should be a goal for researchers and operational centers. Current multi-model techniques generally include models based simply on what is available; continued work is necessary to develop techniques of optimally selecting and weighting ensemble members. Experimentation with MME should not compete with model improvement, but rather, should contribute to the process of identifying areas for model improvement.

(10) To enable assimilation of all available observations of the coupled climate system, operational centers should implement state-of-the-art 4-D Var, Ensemble Kalman Filters, or hybrids of these in their data assimilation systems.

Assimilation methods currently being used are often obsolete, and many observations are not being assimilated as part of the forecast cycle. To enable assimilation of all available observations of the coupled climate system, operational centers should implement state-of-the-art 4-D Var, Ensemble Kalman Filters, or hybrids of these in their data assimilation systems. Priority should be given to expanding operational data assimilation to ocean observations, such as sea surface heights.

---

[19] The biggest advantage of the high resolution models lies in better resolving extreme events (tropical cyclones, heavy rainfall, instability waves etc.) and perhaps their feedback to large scale slow dynamics. On the other hand, it has also been shown that the coupled models' biases in representing mean precipitation and SST (e.g., Pacific cold tongue and double ITCZ) remain when the model resolutions increase. There is a need to continue evaluating the positive impacts of the resolution and comparing with other competing strategies in improving model prediction, such as MME.

Research for Sources of Predictability

(11) Many sources of predictability remain to be fully exploited by ISI forecast systems. To better understand key processes that are likely to contribute to improved ISI predictions, the committee recommends that the scientific community pursue the following six areas as research goals.

To identify research priorities, the committee applied four criteria to the list of sources of predictability listed in Chapter 2. Sources that merited further research were selected based on:

1) Physical principles indicate that the source has an impact on ISI variability and predictability.
2) Empirical or modeling evidence exists to support the case made based on physical principles in (1).
3) The committee could identify gaps in knowledge that have prevented these sources from being exploited by ISI forecast systems.
4) There is potential social value for gaining knowledge of this source of variability. For example, the MJO has significant societal impact in its effect on the Indian monsoon, which determines water supply and agricultural productivity for billions of people.

The following six areas met these criteria, but are not presented with any further prioritization.

MJO

A concerted effort on improving the prediction quality associated with the MJO should be undertaken and coordinated with research activities. The path forward on this should include focused process studies, model improvement, and close collaboration between research and operational communities (e.g., Year of Tropical Convection (YOTC) project, the MJO Task Force). It will be necessary to develop and implement standardized diagnostics and metrics to gauge model improvements and track improvements in forecast quality. MJO influences on other important components of the climate system (e.g,. ENSO, monsoon onsets and breaks, tropical cyclone genesis, etc.) should continue to be explored and exploited for additional predictability.

Stratosphere-Troposphere Interactions

Operational ISI prediction models should be improved to represent stratosphere-troposphere interactions. Relatively long-lived (up to two months) atmospheric anomalies can arise from stratospheric disturbances. In sensitive areas such as Europe in winter, experiments suggest that the influence of stratospheric variability on land surface temperatures can exceed the local effect of sea surface temperature. Additionally, while our weather and climate models do not often resolve or represent the stratospheric Quasi-Biennial Oscillation very well, it is one of the more predictable features in the atmosphere, and it has been found to exhibit a signature in ISI surface climate.

Ocean-atmosphere coupling

Due to the very large heat capacity of sea water, anomalous sea surface temperatures and upper ocean heat content can have significant impacts on the atmosphere above. The impacts of the anomalies associated with ENSO are well known. However, further research is needed to examine the role of extratropical atmosphere-ocean coupling, to investigate the need to represent ocean-atmosphere coupling more realistically over a wide range of spatial scales (including down to the scales of the sharp SST gradients associated with fronts), and to better observe and represent air-sea fluxes more realistically in models.

Land-atmosphere feedbacks

Research should be directed at maximizing prediction quality associated with land-atmosphere feedbacks. Recent research shows that the realistic initialization of soil moisture in dynamical models can increase the accuracy of precipitation and (especially) temperature predictions at intraseasonal timescales. The realistic initialization of snow amount may also yield better quality predictions, though this connection is relatively unexplored. To maximize the impact of land feedbacks on prediction quality, the mechanisms underlying the land-atmosphere coupling (e.g., evaporation, boundary layer dynamics, convection) need to be better understood and better represented in forecast systems.

High impact events affecting atmospheric composition

Operational centers should be prepared to make ISI forecasts following unusual but high impact events such as volcanic eruptions, limited nuclear exchange, or space impacts that can cause a sudden, drastic change to the atmospheric burden of aerosols and trace gases. Research efforts should study the consequences of such high impact events on the climate system over ISI timescales and provide guidance for improving forecast systems.

Non-stationarity

Trends can be an important source of predictability that should be exploited since accurate trends in atmospheric compositions (e.g. greenhouse gases, aerosols) and land cover can influence ISI variability and forecasts. Current statistical techniques (such as Optimal Climate Normals) and dynamical models do not adequately deal with this non-stationarity.

Improved statistical techniques should be developed for exploiting the predictability associated with such non-stationary behavior (e.g., Livezey et al., 2007). The use of dynamical models that include a more comprehensive treatment of radiative processes, such as aerosol effects, and also incorporate trends in land use could help improve the quality of dynamical ISI forecasts on longer timescales. As statistical and dynamical models evolve, it will be important to evaluate how much improvement in forecast quality is derived from the trend and how much is derived from model improvements.

REMARKS ON IMPLEMENTATION

The committee also discussed three issues related to the adoption and implementation of the recommendations: the more effective use of many existing observations through improvements to ISI forecast systems, especially as some research-oriented observations transition to operational observations; the role of ISI forecasting as it relates to seamless forecasting across a wide range of space scales and timescales; and, realistic expectations for the types and rate of improvement in ISI forecast quality.

More Effective Use of Observations

Observations are an essential building block of ISI forecast systems. Observations are required to provide initial values for ISI forecasts, to investigate particular processes and develop parameterizations for use in dynamical models, and to validate and verify models. There are many available observations that are not currently being utilized in data assimilation schemes that could contribute to the initialization of dynamical models. Thus, improving ISI forecasting systems offers opportunities to both collect new observations and utilize existing observations in new ways, which can influence decisions regarding the maintenance and upkeep of observational networks.

In recognition of the need to better understand and predict climate change and variability, the number and types of *in situ* and remotely-sensed observations have grown in the last decade under different national and international programs (e.g., the NOAA Climate Program Office, *http://www.climate.noaa.gov/index.jsp?pg=./cp_oa/description.html*, NASA's Earth Observing System (EOS), *http://eospso.gsfc.nasa.gov/*, and the U.S. and intergovernmental Global Climate Observing System (GCOS) draft plan, *http://ioc-goos.org/gcos-ip10draft*). It is a challenge to make effective use of these observations, both in operations as well as research. It is also a challenge to identify observations initiated under research programs that have merit to be continued on an ongoing basis, potentially past the lifetime of the research programs themselves. For operations, one of the more notable challenges is to make use of as much data as possible in the data assimilation process, and subsequently determine the impact of these observations on forecast quality. This will be facilitated by the use of more advanced methods for data assimilation in ISI forecast systems, such as Ensemble Kalman Filter techniques, that are able to adapt the forecast error covariance to the presence of new types of observations. Efforts to improve ISI prediction should work synergistically with efforts to develop and sustain the observing system.

In situ data has value in a number of ways. Increasing our knowledge of processes that affect climate on ISI timescales will require observations targeted on phenomena that are currently either not sampled or not sampled at the appropriate resolution. The concentrated process studies done in climate research programs, in particular the CLIVAR Climate Process Teams or CPTs, are a means to develop both better understanding of processes that may yield ISI predictability and to improve the representation of key processes not explicitly resolved in dynamic models. Long time series, though sparse (especially in the ocean), yield records that can be used to identify biases in dynamic models and to improve the realism of the models' representation of key physical processes. Observations of coupling between the components of the Earth system, in particular, guide the development of more realistic coupled models. Networks of *in situ* observations, such as radiosondes or drifting buoys, provide data for model initialization. In every case, it is necessary to provide metadata (including estimates of observational uncertainty) with observations and to facilitate the timely access to the observations by the modeling community. It remains an ongoing task to facilitate dialog between the observers and the modelers as well as to build and maintain accessible databases.

There are many new satellite products that have become available in the last decade (e.g., EOS A-Train) that include much more detailed information on clouds and aerosols (e.g., CloudSat, CALIPSO, MISR), atmospheric composition (e.g., TES, MLS), ice and snow, and soil moisture (e.g., AMSR, MODIS, GRACE). Considerations have to be made regarding the use that can be made of these now—despite their primary role as research satellites—as well as which elements of their respective data streams should become operational in the future. For example, despite the research-oriented nature of the Microwave Limb Sounder (MLS) on the EOS Aura platform, ECMWF assimilates radiances from MLS to provide more information about the upper troposphere and lower stratosphere. Operational forecast systems should be nimble enough to take advantage of these types of observations.

In particular, information on clouds has yet to be widely used. With the exception of cloud-tracked winds, the bulk of the satellite data employed is often associated with clear skies. However, cloudy conditions often indicate areas of small-scale gradients and inhomogeneities, i.e., locations where more coarse-scale observations are unrepresentative. The variety of available, high-resolution satellite data sets can provide a wealth of information for cloudy areas. This example with cloud data can be generalized to other data sets that have been developed primarily for research purposes but for which technical observing challenges or challenges in assimilating and incorporating such observations into prediction tools might still remain (e.g., precipitation, integrated surface water/ice mass).

Some of the new data sets can also be used to develop advanced diagnostics and metrics, as discussed above, for assessing model performance and guiding model improvement. Strong support should be given to activities that utilize these resources for model improvement, and interaction of the observing efforts with the pertinent forecast modelers should always be considered.

Conversely, some of the suggested improvements to forecast systems may provide guidance for future measurement campaigns. For example, research regarding expanded data assimilation methods could indicate the types and/or spatial and temporal resolution of data sets that could be the target of future measurement missions.

Seamless Forecasting

ISI prediction is the temporal and spatial bridge between numerical weather prediction and climate prediction and, as such, a key component of a seamless prediction system. It is worth highlighting that a seamless forecasting system is not necessarily one that uses the exact same model at all timescales (which, if only for computational reasons, is hardly practical) but a system that, by using the same modeling framework, allows us to understand and trace model biases and errors across timescales. Of concern, for example, are biases seen in ISI forecasting and how they may impact predictions at longer climate time scales that cannot be verified. As part of this concept, ISI prediction is a perfect platform for model development for all timescales (from the short-range to long-term climate) for the following reasons:

- ISI prediction deals with an intrinsically coupled, multi-scale problem. Therefore, coupled models need to be used for ISI prediction, requiring us to properly understand and represent air-sea-land-ice exchanges and coupled variability.

- ISI prediction deals with natural variability and long-term trends. Therefore it requires initialization of the current state of the earth system (atmosphere/land-surface/sea ice/ocean) and the long-term forcings (such as greenhouse gases or aerosols).
- ISI predictions can be verified (as opposed to long-term climate projections), providing a robust mechanism for model validation and improvement: "Fast" physical processes (e.g., convection), low-frequency phenomena (e.g. MJO), and global teleconnections can all be verified against real-time observations.

In order to move closer to seamless prediction and leverage improvements in ISI prediction, transparency among forecast systems is paramount. Through the adoption of Best Practices, efforts to improve ISI predictions can be related back to model development and process knowledge.

Realistic Expectations for Forecast Improvement

Figure 6.1 displays the evolution of forecast skill for the ECMWF atmospheric prediction system from 1980 through 2009. During this period, there have been huge improvements in the forecast model, the observing system, and the data assimilation system. Many of the changes have been revolutionary, for instance the switch to 4D-Variational data assimilation, the availability of observations of the southern hemisphere via satellite measurements, and the direct assimilation of satellite radiances. However, in general the long-term trend in forecast quality progress is slow but monotonic.

It would not be possible to cleanly reproduce Figure 6.1 for ISI forecasts because of the disparity in sample sizes and the greater importance of episodic events such as ENSO on forecast accuracy. If we were to scale time to be forecast samples rather than years, then the progress over the past 25 years has been rapid and dramatic for ISI forecasts. As recently as 1985, although the prediction community was aware that El Niño was important, the objective incorporation of factors perceived as influential on seasonal climate, such as SSTs and soil moisture, were still in the research realm (Gilman, 1985). It was not until the mid-1990s that the National Weather Service's Climate Prediction Center complemented their subjective forecasts based on statistical prediction guidance with objective methods that also considered dynamical ocean-atmosphere models (O'Lenic et al. 2008). The impact of these improvements on forecast quality has not been quantified, however. One approach to doing so is to compare the quality of forecast systems over a common multi-decade period. Unfortunately, very few such studies exist, and those that do often focus on ENSO. These few studies do show that modest improvements have been seen in the forecasts due to improvements in the observational network (Stockdale et al., 2010; Figure 6.2), improved prediction tools (Saha et al., 2006; Figure 4.2), and the combined effects of improvements in the dynamical model and in the assimilated system that provides the ocean initial conditions (Balmaseda et al. 2009; Figure 6.3). These improvements may be synergistic—Stockdale et al. (2010) point out that the season for which the initial conditions from the completed TAO array has the greatest impact is the season when the model has the smallest errors. Similarly Balmaseda et al. (2009) show that the regions that are least

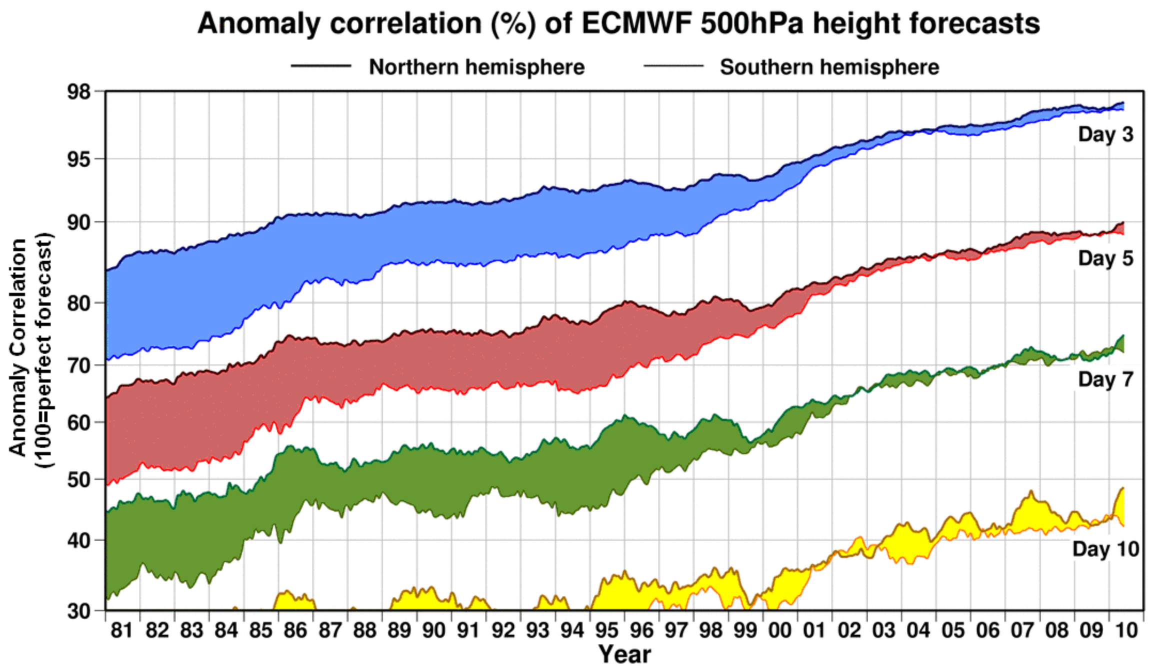

FIGURE 6.1. Evolution of ECMWF forecast skill for varying lead times (3 days in blue; 5 days in red; 7 days in green; 10 days in yellow) as measured by 500-hPa height anomaly correlation. Top line corresponds to the Northern Hemisphere; bottom line corresponds to the Southern Hemisphere. Large improvements have been made, including a reduction in the gap in accuracy between the hemispheres. Identical to Figure 2.1. SOURCE: courtesy of ECMWF, adapted from Simmons and Hollingsworth (2002).

improved by enhancements of the observing network are those where the models have serious biases in the representation of the mean climate.

The earlier sections of this report support the conclusion that ISI forecast quality should continue to slowly improve on average in the future. For example, as operational centers move to more objective methods in translating prediction inputs into issued forecasts (O'Lenic et al. 2008), modest improvements in forecast quality can be expected. The components of the climate system are currently better observed than the tropical Pacific was before the 1980s. It is unlikely, though not impossible, that there are processes with impacts as large as ENSO and the MJO that have not been detected by current observing systems. Similarly, it is unlikely though not inconceivable that available models are failing to simulate some important process that could lead to a revolutionary advance in the quality of ISI predictions. It is more likely that forecasts will improve incrementally with an improved representation or consideration of the sources of predictability (like the MJO or land surface processes); a concerted effort in building better models and better assimilation systems; and, the deployment and use of more observations.

The curves in Figure 6.1 mask one aspect of short-term weather forecasts because they have been monthly-averaged and time-smoothed. There is considerable day-to-day variability in the accuracy of predictions, some of it associated with particular phenomena in the atmosphere.

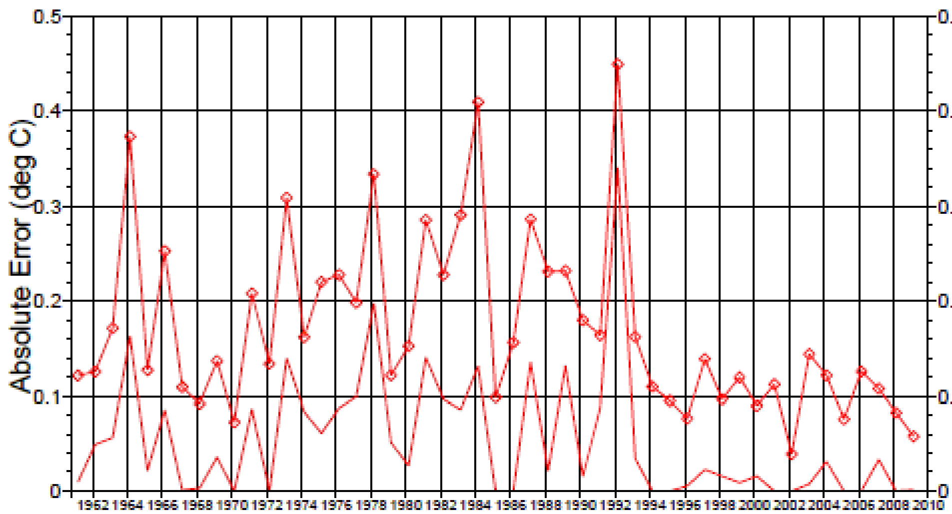

FIGURE 6.2 Time series of Mean Absolute Error (MAE) (thicker line with symbols) for the first three months of NINO3.4 predictions starting 1st February each year. Also shown (thin line, no symbols) is what is referred to as the Best Absolute Error (BAE), which is defined at each lead time as either zero (if the observations lie within the predicted range) or the distance between the observed value and the closest ensemble member, and then averaged over lead times. For a perfect forecasting system with a modest ensemble size, the BAE would be mostly zero, with occasional small positive values. The step change in skill after 1993 is evident. SOURCE: Stockdale et al. (2010), Fig 7a.

In addition, the value of forecasts to users may be far greater in some instances. For instance, the economic value of an accurate 72-hour prediction of hurricane landfall may be far greater than a forecast of fair weather cumulus for the same location. Short-range weather prediction takes advantage of this by committing increased resources to creating and disseminating forecasts of high impact events like hurricanes.

ISI forecasts also exhibit conditional accuracy; for example, forecast quality improves significantly during ENSO events. Forecasts may also be more valuable in certain instances. Given higher predictability and the opportunities as well as catastrophes within the United States associated with ENSO events, a well predicted ENSO event and a reliable forecast of its teleconnections may lead to a very positive net economic impact through reduced disaster losses and increased profits for some sectors (Chagnon, 1999; Goddard and Dilley, 2005). Similarly, MJO events can be intermittent yet have influence over tropical cyclone activity; and thus an accurate MJO forecast can yield valuable foresight into the extremes associated with enhanced or suppressed hurricane activity. More unusual events that impact conditions on intraseasonal and interannual time scales may have even greater economic impact. Unusual events like a major volcanic eruption, the impact of a large body from space, or a nuclear exchange may lead to larger deviations from recent climatology than even the largest ENSO. Producers of ISI predictions could and should be prepared to make short term climate forecasts for such situations (i.e., radical changes in atmospheric composition). Models and forecast generation procedures should be prepared to deal with such events before they happen. Good forecasts of the seasonal

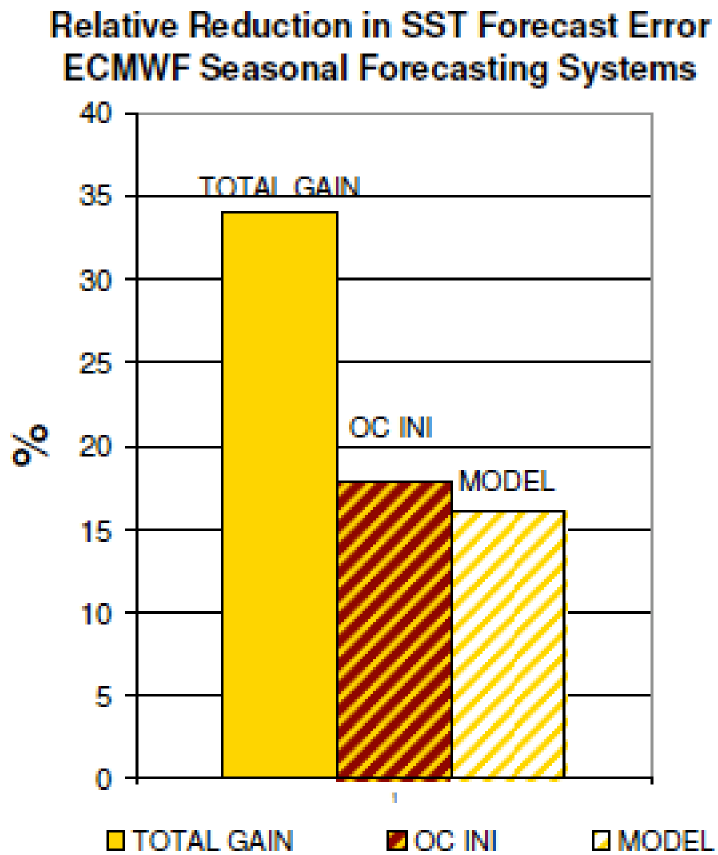

Figure 6.3 Progress in the seasonal forecast skill of the ECMWF operational system during the last decade. The solid bar shows the relative reduction in mean absolute error of forecast of SST in the Eastern Pacific (NINO3). The brown-striped bar shows the contribution from the ocean initialization, and the white-striped bar is the contribution from model improvement. SOURCE: Balmaseda et al. 2009.

response to such unusual events could have far more impact than any forecasts of the undisrupted climate system.

CLOSING REMARKS

The committee's recommendations constitute a strategy to improve the quality of climate predictions at ISI timescales by expanding access to forecasting data and tools; broadening the suite of verification metrics that are used; enhancing collaboration among the operational, research, and user communities; upgrading the building blocks of the ISI forecast systems, which include observations, statistical and dynamical models, and data assimilation techniques; and pursuing research on incompletely understood processes that can contribute to predictability. This strategy is based largely on the lessons learned from historical improvements in the quality of weather and ISI forecasts.

The recommendations have also been crafted to draw on the respective strengths of operational forecast centers and research scientists in the broader community. Considerable expertise in producing and disseminating forecasts exists at the operational centers. Therefore, Best Practices have been designed with their protocols in mind and they will play an integral role upgrading the building blocks of ISI forecast systems. In contrast, the research community is more focused on experimenting with novel ideas, approaches, and techniques. Their role involves expanding our understanding of ISI processes and the tools that are used to measure and simulate these processes. Communication and interaction between these groups will be critical to the improvement of ISI forecast systems.

Finally, the committee stresses that improvements to ISI forecasting systems and improvements in the use of ISI forecasts are possible. In particular, adoption of Best Practices offers a near-term way to aid forecast users and researchers by enhancing access and transparency to forecast information. Incorporating these practices will facilitate more frequent and valuable interaction among these groups. Over the coming years and decades, there are ample opportunities to improve the building blocks of ISI forecast systems and expand our ability to exploit the sources of variability. Although improvements are unlikely to be revolutionary, a coordinated effort by operational centers and the broader research community is likely to yield positive results over time.

References

Agudelo, P. A., C. D. Hoyos, P. J. Webster, and J. A. Curry. 2008. Application of a serial extended forecast experiment using the ECMWF model to interpret the predictive skill of tropical intraseasonal variability. *Climate Dynamics* 32(6):855–872, doi: 10.1007/s00382-008-0447-x.

Allen, M., J. Kettleborough, and D. Stainforth. 2005. Model error in weather and climate forecasting. Pp. 391-427 in *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, eds. Cambridge, UK: Cambridge University Press.

Alves, O., M. Balmaseda, D. Anderson and T. Stockdale. 2004. Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Quarterly Journal of the Royal Meterological Society* 130:647668.

Anderson, J., H. van den Dool, A. Barnston, W. Chen, W. Stern, and J. Ploshay. 1999. Present day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bulletin of the American Meteorological Society* 80:1349–1362.

Anderson, D., T. Stockdale, M. Balmaseda, L. Ferranti, F. Vitart, F.J. Doblas-Reyes, R. Hagedorn, T. Jung, A. Vidard, A. Troccoli, and T.N. Palmer, 2003: Comparison of the ECMWF seasonal forecast systems 1 and 2, including the relative performance for the 1997/8 El Niño. Tech. Memo.404, ECMWF, Reading, United Kingdom, 93 pp.

Anderson D., M. Balmaseda, T. Stockdale, L. Ferranti, F. Vitart, K. Mogensen, F. Molteni, F. Doblas-Reyes, and A. Vidard. 2007. *Development of the ECMWF seasonal forecast System 3.* ECMWF Technical Memorandum 503. Reading, UK: European Centre for Medium-Range Weather Forecasts. Available at *http://www.ecmwf.int/publications/library/do/references/show?id=87744*, accessed June 23, 2010.

Applequist, S., G. E. Gahrs, R. L. Pfeffer, and X-F Niu. 2002. Comparison of methodologies for probabilistic quantitative precipitation forecasting. *Weather and Forecasting* 17:783–799.

Athanasiadis, P. J. and M. H. P. Ambaum. 2009. Linear contribution of different time scales to teleconnectivity. *Journal of Climate* 22, 3720–3728.

Atlas, R., N. Wolfson, and J. Terry. 1993. The effect of SST and soil moisture anomalies on GLA model simulations of the 1988 U.S. summer drought. *Journal of Climate* 6:2034–2048.

Baldwin, M. P., and T. J. Dunkerton. 1999. Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *Journal of Geophysical Research* 104:30937–30946.

Baldwin, M. P., D. W. J. Thompson, E. F.Shuckburgh, W. A. Norton, and N. P. Gillett. 2003a. Weather from the Stratosphere? *Science* 301:317–318.

Baldwin, M. P., D. B. Stephenson, D. W. J.Thompson, T. J. Dunkerton, A. J. Charlton,A. O'Neill. 2003b. Stratospheric memory and extended-range weather forecasts. *Science* 301:636–640.

Balmaseda, M.A., M. K. Davey, and D. L. T. Anderson. 1995. Decadal and Seasonal

Dependence of ENSO Prediction Skill. *Journal of Climate* 8:2705–2715.

Balmaseda, M., A. Vidard, and D. Anderson, 2008: The ECMWF ocean analysis system: ORA-S3. *Monthly Weather Review* 136:3018–3034.

Balmaseda, M. Y. Fujii, O. Alves, T. Lee, M. Rienecker, T. Rosati, D. Stammer, Y. Xue, H. Freeland, M. J. McPhaden, L. Goddard, and C. Coelho. 2009. Role of the ocean observing system in an end-to-end seasonal forecasting system. In *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society* (Vol. 2), Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison, and D. Stammer, eds. ESA Publication WPP-306.

Barnett, T.P. and R. Preisendorfer. 1987. Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Monthly Weather Review* 115:1825–1850.

Barnston, A. G. and R. E. Livezey. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review* 115:1083–1126.

Barnston, A. G. and C. F. Ropelewski. 1992. Prediction of ENSO episodes using canonical correlation analysis. *Journal of Climate* 5:1316–1345.

Barnston, A. G. and Y. He. 1996. Skill of canonical correlation analysis forecasts of 3-month mean surface climate in Hawaii and Alaska. *Journal of Climate* 9:2579–2605.

Barnston, A. G., A. Leetmaa, V. E. Kousky, R. E. Livezey, E. O'Lenic, H. van den Dool, A. J. Wagner, and D. A. Unger. 1999. NCEP Forecasts of the El Niño of 1997–98 and its U.S. Impacts. *Bulletin of the American Meteorological Society* 80:1829–1852.

Barnston, A. G., Y. He, and D. A. Unger. 2000. A forecast product that maximizes utility for state-of-the-art seasonal climate prediction. *Bulletin of the American Meteorological Society*. 81:1271–1279.

Barnston, A. G. and M. Tippett. 2009. An Evaluation of the Skill of ENSO Forecasts during 2002–2009. 34th Annual Climate diagnostics and Prediction Workshop, October 26–30, 2009, Monterey, California, USA.

Barnston, A. G., S. Li, S. J. Mason, D. G. DeWitt, L. Goddard, and X. Gong. 2009. Verification of the first 11 years of IRI's seasonal climate forecasts. *Journal of Climate* (in press).

Barry, R. G. 1995. Observing systems and data sets related to the cryosphere in Canada: A contribution to planning for the Global Climate Observing System. *Atmosphere-Ocean* 33:771–807.

Battisti, D.S. and A.C. Hirst, 1989. Interannual Variability in the Tropical Atmosphere-Ocean System: Influences of the Basic State, Ocean Geometry and Nonlinearity. *Journal of the Atmospheric Sciences* 46:1687-1712.

Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. L. becher, M. J. Rodwell, F. Vitart, and G. Balsamo. 2008. Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society* 134:1337–1351.

Behringer, D.W., M. Ji, and A. Leetmaa. 1998. An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. *Monthly Weather Review* 126(4):1013–1021.

Behringer, D. 2007. Evaluation of the Global Ocean Data Assimilation System (GODAS) at NCEP. 11th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface, American Meteorological Society, San Antonio, Texas, January 14-18.

Berbery, E. H. and J. Nogués-Paegle. 1993. Intraseasonal Interactions between the Tropics and Extratropics in the Southern Hemisphere. *Journal of the Atmospheric Sciences* 50(13):1950-1965.

Berri, G. J., P. L.Antico, and L. Goddard. 2005. Evaluation of the Climate Outlook Forum seasonal precipitation forecasts of Southeast South America during 1998-2002. *International Journal of Climatology* 25:365-377.

Betts, A. K., J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. Viterbo. 1994. Coupling between Land-Surface, Boundary-Layer Parameterizations and Rainfall on Local and Regional Scales: Lessons from the Wet Summer of 1993. Fifth Conference on Global Change Studies, 74th Annual Meeting, American Meteorological Society, Nashville, TN, January 23–28.

Bjerknes, J. 1969. Atmospheric teleconnections from the equatorial Pacific. *Monthly Weather Review* 97:163–172.

Blackmon M. L., Y.-H. Lee, and J. M. Wallace. 1984. Horizontal structure of 500 mb height fluctuations with long, intermediate and short time scales. *Journal of the Atmospheric Sciences* 41:961–980.

Blanke, B., J. D. Neelin, and D. Gutzler. 1997. Estimating the effect of stochastic wind stress forcing on ENSO irregularity. *Journal of Climate* 10:1473–1486.

Boer, G.J., and K. Hamilton. 2008. QBO influence on extratropical predictive skill. *Climate Dynamics* 31:987–1000.

Bollasina, M. and S.Nigam. 2009. Absorbing aerosols and pre-summer monsoon hydroclimate variability over the Indian subcontinent: The challenge in investigating links. *Atmospheric Research* 94:338–344.

Bond, N. A. and G. A. Vecchi. 2003. The influence of the Madden-Julian oscillation on precipitation in Oregon and Washington. *Weather and Forecasting* 18(4):600–613.

Boulanger, J.-P., E. Durand, J.-P. Duvel, C. Menkes, P. Delecluse, M. Imbard, M. Lengaigne, G. Madec, and S. Masson. 2001. Role of non-linear oceanic processes in the response to westerly wind events: New implications for the 1997 El Niño onset. Geophysical Research Letters 28(8):1603–1606.

Bowler, N. E., A. Arribas, K.R. Mylne, K.B. Robertson, and S.E. Beare. 2008. The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society* 134:703–722.

Bradley, R. S. 1988. The explosive volcanic eruption signal in northern hemisphere continental temperature records. *Climatic Change* 12(3):221–243.

Brankovic, C., T. N. Palmer, and L. Ferranti. 1994. Predictability of seasonal atmospheric variations. *Journal of Climate* 7:217–237.

Branstator, G. 1987: A striking example of the atmosphere's leading traveling pattern. *Journal of the Atmospheric Sciences* 44:2310–2323.

Brönnimann S, E. Xoplaki, C. Casty, A. Pauling, and J. Luterbacher. 2007. ENSO influence on Europe during the last centuries. *Climate Dynamics* 28(2-3):181-197.

Brunet, N., R. Verret, and N. Yacowar. 1988. An objective comparison of model output statistics and "perfect prog" systems in producing numerical weather elements forecasts. *Weather and Forecasting* 3:273–283.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He. 2009a. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic

NWP. Part I: Description of configurations and results from idealized experiments. *Monthly Weather Review* (in press) doi: 10.1175/2009MWR3157.1.
Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He. 2009b. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: Results from 1-month experiments with real observations. *Monthly Weather Review* (in press) doi: 10.1175/2009MWR3158.1.
Buizza, R. 1997. Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Monthly Weather Review* 125:99–119.
Buizza, R., M. Miller, and T. N. Palmer. 1999. Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society* 125:2887–2908.
Burgers , G., F.-F. Jin, and G. J. van Oldenborgh. 2005. The simplest ENSO recharge oscillator. *Geophysical Research Letters* 32:L13706, doi:10.1029/2005GL022951.
Cane, M. A. 1986. El-Niño. *Annual Review of Earth and Planetary Sciences* 14:43–70.
Cane, M. A., S. E. Zebiak, and S. C. Dolan. 1986. Experimental Forecasts of El-Niño. *Nature* 321(6073):827–832.
Cannon, A. J., P. H. Whitfield, and E. R. Lord. 2002. Synoptic map-pattern classification using recursive partitioning and principal component analysis. *Monthly Weather Review* 130:1187–1206.
Cavazos, T., A. C. Comrie, and D. M. Liverman. 2002. Intraseasonal variability associated with wet monsoons in southeast Arizona. *Journal of Climate* 15:2477–2490.
CCSP (U.S. Climate Change Science Program). 2008. Decision-Support Experiments and Evaluations using Seasonal-to-Interannual Forecasts and Observational Data: A Focus on Water Resources. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research, N. Beller-Simms, H. Ingram, D. Feldman, N. Mantua, K. L. Jacobs, and A. M. Waple, eds. Asheville, NC: NOAA National Climatic Data Center. 192 pp.
Changnon, S. A. 1999. Impacts of 1997-98 El Niño-generated weather in the United States. *Bulletin of the American Meteorological Society* 80:1819–1927.
Charney, J. G. and J. Shukla. 1981. Predictability of monsoons. Pp. 99–109 in *Monsoon Dynamics*, J. Lighthill and R. P. Pearce, eds. New York: Cambridge University Press.
Chelton, D. B., S. K. Esbensen, M. G. Schlax, N. Thum, and M. H. Frieleich. 2001. Observations of coupling between the surface wind stress and sea surface temperature in the eastern tropical Pacific. *Journal of Climate* 14:1479–1498.
Chen, D., S.E. Zebiak., A.J. Busalacchi, M.A. Cane. 1999. An Improved Procedure for EI Nino Forecasting: Implications for Predictability. *Science* 269(5231):1699-702.
Chen, D. and X. Yuan. 2004. A Markov Model for Seasonal Forecast of Antarctic Sea Ice. *Journal of Climate* 17:3156–3168.
Chen, S. S., R. A. Houze, and B. E. Mapes. 1996. Multiscale variability of deep convection in relation to large-scale circulation in TOGA COARE. *Journal of the Atmospheric Sciences* 53:1380–1409.
Chen, T. C., and J. C. Alpert. 1990. Systematic-Errors in the Annual and Intraseasonal Variations of the Planetary-Scale Divergent Circulation in NMC Medium-Range Forecasts. *Monthly Weather Review* 118(12):2607–2623.

Chen, W. Y. and H. M. van den Dool. 1997. Atmospheric predictability of seasonal, annual, and decadal climate means and the role of the ENSO cycle: A model study. *Journal of Climate* 10:1236–1254.

Cherry, S. 1996. Singular value decomposition analysis and canonical correlation analysis. *Journal of Climate* 9:2003–2009.

Chu, P.C. 1999. Two Kinds of Predictability in the Lorenz System. *Journal of the Atmospheric Sciences* 56:1427-1432.

CLIVAR Madden-Julian Oscillation Working Group. 2009. Establishing Operational Model MJO Forecasts. *Bulletin of the American Meteorological Society* (In preparation).

Cobb, K. M., C. D. Charles, H. Cheng, and R. L. Edwards. 2003. El Niño/Southern Oscillation and tropical Pacific climate during the last millennium. *Nature* 424:271–276.

Codron, F., A. Vintzileos, and R. Sadourny. 2001. Influence of Mean State Changes on the Structure of ENSO in a Tropical Coupled GCM. *Journal of Climate* 14(5):730–742.

Cohen, J. and C. Fletcher. 2007. Improved skill of Northern Hemisphere winter surface temperature predictions based on land-atmosphere fall anomalies. *Journal of Climate* 20:4118–4132.

Collins, M. 2000. Understanding uncertainties in the response of ENSO to greenhouse warming. *Geophysical Research Letters* 27(21):3509-3512.

Cook, B. I., G. B. Bonan, and S. Levis. 2006. Soil moisture feedbacks to precipitation in Southern Africa. *Journal of Climate* 19:4198–4206.

Courtier, P., E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quarterly Journal of the Royal Meteorological Society* 550:1783–1807.

Daley, R. 1993. *Atmospheric Data Analysis*. New York: Cambridge University Press.

Davison, P. S., D. L. Roberts, R. T. Arnold, and R. N. Colvile. 2004. Estimating the direct radiative forcing due to haze from the 1997 forest fires in Indonesia. *Journal of Geophysical Research* 109:D10207, doi:10.1029/2003JD004264.

DelSole, T. 2007. A Bayesian Framework for Multimodel Regression. *Journal of Climate* 20:2810–2826.

Delworth, T. L. and S. Manabe. 1989. The influence of soil wetness on near-surface atmospheric variability. *Journal of Climate* 2:1447–1462.

Derber, J. and A. Rosati. 1989. A global oceanic data assimilation system. *Journal of Physical Oceanography* 19:1333–1347.

Deser, C. and M. L. Blackmon. 1993. Surface climate variations over the North Atlantic ocean during winter: 1900–1989. *Journal of Climate* 6:1743–1753.

DeWitt, D. G. 2005. Retrospective Forecasts of Interannual Sea Surface Temperature Anomalies from 1982 to Present Using a Directly Coupled Atmosphere–Ocean General Circulation Model. *Monthly Weather Review* 133:2972–2995.

Dickson, R. R. and Integrated Arctic Ocean Observing System (iAOOS) PIs. 2009. *The Integrated Arctic Ocean Observing System (iAOOS) in 2008*. Vienna, VA: Arctic Ocean Sciences Board. Available at *http://aosb.arcticportal.org/pdf/iAOOS2008_final.pdf*, accessed August 19, 2010.

Ding, Q. and B. Wang. 2005. Circumglobal teleconnection in Northern Hemisphere summer. *Journal of Climate* 18:3483–3505.

Dirmeyer, P. A. 2000. Using a global soil wetness dataset to improve seasonal climate simulation. *Journal of Climate* 13:2900–2922.

Doblas-Reyes, F. J., M. Déqué, and J.-Ph. Piedelievre. 2000. Model and multimodel spread and probabilistic seasonal forecasts in PROVOST. *Quarterly Journal of the Royal Meteorological Society* 126:2069–2088.

Doblas-Reyes, F.J., R. Hagedorn and T.N. Palmer. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting. Part II: Calibration and combination. *Tellus A* 57:234-252.

Donlon, C., K. Casey, C. Gentemann, P. LeBorgne, I. Robinson, R. Reynolds, C. Merchant, D. Llewellyn-Jones, P. Minnett, J. Piolle, P. Cornillon, N. Rayner, T. Brandon, J. Vazquez, E. Armstrong, H. Beggs, I. Barton, G. Wick, S. Castro, J. Hoeyer, D. May, O. Arino, D. Poulter, R. Evans, C. Mutlow, A. Bingham, and A. Harris. 2009. Successes and Challenges for the Modern Sea Surface Temperature Observing System. In *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society*, Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison, and D. Stammer, eds. ESA Publication WPP-306.

Douville, H. and F. Chauvin. 2000. Relevance of soil moisture for seasonal climate predictions, A preliminary study. *Climate Dynamics* 16:719–736.

Douville, H., F. Chauvin, and H. Broqua. 2001. Influence of soil moisture on the Asian and African monsoons, Part 1, Mean monsoon and daily precipitation. *Journal of Climate* 14:2381–2403.

Drusch, M. and P. Viterbo. 2007. Assimilation of Screen-Level Variables in ECMWF's Integrated Forecast System: A Study on the Impact on the Forecast Quality and Analyzed Soil Moisture. *Monthly Weather Review* 135:300–314.

ECMWF (European Centre for Medium-range Weather Forecasts). 2004. ECMWF/CLIVAR Workshop on simulation and prediction of intra-seasonal variability with emphasis on the MJO, Reading, UK, 3-6 November 2003. ECMWF Workshop Proceedings. Reading, UK, 2004. 269pp. Available at *http://www.ecmwf.int/publications/library/do/references/list/17124*, accessed August 19, 2010.

Efron, B. and R. J. Tibshirani. 1993. *An introduction to the bootstrap.* Boca Raton, FL: CRC Press.

Eisenman, I., L. Yu, and E. Tziperman. 2005. Westerly wind bursts: ENSO's tail rather than the dog? *Journal of Climate* 18:5224–5238.

English, S. J., R. J. Renshaw, P. C. Dibben, A. J. Smith, P. J. Rayer, C. Poulsen, F. W. Saunders, and J. R. Eyre. 2000. A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Quarterly Journal of the Royal Meteorological Society* 126:2911–2931.

Enloe, J., J.J. O'Brien, S.R. Smith, 2004. ENSO impacts on peak wind gusts in the United States. *Journal of Climate* 17:1728–1737.

Entin, J. K, A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkhai. 2000. Temporal and spatial scales of observed soil moisture variations in the extratropics. *Journal of Geophysical Research* 105:11865–11877.

Feddersen, H., A. Navarra, and M. N. Ward. 1999. A statistical correction approach for dynamical seasonal predictions. *Journal of Climate* 12:1974–1989.

Fedorov, A. V. and S. G. Philander. 2001. A Stability Analysis of Tropical Ocean–Atmosphere Interactions: Bridging Measurements and Theory for El Niño. *Journal of Climate* 14:3086–3101.

Fennessey, M. J., and J. Shukla. 1999. Impact of initial soil wetness on seasonal atmospheric prediction. *Journal of Climate* 12:3167–3180.

Ferranti, L., T. N. Palmer, F. Molteni, and K. Klinker. 1990. Tropical-extratropical interaction associated with the 30-60-day oscillation and its impact on medium and extended range prediction. *Journal of the Atmospheric Sciences* 47:2177–2199.

Findell, K. L., and E. A. B. Eltahir. 2003. Atmospheric controls on soil moisture—boundary layer interactions, Part II, Feedbacks withing the continental United States. *Journal of Hydrometeorology* 4:570–583.

Fischer, E. M., S. I. Seneviratne, P. L. Vidale, D. Luthi, and C. Schar. 2007. Soil moisture—atmosphere interactions during the 2003 European summer heat wave. *Journal of Climate* 20:5081–5099.

Flügel, M., P. Chang, C. Penland. 2004. The Role of Stochastic Forcing in Modulating ENSO Predictability. *Journal of Climate* 17:3125-3140.

Freeland, H., D. Roemmich, S. Garzoli, P. LeTraon, M. Ravichandran, S. Riser, V. Thierry, S. Wijffels, M. Belbéoch, J. Gould, F. Grant, M. Ignazewski, B. King, B. Klein, K. Mork, B. Owens, S. Pouliquen, A. Sterl, T. Suga, M. Suk, P. Sutton, A. Troisi, P. Vélez-Belchi, and J. Xu. 2009. Argo: A Decade of Progress. In *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society* (Vol. 2), Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison, and D. Stammer, eds. ESA Publication WPP-306.

Frías, M. D., S. Herrera, A. S. Cofiño, J. M. Gutiérrez. 2010. Assessing the Skill of Precipitation and Temperature Seasonal Forecasts in Spain: Windows of Opportunity Related to ENSO Events. *Journal of Climate* 23:209-220.

Fu, X., B. Wang, T. Li, and J. P. McCreary. 2003. Coupling between Northward-Propagating, Intraseasonal Oscillations and Sea Surface Temperature in the Indian Ocean. *Journal of the Atmospheric Sciences* 60:1733–1753.

Fu, X., B. Wang, D. E. Waliser, and T. Li. 2006. Impact of Atmosphere-Ocean Coupling on the Predictability of Monsoon Intraseasonal Oscillations (MISO). *Journal of the Atmospheric Sciences* 64:157–174.

Fu, X., and B. Wang. 2009. Critical Roles of the Stratiform Rainfall in Sustaining the Madden-Julian Oscillation: GCM Experiments. *Journal of Climate* 22:3939-3959.

Fu, X., B. Wang, Q. Bao, P. Liu, and J.-Y. Lee. 2009. Impacts of initial conditions on monsoon intraseasonal forecasting. *Geophysical Research Letters* 36:L08801, doi:10.1029/2009GL037166.

Gebbie, G., I. Eisenman, A. Wittenberg, and E. Tziperman. 2007. Modulation of Westerly Wind Bursts by Sea Surface Temperature: A Semistochastic Feedback for ENSO. *Journal of the Atmospheric Sciences* 64:3281–3295.

Gent, P. R., and J. C. McWilliams. 1990. Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography* 20:150–155.

Gerrity, J. P. 1992. A note on Gandin and Murphy's equitable skill score. *Monthly Weather Review* 120:2709–2712.

Gill, A. E. 1982. Atmosphere-Ocean Dynamics. San Diego, CA: Academic Press.

Gilman, D.L. 1985. Long-range forecasting: The present and the future. *Bulletin of the American Meteorological Society* 66:159–164.

Giorgi, F. and R. Francisco. 2000. Uncertainties in regional climate change predictions. A regional analysis of ensemble simulations with the HadCM2 coupled AOGCM. *Climate Dynamics* 16:169–182.

Glahn, H. R. and D. A. Lowry 1972. The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology* 11:1203–1211.

Glahn. H, R. 1984. Statistical weather forecasting. In *Probability, Statistics,and Decision Making in the Atmospheric Sciences,* A. H. Murphy and R. W. Katz, eds. Boulder, CO: Westview Press.

Glantz, M. H. 1996. *Currents of Change: El Niño's Impact on Climate and Society*. New York: Cambridge University Press.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux. 2008. Performance metrics for climate models. *Journal of Geophysical Research* 113:D06104, doi:10.1029/2007JD008972.

Goddard, L. and S. J. Mason. 2002. Sensitivity of seasonal climate forecasts to persisted SST anomalies. *Climate Dynamics* 19:619-632, doi:10.1007/s00382-002-0251-y.

Goddard, L. and D. G. DeWitt. 2005. Seeking progress in El Niño Prediction. *U.S. CLIVAR Variations* 3(1).

Goddard, L. and M. Dilley. 2005. El Niño: Catastrophe or opportunity? *Journal of Climate* 18:651–665.

Goddard, L. and M. P. Hoerling. 2006. Practices for Seasonal-to-Interannual Climate Prediction. *U.S. CLIVAR Variations* 4(3).

Gong, X. and M. B. Richman. 1995. On the application of cluster analysis to growing season precipitation in North America east of the Rockies. *Journal of Climate* 8:897–931.

Gottschalck, J., M. Wheeler, K. Weickmann, F. Vitart, N. Savage, H. Lin, H. Hendon, D. Waliser, K. Sperber, M. Nakagawa, C. Prestrelo, M. Flatau, and W. Higgins. 2010. A Framework for Assessing Operational Model MJO Forecasts: A Project of the CLIVAR Madden-Julian Oscillation Working Group. *Bulletin of the American Meteorological Society* Early Online Release, available at *http://journals.ametsoc.org/doi/pdf/10.1175/2010BAMS2816.1*. Accessed July 7, 2010.

Goswami, B. N. and J. Shukla. 1991. Predictability of a coupled ocean-atmosphere model. *Journal of Climate* 4:3–22.

Goswami, B. N. and P. Xavier. 2003. Potential predictability and extended range prediction of Indian summer monsoon breaks. *Geophysical Research Letters* 30(18) doi:1910.1029/2003GL1017,1810.

Goswami, B. N. 2005. South Asian Summer Monsoon. In *Intraseasonal Variability of the Atmosphere-Ocean Climate System* W. K. M. Lau and D. E. Waliser, eds. Heidelberg, Germany: Springer.

Graf, H.-F., J. Yang, and T. M. Wagner. 2009. Aerosol effects on clouds and precipitation during the 1997 smoke episode in Indonesia. *Atmospheric Chemistry and Physics* 9:743–756, doi:10.5194/acp-9-743-2009.

Graham, N. E., J. Michaelsen, and T. P. Barnett. 1987. An Investigation of the El Niño-Southern Oscillation Cycle with Statistical Models 2. Model Results. *Journal of Geophysical Research* 92(C13):14,271–14,289.

Graham, R., C. Gordon, M. Huddleston, M. Davey, W. Norton, A. Colman, A. Scaife, A. Brookshaw, B. Ingleby, P. McLean, K. Groves, D. Cotgrove, D. Robinson, E. McCallum, and W. Elliot. 2006. The 2005/6 winter in Europe and the United Kingdom: Part 1—How the Met Office forecast was produced and communicated. *Weather* 61(12):327–336.

Gualdi, S., A. Alessandri, and A. Navarra. 2005. Impact of atmospheric horizontal resolution on El Niño Southern Oscillation forecasts. *Tellus* 57A:357–374.

Guilyardi, E. 2006. El Niño–mean state–seasonal cycle interactions in a multi-model ensemble. *Climate Dynamics* 26:329–348.

Guo, Z., P. A. Dirmeyer, R. D. Koster, Y. C. Sud, G. Bonan, K. W. Oleson, E. Chan, D. Verseghy, P. Cox, C. T. Gordon, J. L. McGregor, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu and D. Mocko, C.-H. Lu, K. Mitchell, S. Malyshev, B. McAvaney, T. Oki, T. Yamada, A. Pitman, C. M. Taylor, R. Vasic, and Y. Xue. 2006. GLACE, The global land-atmosphere coupling experiment. Part II: Analysis. *Journal of Hydrometeorology* 7:611–625.

Hagedorn, R., F. J. Doblas-Reyes, T. N. Palmer. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting. Part I: The basic concept. *Tellus* 57A:219–233.

Hamill, T. M., J. S. Whitaker and X. Wei. 2004. Ensemble reforecasting: improving medium range forecast skill using retrospective forecasts. *Monthly Weather Review* 132:1434–1447.

Haigh, J. D., M. Blackburn, and R. Day. 2005. The Response of Tropospheric Circulation to Perturbations in Lower-Stratospheric Temperature. *Journal of Climate* 18:3672–3685.

Hanley, D. E., M. A. Bourassa, J. J. O'Brien, S. R. Smith, and E. R. Spade. 2003. A quantitative evaluation of ENSO indices. *Journal of Climate* 16:1249–1258.

Haupt, S. E., A. Pasini, and C. Marzban. 2009. *Artificial Intelligence Methods in the Environmental Sciences*. Heidelberg, Germany: Springer.

Hayes, S. P., L. J. Mangum, J. Picaut, A. Sumi, and K. Takeuchi. 1991. TOGA-TAO: A moored array for real-time measurements in the tropical Pacific Ocean. *Bulletin of the American Meteorological Society* 72:339–347.

Hendon, H. H., and M. L. Salby. 1994. The Life-Cycle of the Madden-Julian Oscillation. *Journal of the Atmospheric Sciences* 51(15):2225–2237.

Hendon, H. H. and J. Glick. 1997. Intraseasonal air-sea interaction in the tropical Indian and western Pacific Oceans. *Journal of Climate* 10:647–661 .

Hendon, H. H., B. Liebmann, M. Newman, and J. D. Glick. 2000. Medium-range forecast errors associated with active episodes of the Madden-Julian oscillation. *Monthly Weather Review* 128(1):69.

Higgins, R. W., J. K. E. Schemm, W. Shi, and A. Leetmaa. 2000. Extreme precipitation events in the western United States related to tropical forcing. *Journal of Climate* 13(4):793–820.

Higgins, R. W., A. Leetmaa, and V. E. Kousky. 2002. Relationships between climate variability and winter temperature extremes in the United States. *Journal of Climate* 15:1555–1572.

Higgins, R. W., V. B. S. Silva, W. Shi, and J. Larson. 2007. Relationships between climate variability and fluctuations in daily precipitation over the United States. *Journal of Climate* 20:3561–3579.

Hirsch, M. E., A. T. DeGaetano, and S. J. Colucci. 2001. An east coast winter storm climatology. *Journal of Climate* 14:882–899.

Hoffman, M. J., E. Kalnay, J. A. Carton, and S.-C. Yang. 2009. Use of breeding to detect and explain instabilities in the global ocean. *Geophysical Research Letters* 36:L12068.

Hong, S.-Y. and E. Kalnay. 2000. Role of sea surface temperature and soil-moisture feedback in the 1998 Oklahoma-Texas drought. *Nature* 408:842–844.

Horel, J. D. and J. M. Wallace. 1981. Planetary-Scale Atmospheric Phenomena Associated with the Southern Oscillation. *Monthly Weather Review* 109:813–829.

Hoskins, B.J., and D.J. Karoly. 1981. The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing. *Journal of the Atmospheric Sciences* 38:1179–1196.

Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24:417–520.

Houtekamer, P. L., and H. L. Mitchell. 2005. Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society* 131:3269–3289.

Huang, B., Y. Xue, and D. Behringer. 2008. Impacts of Argo salinity in NCEP Global Ocean Data Assimilation System: The tropical Indian Ocean. *Journal of Geophysical Research* 113, C08002.

Huang, J., H. M. van den Dool, and A. G. Barnston. 1996a. Long-lead seasonal temperature prediction using optimal climate normals. *Journal of Climate* 9:809–817.

Huang, J., H. M. van den Dool, and K. Georgakakos. 1996b. Analysis of model-calculated soil moisture over the United States (1931–93) and applications to long-range temperature forecasts. *Journal of Climate* 9:1350–1362.

Huang, J., D. Wang, and C.-H. Zhang. 2005. A two-way semilinear model for normalization and analysis of cDNA microarray data. *Journal of the American Statistical Association* 100:814-829.

Hurrell, J. W. 1995. Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science* 269:676–679.

Hsu, H. H. 2005. East Asian Monsoon. In *Intraseasonal Variability of the Atmosphere-Ocean Climate System*, W. K. M. Lau and D. E. Waliser, eds. Heidelberg, Germany: Springer.

ICTP (International Centre for Theoretical Physics). 2006. *Workshop on the Organization and Maintenance of Tropical Convection and the Madden Julian Oscillation*. International Centre for Theoretical Physics, Trieste, Italy, March 13-17, 2006.

IGOS (Integrated Global Observing Strategy). 2007. *For the Monitoring of our Environment from Space and from Earth: An international partnership for cooperation in Earth observations*. WMO/TD-No. 1405. Geneva: World Meteorological Organization. 100 pp.

Ineson, S. and A. A. Scaife. 2009. The role of the stratosphere in the European climate response to El Niño. *Nature Geoscience* 2:32–36.

IPCC (Intergovernmental Panel on Climate Change). 2007. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller. New York: Cambridge University Press.

Jarvis, C. H. and N. Stuart, 2001: A comparison among strategies for interpolating maximum and minimum daily air temperature. Part II: the interaction between number of guiding variables and the type of interpolation method. *Journal of Applied Meteorology* 40:1075–1084.

Jhun, J.-G. and E.-J. Lee. 2004. A new East Asian winter monsoon index and associated characteristics of winter monsoon. *Journal of Climate* 17:711–726.

Ji, M., A. Kumar, and A. Leetmaa. 1994. An experimental coupled forecast system at the National Meteorological Center: Some early results. *Tellus* 46A:398–418.

Ji, M., A. Leetmaa, and J. Derber. 1995. An ocean analysis system for seasonal to interannual climate studies. *Monthly Weather Review* 123:460–481.

Ji, M., and V. E. Kousky. 1996. Coupled model prediction of ENSO during the 1980s and the 1990s at the National Centers for Environmental Prediction. *Journal of Climate* 9:3105–3120.

Ji, M., D. W. Behringer and A. Leetmaa. 1998. An improved coupled model for ENSO prediction and implications for ocean initialization. Part II: The coupled model. *Monthly Weather Review* 126:1022–1034.

Jiang, X., D. E. Waliser, W. S. Olson, W.-K. Tao, T. S. L'Ecuyer, J.-L. Li, B. Tian, Y. L. Yung, A. M. Tompkins, S. E. Lang, and M. Grecu. 2009. Vertical Heating Structures Associated with the MJO as Characterized by TRMM Estimates, ECMWF Reanalyses, and Forecasts: A Case Study during 1998/99 Winter. *Journal of Climate*, 22:6001-6020.

Jin, E. K., J. K. Kinter III, B. Wang, C.-K. Park, L.-S. Kang, B. P. Kirtman, J.-S. Kug, A. Kumar, J.-J. Luo, J. Schemm, J. Shukla, and T. Yamagata. 2008. Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Climate Dynamics* 31:647–664, doi:10.1007/s00382-008-03973.

Jin, F.-F. 1997. An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. *Journal of the Atmospheric Sciences* 54:811–829.

Jin, F.-F., L. Lin, A. Timmermann, and J. Zhao. 2007. Ensemble-mean dynamics of the ENSO recharge oscillator under state-dependent stochastic forcing. *Geophysical Research Letters* 34:L03807, doi:10.1029/2006GL027372.

Jolliffe, I. T. and D. B. Stephenson. 2003. Forecast Verification: A Practitioner's Guide in Atmospheric Science. Chichester, U.K.: John Wiley and Sons Ltd.

Johansson, A., A. Barnston, S. Saha, and H. van den Dool. 1998. On the level and origin of seasonal forecast skill in northern Europe. *Journal of the Atmospheric Sciences* 55:103–127.

Johansson, A. 2007. Prediction Skill of NAO and PNA from Daily to Seasonal Timescales. *Journal of Climate* 20, 1957–1975.

Jones, C., D. E. Waliser, J.-K. E. Schemm, and W. K. M. Lau. 2000. Prediction skill of the Madden and Julian Oscillation in dynamical extended range forecasts. *Climate Dynamics* 16(4):273–289.

Jones, C., L. M. V. Carvalho, R. W. Higgins, D. E. Waliser, and J.-K. E. Schemm. 2003. Climatology of Tropical Intraseasonal Convective Anomalies: 1979-2002. *Journal of Climate* 17:523–539.

Jones, C., D. E. Waliser, K.-M. Lau, and W. Stern. 2004a. Global Occurrences of Extreme Precipitation and the Madden-Julian Oscillation: Observations and Predictability. *Journal of Climate* 17:4575–4589.

Jones, C., D. E. Waliser, K.-M. Lau, and W. Stern. 2004b. The Madden-Julian Oscillation and its Impact on Northern Hemisphere Weather Predictability. *Monthly Weather Review* 132(6):1462–1471.

Kalnay, E. 2003. Atmospheric modeling, data assimilation and predictability. New York: Cambridge University Press.

Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera. 2007. 4D-Var or Ensemble Kalman Filter? *Tellus* 59A:758–773.

Kang, I.-S. and J. Shukla. 2005. Dynamical seasonal prediction and predictability of monsoon. Pp. 585–612 in *The Asian Monsoon*, B. Wang, ed. Berlin: Springer.

Karl, T. R. 1986. The Relationship of Soil Moisture Parameterizations to Subsequent Seasonal and Monthly Mean Temperature in the United States. *Monthly Weather Review* 114:675-

686.

Katz, R. W. 2002. Sir Gilbert Walker and a connection between El Niño and statistics. *Statistical Science* 17:97–112.

Keenlyside, N., M. Latif, M. Botzet, J. Jungclaus, and U. Schulzweide. 2005. A coupled method for initializing El Niño Southern Oscillation forecasts using sea surface temperature. *Tellus* 57A:340–356.

Kemball-Cook, S., and B. Wang. 2001. Equatorial waves and air-sea interaction in the boreal summer intraseasonal oscillation. *Journal of Climate* 14:2923-2942.

Kemball-Cook, S., B. Wang, and X. Fu. 2002. Simulation of the intraseasonal oscillation in the ECHAM-4 model: The impact of coupling with an ocean model. *Journal of the Atmospheric Sciences* 59(9):1433–1453.

Keppenne, C. L. and M. M. Rienecker. 2002. Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Monthly Weather Review* 130:2951–2965.

Kessler, W.S. and R. Kleeman, 2000. Rectification of the Madden–Julian Oscillation into the ENSO Cycle. *Journal of Climate* 13:3560–3575.

Kessler, W.S. 2001. EOF representations of the Madden-Julian Oscillation and its connection with ENSO. *Journal of Climate* 14:3055-3061.

Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy. 2009. Convectively coupled equatorial waves. *Reviews of Geophysics* 47:RG2003, doi:10.1029/2008RG000266.

Kim, D., K. Sperber, W. Stern, D. Waliser, I.-S. Kang, E. Maloney, S. Schubert, W. Wang, K. Weickmann, J. Benedict, M. Khairoutdinov, M.-I. Lee, R. Neale, M. Suarez, K. Thayer-Calder, and G. Zhang. 2009: Application of MJO Simulation Diagnostics to Climate Models, *Journal of Climate* 22: 6413-6436.

Kirtman, B. P. 2003. The COLA anomaly coupled model: Ensemble ENSO prediction. *Monthly Weather Review* 131:2324–2341.

Kirtman, B. P., 1997: Oceanic Rossby Wave Dynamics and the ENSO Period in a Coupled Model. *Journal of Climate* 10:1690-1704.

Kirtman, B. P., J. Shukla, B. Huang, Z. Zhu, and E. K. Schneider 1997. Multiseasonal predictions with a coupled tropical ocean global atmosphere system. *Monthly Weather Review* 125:789–808.

Kirtman, B. P. and P. S. Schopf. 1998. Decadal variability in ENSO predictability and prediction. *Journal of Climate* 11:2804–2822.

Kirtman, B. P., K. Pegion, and S. Kinter. 2005. Internal atmospheric dynamics and climate variability. *Journal of the Atmospheric Sciences* 62:2220–2233.

Kirtman, B. and A. Pirani. 2008. WCRP Position Paper on Seasonal Prediction: Report from the First WCRP Seasonal Prediction Workshop, June 4–7, 2007, Barcelona, Spain. WCRP Informal Report No. 3/2008, ICPO Publication No. 127.

Kirtman, B. and A. Pirani. 2009. The State of the Art of Seasonal Prediction: Outcomes and Recommendations from the First World Climate Research Program Workshop on Seasonal Prediction. *Bulletin of the American Meteorological Society* 90:455–458.

Kirtman, B. P. and D. Min. 2009. Multimodel Ensemble ENSO Prediction with CCSM and CFS. *Monthly Weather Review* 137(9):2908–2930.

Kleeman, R., A. M. Moore, N. R. Smith. 1995. Assimilation of subsurface thermal data into a simple ocean model for the initialization of an intermediate tropical coupled ocean-atmosphere forecast model. *Monthly Weather Review* 123:3103–3113.

Kleeman, R., Y. Tang, and A. M. Moore. 2003. The calculation of climatically relevant singular vectors in the presence of weather noise as applied to the ENSO problem. *Journal of the Atmospheric Sciences* 60:2856–2868.

Klein, B. M., F. Lewis, and I. Enger. 1959. Objective prediction of five-day mean temperature during winter. *Journal of Meteorology* 16:672–682.

Koster, R. D., M. J. Suarez, and M. Heiser. 2000. Variance and predictability of precipitation at seasonal-to-interannual timescales. *Journal of Hydrometeorology* 1:26–46.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, A. Berg, M. Kistler, R. Reichle, M. Rodell, and J. Famiglietti. 2004. Realistic initialization of land surface states: Impacts on subseasonal forecast skill. *Journal of Hydrometeorology* 5:1049–1063.

Koster, R. D., Z. Guo, P. A. Dirmeyer, G. Bonan, E. Chan, P. Cox, H. Davies, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C.-H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. W. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada. 2006. GLACE: The global land-atmosphere coupling experiment. Part I: Overview. *Journal of Hydrometeorology* 7:590–610.

Koster, R. D., S. P. P. Mahanama, T. J. Yamada, G. Balsamo, A. A. Berg, M. Boisserie, P. A. Dirmeyer, F. J. Doblas-Reyes, G. Drewitt, C. T. Gordon, Z. Guo, J.-H. Jeong, D. M. Lawrence, W.-S. Lee, Z. Li, L. Luo, S. Malyshev, W. J. Merryfield, S. I. Seneviratne, T. Stanelle, B. J. J. M. van den Hurk, F. Vitart, and E. F. Wood. 2010. Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters* 37:L02402, doi:10.1029/2009GL041677.

Krasnopolsky, V. M., M. S. Fox-Rabinovitz, and D. V. Chalikov. 2005. New approach to calculation of atmospheric model physics: accurate and fast neural network emulation of longwave radiation in a climate model. *Monthly Weather Review* 133:1370–1383.

Krishnamurti, T. N., D. K. Oosterhof, A. V. Mehta. 1988. Air–Sea Interaction on the Time Scale of 30 to 50 Days. *Journal of the Atmospheric Sciences* 45:1304-1322.

Krishnamurti, T. N., M. Subramaniam, D. K. Oosterhof, and G. Daughenbaugh. 1990. Predictability of Low-Frequency Modes. *Meteorology and Atmospheric Physics* 44(1–4):63–83.

Krishnamurti, T. N., M. Subramaniam, G. Daughenbaugh, D. Oosterhof, J. Xue. 1992. One-Month Forecasts of Wet and Dry Spells of the Monsoon. *Monthly Weather Review* 120(7):1191–1223.

Krishnamurti, T. N., S.-O. Han, V. Misra. 1995. Prediction of the Dry and Wet Spell of the Australian Monsoon. *International Journal of Climatology* 15(7):753–771.

Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, S. Surendran. 1999. Improved weather and seasonal climate forecasts from multi-model superensemble. *Science* 285:1548–1550.

Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran. 2000. Multi-model superensemble forecasts for weather and seasonal climate. *Journal of Climate* 13:4196–4216.

Kucharski, F., I. Kang, D. Straus, and M. King. 2010. Teleconnections in the atmosphere and oceans. *Bulletin of the American Meteorological Society* 91:381–383.

Kug, J.-S., F.-F. Jin, K. P. Sooraj, and I.-S. Kang. 2008. State-dependent atmospheric noise associated with ENSO. *Geophysical Research Letters* 35:L05701, doi:10.1029/2007GL032017.

Kug, J. S., Y.-G. Ham, M. Kimoto, F.-F. Jin, and I.-S. Kang. 2009. New approach for optimal perturbation method in ensemble climate prediction with empirical singular vector. *Climate Dynamics*, DOI 10.1007/s00382-009-0664-y. Available at *http://www.springerlink.com/content/y131427342247045/fulltext.pdf*, accessed July 19, 2010.

Kumar, A, A. G. Barnston, P. Peng, M.P. Hoerling, and L. Goddard. 2000. Changes in the spread of the variability of the seasonal mean atmospheric states associated with ENSO. *Journal of Climate* 13:3139–3151.

Kumar, K. K., M. Hoerling, and B. Rajagopalan. 2005. Advancing dynamical prediction of Indian monsoon rainfall. *Geophysical Research Letters* 32:L08704, doi:10.1029/2004GL021979.

Lamb, P. J. and S. A. Changnon. 1981. On the "Best" Temperature and Precipitation Normals: The Illinois Situation. *Journal of Applied Meteorology* 20:1383–1390.

Landman, W. A. and L. Goddard. 2002. Statistical recalibration of GCM forecasts over southern Africa using model output statistics. *Journal of Climate* 15:2038–2055.

Large, W. G., J. C. McWilliams, and S. C. Doney. 1994. Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Reviews of Geophysics* 32:363–403.

Lau, K. M. and P. H. Chan. 1986. Aspects of the 40-50 Day Oscillation During the Northern Summer as Inferred from Outgoing Longwave Radiation. *Monthly Weather Review* 114(7):1354–1367.

Lau, K. M. and T. J. Phillips. 1986. Coherent Fluctuations of Extratropical Geopotential Height and Tropical Convection in Intraseasonal Time Scales. *Journal of the Atmospheric Sciences* 43(11):1164–1181.

Lau, K. M., T. Nakazawa, and C. H. Sui. 1991. Observations of Cloud Cluster Hierarchies over the Tropical Western Pacific. *Journal of Geophysical Research-Oceans* 96:3197–3208.

Lau, K. M., and F. C. Chang. 1992. Tropical Intraseasonal Oscillation and Its Prediction by the NMC Operational Model. *Journal of Climate* 5(12):1365–1378.

Lau, K.M. and C. H. Sui. 1997. Mechanisms of Short-Term Sea Surface Temperature Regulation: Observations during TOGA COARE. *Journal of Climate* 10:465-472.

Lau, W. K. M. and D. E. Waliser (Eds). 2005. *Intraseasonal Variability of the Atmosphere-Ocean Climate System.* Springer, Heidelberg, Germany, 474pp.

Lawrence, D. M., and P. J. Webster. 2002. The boreal summer intraseasonal oscillation: Relationship between northward and eastward movement of convection. *Journal of the Atmospheric Sciences* 59(9):1593–1606.

LeDimet, F.-X. and O. Talagrand. 1986. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus A* 38:97–110.

Lee, M. I., I. S. Kang, J. K. Kim, and B. E. Mapes. 2001. Influence of cloud-radiation interaction on simulating tropical intraseasonal oscillation with an atmospheric general circulation model. *Journal of Geophysical Research—Atmospheres* 106(D13):14219–14233.

Lee, Y., G. Wahba, and S. A. Ackerman. 2004. Cloud classification of satellite radiance data by multicategory support vector machines. *Journal of Atmospheric and Oceanic Technology* 21:159–169.

Leith, C. E. 1971. Atmospheric predictability and two-dimensional turbulence. *Journal of the Atmospheric Sciences* 28:145–161.

Lengaigne, M., J.-P. Boulanger, C. Menkes, G. Madec, and P. Delecluse. 2003. The March 1997 Westerly Wind Event and the Onset of the 1997/98 El Niño: Understanding the Role of the Atmospheric Response. *Journal of Climate* 16:3330–3343.

Levitus, S. 1982. Climatological Atlas of the World Ocean. NOAA Professional Paper 13. 190 pp. Silver Spring, MD: Naitonal Oceanic and Atmospheric Administration.

Li, T., B. Wang, C.P. Chang, Y.S. Zhang. 2003. A theory for the Indian Ocean dipole mode. *Journal of the Atmospheric Sciences* 60: 2119–2135.

Liebmann, B., and D.L. Hartmann, 1984. Low frequency variability and tropical extratropical interaction during winter. *Journal of the Atmospheric Sciences* 41:3333-3350.

Liess, S., D. E. Waliser, and S. D. Schubert. 2004. Predictability studies of the intraseasonal oscillation with the ECHAM5 GCM. *Journal of the Atmospheric Sciences* 62:3320–3336.

Lima, C. H. R., U. Lall, T. Jebara, and A. G. Barnston. 2009. Statistical prediction of ENSO from subsurface sea temperature using a nonlinear dimensionality reduction. *Journal of Climate* 22:4501–4519.

Limpasuvan, V., D. W. J. Thompson, and D. L. Hartmann. 2004. The life cycle of the northern hemisphere sudden stratospheric warmings. *Journal of Climate* 17:2584–2596.

Lin, J. L., B. Mapes, M. H. Zhang, and M. Newman. 2004. Stratiform precipitation, vertical heating profiles, and the Madden-Julian oscillation. *Journal of the Atmospheric Sciences* 61(3):296–309.

Livezey, R. E., M. Masutani, A. Leetmaa, H. Rui, M. Ji, and A. Kumar. 1997. Teleconnective Response of the Pacific-North American Region Atmosphere to Large Central Equatorial Pacific SST Anomalies. *Journal of Climate* 10:1787–1820.

Livezey, R. E. and T. M. Smith. 1999. Considerations for the use of the Barnett and Preisendorfer (1987) algorithm for canonical correlation analysis of climate variations. *Journal of Climate* 12:303–305.

Livezey, R. E., K. Y. Vinnikov, M. M. Timofeyeva, R. Tinker, and H. M. van den Dool. 2007. Estimation and extrapolation of climate normals and climatic trends. *Journal of Applied Meteorology and Climatology* 46:1759-1776.

Livezey, R.E. and M. M. Timofeyeva. 2008. The first decade of long-lead U.S. seasonal forecasts: Insights from a skill analysis. *Bulletin of the American Meteorological Society*, 89:843–854.

Lorenc, A. 1986. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society* 112:1177–1194.

Lorenz, E. N. 1956. *Empirical Orthogonal Functions and Statistical Weather Prediction*. Scientific Report 1, Statistical Forecasting Project. Cambridge, MA: Massachusetts Institute of Technology.

Lorenz, E. N. 1965. A study of the predictability of a 28-variable atmospheric model. *Tellus 17:*321–333.

Lorenz, E. N. 1969. Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences* 26:636–646.

Lorenz, E. N. 1975. Climatic Predictability. Pp.133-136 in *The Physical Basis of Climate and Climate Modelling*, GAPR Publication Series No. 16. Geneva, World Meteorological Organization.

Lorenz, E. N. 2006. *Predictability—a problem partly solved*. Pp. 40-58 in *Predictability of*

Weather and Climate, T. Palmer and R. Hagedorn, eds. New York: Cambridge University Press.

Lu, B.-W., L. Pandolfo, and K. Hamilton. 2009. Nonlinear Representation of the Quasi-Biennial Oscillation. *Journal of the Atmospheric Sciences* 66:1886-1904.

Majda, A. J., and S. N. Stechmann. 2009. The skeleton of tropical intraseasonal oscillations. *Proceedings of the National Academy of Sciences* 106:8417–8422, doi:10.1073/pnas.0903367106.

Maloney, E. D. and D. L. Hartmann. 2000. Modulation of hurricane activity in the Gulf of Mexico by the Madden-Julian oscillation. *Science* 287(5460):2002–2004.

Maloney, E. D. 2002. An intraseasonal oscillation composite life cycle in the NCAR CCM3.6 with modified convection. *Journal of Climate* 15(9):964–982.

Maloney, E. D. and A. H. Sobel. 2004. Surface Fluxes and Ocean Coupling in the Tropical Intraseasonal Oscillation. *Journal of Climate* 17:4368–4386.

Manzato, A. 2005. The use of sounding-derived indices for a neural network short-term thunderstorm forecast. *Weather and Forecasting* 20:896–917.

Mason, S. J., L. Goddard, N. E. Graham, E. Yulaeva, L. Sun, and P. A. Arkin. 1999. The IRI Seasonal Climate Prediction System and the 1997/98 El Niño Event. *Bulletin of the American Meteorological Society* 80:1853–1873.

Mason, S. J., and N. E. Graham. 1999. Conditional probabilities, relative operating characteristics, and relative operating levels. *Weather and Forecasting* 14:713–725.

Mason, S. J., and A. P. Weigel. 2009. A generic forecast verification framework for administrative purposes. *Monthly Weather Review* 137:331–349.

Matsuno, T. 1966. Quasi-Geostrophic Motions in the Equatorial Area. *Journal of the Meteorological Society of Japan* 44:25–43.

McPhaden, M. J., A. J. Busalacchi, R. Cheney, J. R. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G. T. Mitchum, P. P. Niiler, J. Picaut, R.W. Reynolds, N. Smith, K. Takeuchi. 1998. The Tropical Ocean-Global Atmosphere (TOGA) observing system: A decade of progress. *Journal of Geophysical Research* 103:14,169–14,240.

McPhaden, M. J., T. Delcroix, K. Hanawa, Y. Kuroda, G. Meyers, J. Picaut, and M. Swenson. 2001. The El Niño/Southern Oscillation (ENSO) Observing System. Pp. 231–246 in *Observing the Ocean in the 21st Century*, C. J. Koblinsky and N. R. Smith, eds. Melbourne, Australia: Australian Bureau of Meteorology.

McPhaden, M. J., G. Meyers, K. Ando,Y. Masumoto, V. S. N. Murty, M. Ravichandran, F. Syamsudin, J. Vialard, L. Yu, and W. Yu. 2009. RAMA: The Research Moored Array for African-Asian-Australian monsoon analysis and prediction. *Bulletin of the American Meteorological Society* 90:459–480.

McPhaden, M.J., K. Ando, B. Bourles, H. P. Freitag , R. Lumpkin, Y. Masumoto, V. S. N. Murty, P. Nobre, M. Ravichandran, J. Vialard, D. Vousden, and W. Yu. 2010. The Global Tropical Moored Buoy Array. In *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society* (Vol. 2), Venice, Italy, 21–25 September 2009, J. Hall, D. E. Harrison, and D. Stammer, eds. ESA Publication WPP-306.

Meehl, G. A. 1997. The South Asian monsoon and the tropospheric biennial oscillation (TBO). *Journal of Climate* 10:1921–1943.

Mercer, A. E. M. B. Richman, H. B. Bluestein, and J. M. Brown. 2008. Statistical modeling of downslope windstorms in Boulder, Colorado. *Weather and Forecasting* 23:1176–1194.

Mercer, A. E., C. M. Shafer, C. A. Doswell, L. M. Leslie, and M. B. Richman. 2009. Objective classification of tornadic and non-tornadic severe weather outbreaks. *Monthly Weather Review* 137:12.

Michaelson, J. 1987. Cross-validation in statistical climate forecast models. *Journal of Climate and Applied Meteorology* 26:1589–1600.

Mielke, P.W., K. J. Berry and G. W. Brier. 1981. Application of multi-response permutation procedures for examining changes in monthly mean sea-level pressure patterns. *Monthly Weather Review* 109:120–126.

Miyakoda, K., J. Smagorinsky, R. F. Strickler, G. D. Hembree. 1969. Experimental Extended Predictions With A Nine-Level Hemispheric Model. *Monthly Weather Review* 97:1-76.

Miyakoda, K., G. D. Hembree, and R. F. Strickler. 1979. Cumulative results of extended forecast experiments II: Model performance for summer cases. *Monthly Weather Review* 107: 395-420.

Miyakoda, K., T. Gordon, R. Caverly, W. Stern, J. Sirutis, W. Bourke. 1983. Simulation of a Blocking Event in January 1977. *Monthly Weather Review* 111:846-869.

Mo, K. C. 2001. Adaptive filtering and prediction of intraseasonal oscillations. *Monthly Weather Review* 129(4):802–817.

Molteni, F., R. Buizza, T. N. Palmer and T. Petroliagis. 1996. The ECMWF ensemble prediction system: methodology and validation. *Quarterly Journal of the Royal Meteorological Society* 122:73–119.

Monahan, A. H. and A. Dai. 2004. The Spatial and Temporal Structure of ENSO Nonlinearity. *Journal of Climate* 17:3026-3036.

Moncrieff, M. W. 2004. Analytic representation of the large-scale organization of tropical convection. *Journal of the Atmospheric Sciences* 61(13):1521–1538.

Moncrieff, M. W. and C. Liu. 2006. Representing convective organization in prediction models by a hybrid strategy. *Journal of the Atmospheric Sciences* 63:3404–3420.

Moncrieff, M. W., M. Shapiro, J. Slingo, and F. Molteni. 2007. Collaborative research at the intersection of weather and climate. *WMO Bulletin* 56:204–211.

Montroy, D. L., M. B. Richman, and P. J. Lamb. 1998. Observed nonlinearities of monthly teleconnections between tropical Pacific sea surface temperature anomalies and central and eastern North American precipitation. *Journal of Climate* 11:1812–1835.

Moore, D. W. 1968. *Planetary gravity waves in an equatorial ocean*, Ph.D. Thesis, Harvard University 207pp.

Moura, A. D., and S. Hastenrath. 2004. Climate Prediction for Brazil's Nordeste: Performance of Empirical and Numerical Modeling Methods. *Journal of Climate* 17:2667–2672.

Murphy, A. H. and B. G. Brown. 1984. A comparative evaluation of objective and subjective weather forecasts in the United States. *Journal of Forecasting* 3:369–393, DOI: 10.1002/for.3980030402

Murphy, A. H. 1991. Forecast verification: its complexity and dimensionality. *Monthly Weather Review* 119:1590–1601.

Murphy, A. H. 1993. What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and Forecasting* 8:281–293.

Murphy, A. H. 1995. The coefficients of correlation and determination as measures of performance in forecast verification. *Weather and Forecasting* 10:681–688.

Myers, D. S. and D. E. Waliser. 2003. Three-dimensional water vapor and cloud variations associated with the Madden-Julian Oscillation during the Northern Hemisphere winter. *Journal of Climate* 16:929–950.

Nakazawa, T. 1988. Tropical Super Clusters within Intraseasonal Variations over the Western Pacific. *Journal of the Meteorological Society of Japan* 66(6):823–839.

Neelin, J., D. Battisti, A. Hirst, F.-F. Jin, Y. Wakata, T. Yamagata, and S. Zebiak. 1998. ENSO theory. *Journal of Geophysical Research* 103(C7):14261–14290.

Newman, M. and P. D. Sardeshmukh. 1995. A caveat concerning singular value decomposition. *Journal of Climate* 8:352–360.

Njoku, E.G., T.J. Jackson, V. Lakshmi, T.K. Chan, and S.V. Nghiem. 2003. Soil moisture retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing* 41(2): 215 – 229.

NRC (National Research Council). 1996. *Learning to Predict Climate Variations Associated with El Niño and the Southern Oscillation: Accomplishments and Legacies of the TOGA Program*. Washington, D.C.: National Academies Press.

NRC. 1998. *A Scientific Strategy for U.S. Participation in the GOALS (Global Ocean-Atmosphere-Land System) Component of the CLIVAR (Climate Variability and Predictability) Programme.* Washington, D.C.: National Academies Press.

NRC. 2004. *Review of the U.S. CLIVAR Project Office.* Washington, D.C.: National Academies Press.

NRC. 2006. *Toward an Integrated Arctic Observing Network*. Washington, D.C.: National Academies Press.

Oglesby, R. J. and D. J. Erickson III. 1989. Soil moisture and the persistence of North American drought. *Journal of Climate* 2:1362–1380.

Ohring, G. 1972. Application of stepwise multiple regression techniques to inversion of Nimbus "IRIS" observations. *Monthly Weather Review* 100:336–344.

O'Lenic, E. A., D. A. Unger, M. S. Halpert, and K. S. Pelman. 2008. Developments in operational long-range climate prediction at CPC. *Weather and Forecasting* 23:496–515.

Oreskes, N, K. Shrader-Frechette, and K. Beltz. 1994. Verification, validation and confirmation of numerical models in earth sciences. *Science* 263:641–646.

Orrell, D., L.Smith, J. Barkmeijer, and T. N. Palmer. 2001. Model error in weather forecasting. *Nonlinear Processes in Geophysics* 8:357–371.

Owe, M., R. de Jeu, and J. Walker. 2001. A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. *IEEE Transactions on Geoscience and Remote Sensing* 39:1643–1654.

Palmer, T. N. 2002. The economic value of ensemble forecasts as a tool for risk assessment: from days to decades. *Quarterly Journal of the Royal Meteorological Society* 128:747–774.

Palmer, T. N., A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, P. Délécluse, M. Déqué, E. Díez, F. J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J. F. Guérémy, R. Hagedorn, M. Hoshen, N. Keenlyside, M. Latif, A. Lazar, E. Maisonnave, V. Marletto, A. P. Morse, B. Orfila, P. Rogel, J. M. Terres, and M. C. Thomson. 2004. Development of a European Multi-model Ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society* 85:853–872.

Parrish, D. F., and J. C. Derber. 1992. The National Meteorological Center's Spectral Statistical-interpolation Analysis System. *Monthly Weather Review* 120:1747-1763.

Pegion, K. and B. P. Kirtman. 2008. The impact of air-sea interactions on the simulation of tropical intraseasonal variability. *Journal of Climate* 21:616–6635.

Peña, M. and H. van den Dool, 2008: Consolidation of multimodel forecasts by ridge regression: Application to Pacific sea surface temperature. *Journal of Climate* 21:6521–6538.

Philander, S. G. H., T. Yamagata, and R. C. Pacanowski. 1984. Unstable Air-Sea Interactions in the Tropics. *Journal of the Atmospheric Sciences* 41(4):604–613.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons. 1999. The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quarterly Journal of the Royal Meteorological Society* 126:1143–1170.

Rajagopalan, B., U. Lall, and S. E. Zebiak. 2002. Categorical climate forecasts through Regularization and Optimal Combination of Multiple GCM Ensembles. *Monthly Weather Review* 13:1792–1811.

Ramachandran, S., V. Ramaswamy, G. L. Stenchikov, and A. Robock. 2000. Radiative impact of the Mount Pinatubo volcanic eruption: Lower stratospheric response. *Journal of Geophysical Research* 105:24 409–24 429.

Rasmusson, E. M. and T. H. Carpenter. 1982. Variations in Tropical Sea Surface Temperature and Surface Wind Fields Associated with the Southern Oscillation/El Niño. *Monthly. Weather Review* 110:354-384.

Reichle, R. H., R. D. Koster, J. Dong, and A. A. Berg. 2004. Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation. *Journal of Hydrometeorology* 5:430–442.

Reichle, R. H., R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe. 2007. Comparison and assimilation of global soil moisture retrievals from AMSR-E and SMMR. *Journal of Geophysical Research* 112:D09108.

Reichler, T. and J. Kim. 2008. How Well Do Coupled Models Simulate Today's Climate? *Bulletin of the American Meteorological Society* 89:303–311.

Reynolds, C., R. Gelaro, and T. Murphree. 1996. Observed and simulated Northern Hemisphere intraseasonal circulation anomalies and the effluence of model bias. *Journal of Climate* 124:1100–1118.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang. 2002. An improved in situ and satellite SST analysis for climate change. *Journal of Climate* 15:1609–1625.

Richman, M. B. 1986. Rotation of principal components. *International Journal of Climatology* 6:293–335.

Richman, M. B. and I. Adrianto. 2010. Classification and regionalization through kernel principal component analysis. *Physics and Chemistry of the Earth* 35:316-328, doi:10.1016/j.pce.2010.02.001.

Rind, D. 1982. The influence of ground moisture conditions in North America on summer climate as modeled in the GISS GCM. *Monthly Weather Review* 110:1487–1494.

Robertson, A. W., S. E. Zebiak, U. Lall, and L. Goddard. 2004. Optimal combination of multiple atmospheric GCM ensembles for seasonal prediction. *Monthly Weather Review* 132:2732–2744, doi: 10.1175/MWR2818.1.

Robertson, A. W., V. Moron, and Y. Swarinoto. 2009. Seasonal predictability of daily rainfall statistics over Indramayu district, Indonesia. *International Journal of Climatology* 29:1449–1462, doi: 10.1002/joc.1816

Robock, A. and J. Mao. 1995. The volcanic signal in surface temperature observations. *Journal of Climate* 8:1086–1103.

Robock, A., L. Oman, G. L. Stenchikov, O. B. Toon, C. Bardeen, and R. P. Turco. 2007. Climatic consequences of regional nuclear conflicts. *Atmospheric Chemistry and Physics* 7:2003–2012.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll. 2004. The global land data assimilation system. *Bulletin of the American Meteorological Society* 85:381–394.

Rogers, J.C. 1990. Patterns of low-frequency monthly sea-level pressure variability (1899-1986) and associated wave cyclone frequencies. *Journal of Climate* 3(12):1364–1379

Rosati, A., K. Miyakoda, and R. Gudgel. 1997. The impact of ocean initial conditions on ENSO forecasting with a coupled model. *Monthly Weather Review* 125:754–772.

Rotstayn, L. D., M. A. Collier, M. R. Dix, Y. Feng, H. B. Gordon, S. P. O'Farrell, I. N. Smith, and J. Syktus. 2010. Improved simulation of Australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment. *International Journal of Climatology* 30:1067–1088.

Rui, H. and B. Wang. 1990. Development characteristics and dynamic structure of tropical intraseasonal convection anomalies. *Journal of the Atmospheric Sciences* 47(3):357–379.

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Pena, S. Lord, G. White, W. Ebisuzaki, P. Peng, and P. Xie. 2006. The NCEP Climate Forecast System. *Journal of Climate* 19: 3483–3517.

Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y.-T. Hou, H.-y. Chuang, H.-M. H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H. Wei, R. Yang, S. Lord, H. van den Dool, A. Kumar, W. Wang, C. Long, M. Chelliah, Y. Xue, B. Huang, J.-K. Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C.-Z. Zou, Q. Liu, Y. Chen, Y. Han, L. Cucurull, R. W. Reynolds, G. Rutledge, M. Goldberg. 2010. The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society.* Early Online Release, April 23, 2010. doi: 10.1175/2010BAMS3001.1

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata. 1999. A dipole mode in the tropical Indian Ocean. *Nature* 401:360–363.

Sarachik, E. S., and M. A. Cane. 2010. *The El Niño-Southern Oscillation Phenomenon*. New York: Cambridge University Press.

Saravanan, R., G. Danabasoglu, S. C. Doney, and J. C. McWilliams. 2000. Decadal variability and predictability in the midlatitude ocean-atmosphere system. *Journal of Climate* 13:1073–1097.

Sardeshmukh, P. D. and B. J. Hoskins. 1988. The Generation of Global Rotational Flow by Steady Idealized Tropical Divergence. *Journal of the Atmospheric Sciences* 45(7):1228–1251.

Schlosser, A. and P. C. D. Milly. 2002. A model-based investigation of soil-moisture predictability and associated climate predictability. *Journal of Hydrometeorology* 3:483–501.

Schneider, E. K., B. Huang, Z. Zhu, D. G. DeWitt, J. L. Kinter III, B. P. Kirtman, and J. Shukla. 1999. Ocean data assimilation, initialization and prediction with a coupled GCM. *Monthly Weather Review* 127:1187–1207.

Schubert, S., R. Dole, H. van den Dool, M. Suarez, and D.Walliser. 2002. *Proceedings from a Workshop on Prospects for Improved forecasts of Weather and Short-term Climate Variability on Subseasonal (2 week to 2 month) Time Scales*. April 16–18, 2002, Mitchellville, MD. NASA/TM 2002-104606, vol. 23, 171 pp.

Seneviratne, S., R. D. Koster, Z. Guo, P. A. Dirmeyer, E. Kowalczyk, D. Lawrence, P.Liu, C.-H. Lu, D. Mocko, K. W. Oleson, and D. Verseghy. 2006a. Soil moisture memory in AGCM simulations, Analysis of Global Land-Atmosphere Coupling Experiment (GLACE) data. *Journal of Hydrometeorology* 7:1090–1112.

Seneviratne, S. I., D. Luthi, M. Litschi, and C. Schar. 2006b. Land-atmosphere coupling and climate change in Europe. *Nature* 443:205–209.

Seneviratne, S.I., T. Corti, E.L. Davin, M. Hirschi, E.B. Jaeger, I. Lehner, B. Orlowsky, and A.J. Teuling. 2010. Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews* 99, 3-4:125-161, doi:10.1016/j.earscirev.2010.02.004

Servain, J., A. J. Busalacchi, M. J. McPhaden, A. D. Moura, G. Reverdin, M. Vianna, and S. E. Zebiak. 1998. A Pilot Research Moored Array in the Tropical Atlantic (PIRATA). *Bulletin of the American Meteorological Society* 79(10):2019–2031.

Shukla, J. 1985. Predictability. *Advances in Geophysics* 28B:87–122.

Shukla, J. and Y. Mintz. 1982. Influence of land-surface evapotranspiration on the earth's climate. *Science* 215:1498–1501.

Shukla, J. 1998. Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science* 282:728–731.

Shutts, G. 2005. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society* 131:3079–3102.

Silverman, D. and J. A. Dracup. 2000. Artificial neural networks and long-range precipitation prediction in California. *Journal of Applied Meteorology* 39:57–66.

Simmons, A. J. and A. Hollingsworth. 2002. Some aspects of the improvement in skill of numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society* 128:647–677.

Slingo, J. M., K. R. Sperber, J. S. Boyle, J.-P. Ceron, M. Dix, B. Dugas, W. Ebisuzaki, J. Fyfe, D. Gregory, J.-F. Gueremy, J. Hack, A. Harzallah, P. Inness, A. Kitoh, W. K.-M. Lau, B. McAvaney, R. Madden, A. Matthews, T. N. Palmer, C.-K. Park, D. Randall, and N. Renno. 1996. Intraseasonal oscillations in 15 atmospheric general circulation models: Results from an AMIP diagnostic subproject. *Climate Dynamics* 12(5):325–357.

Slingo, J. M., P. M. Inness, and K. R. Sperber. 2005. Modeling. Pp. 361–388 in *Intraseasonal Variability of the Atmosphere-Ocean Climate System*, W. K.-M. Lau and D. E. Waliser, eds. Heidelberg, Germany: Springer.

Sobel, A. H., E. D. Maloney, G. Bellon, and D. M. Frierson. 2008. The role of surface fluxes in tropical intraseasonal oscillations. *Nature Geoscience* 1:653–657.

Song, Q., D. B. Chleton, S. K. Esbensen, N. Thum, and L. W. O'Neill. 2009. Coupling between sea surface temperature and low-level winds in mesoscale numerical models. *Journal of Climate* 22:146–164.

Sperber, K. R., J. M. Slingo, P. M. Inness, and W. K.-M. Lau. 1997. On the maintenance and initiation of the intraseasonal oscillation in the NCEP/NCAR reanalysis and in the GLA and UKMO AMIP simulations. *Climate Dynamics* 13(11):769–795.

Sperber, K. R. and D. Waliser. 2008. New Approaches to Understanding, Simulating, and Forecasting the Madden-Julian Oscillation. *Bulletin of the American Meteorological Society* doi: 10.1175/2008BAMS2700.1.

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C. N. Hill, and J. Marshall. 2002. Global ocean circulation during 1992-1997, estimated from ocean observations and a general circulation model. *Journal of Geophysical Research* 107(C9):3118.

Stewart, R. H. 2005. *Introduction to Physical Oceanography*. Self-published. Available at *http://oceanworld.tamu.edu/ocean410/ocng410_text_book.html*, accesed July 23, 2010.

Stockdale, T. N., D. L. T. Anderson, J. O. S. Alves, and M. A. Balmaseda. 1998. Global seasonal rainfall forecasts using a coupled ocean–atmosphere model. *Nature* 392:370–373.

Stockdale, T. N. D. L. T. Anderson, M. A. Balmaseda, F. Doblas-Reyes, L. Ferranti, K. Mogensen, T. N. Palmer, F. Molteni, and F. Vitart. 2010. ECMWF Seasonal Forecast System 3 and its prediction of sea surface temperature. *Climate Dynamics* (in review).

Suarez, M. J. and P.S. Schopf. 1988. A delayed action oscillator for ENSO. *Journal of the Atmospheric Sciences* 45:3283-3287.

Sun, L., D. F. Moncunill, H. Li, A. D. Moura, and F. A. S. Filho. 2005. Climate Downscaling over Nordeste, Brazil, using the NCEP RSM97. *Journal of Climate* 18:551–567.

Tang, B. W. W. Hsieh, A. H. Monahan, and F. T. Tangang. 2000. Skill comparisons between neural networks and canonical correlation analysis in predicting the equatorial Pacific sea surface temperatures. *Journal of Climate* 13:287–293.

Tangang, F. T., B. Tang A. H. Monahan, and W. W. Hsieh. 1998. Forecasting ENSO events: a neural network-extended EOF approach. *Journal of Climate* 11:29–41.

Thompson, D. W. J. and J. M. Wallace. 2000. Annular modes in the extratropical circulation. Part I: Month-to-month variability. *Journal of Climate* 13:1000–1016.

Thompson, D. W. J. and J. M. Wallace. 2001. Regional Climate Impacts of the Northern Hemisphere Annular Mode. *Science* 293(5527):85–89.

Thompson, D. W. J, S. Lee, and M. P. Baldwin. 2002. Atmospheric processes governing the Northern Hemisphere annular mode/North Atlantic Oscillation. Pp. 81–112 in *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, J. W. Hurrell, Y. Kushnir, G. Ottersen, M. Visbeck, eds. Washington, D.C.: American Geophysical Union.

Tian, B., Y. L. Yung, D. E. Waliser, T. Tyranowski, L. Kuai, E. J. Fetzer, and F. W. Irion. 2007. Intraseasonal variations of the tropical total ozone and their connection to the MJO. *Geophysical Research Letters* 34:L08704, doi: 10.1029/2007GL029471.

Tippett, M. K., L. Goddard, and A. G. Barnston. 2005. Statistical-dynamical seasonal forecasts of Central Southwest Asia winter precipitation. *Journal of Climate* 18:1831–1843.

Tippett, M. K., T. DelSole, S. J. Mason, and A. G. Barnston. 2008. Regression-based methods for finding coupled patterns. *Journal of Climate* 21:4384–4398.

Tompkins, A. M., K. Gierens, and G. Rädel. 2007. Ice supersaturation in the ECMWF integrated forecasting system. *Quarterly Journal of the Royal Meteorological Society* 133:53–56.

Toth, Z. and E. Kalnay. 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society* 74:2330–2371.

Trafalis, T. B., B. Santosa, and M. B. Richman. 2005. Learning networks in rainfall estimation. *Computational Management Science* 2:229–251 doi: 10.1007/s10287-005-0026-0.

Trenberth, K. E., P. D. Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. Parker, F. Rahimzadeh, J.A. Renwick, M. Rusticucci, B. Soden, and P. Zhai. 2007. Observations: Surface and Atmospheric Climate Change. In: *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, eds. Cambridge, UK: Cambridge University Press.

Tribbia, J. J. and D. P. Baumhefner. 2004. Scale interactions and atmospheric predictability: An updated perspective. *Monthly Weather Review* 132:703–713.

Troccoli, A. and P. Kållberg. 2004. *Precipitation correction in the ERA-40 reanalyses*. ERA-40 Project Report Series No. 13, ECMWF.

van den Dool, H. M. 1994. Searching for analogues, how long must we wait? *Tellus* 46A:314–324.

van den Dool, H. M. and S. Saha. 2002. *Analysis of Propagating Modes in the Tropics in Short AMIP runs*. AMIPII workshop, World Meteorological Organization, Toulouse, France, November 12-15, 2002.

van den Dool, H. M., J. Huang and Y. Fan. 2003. Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981-2001. *Journal of Geophysical Research* 108:8617, doi:10.1029/2002JD003114.

Vecchi, G. A. and D. E. Harrison, 2000. Tropical Pacific Sea Surface Temperature Anomalies, El Niño, and Equatorial Westerly Wind Events. *Journal of Climate* 13:1814-1830.

Vecchi, G. A. and N. A. Bond. 2004. The Madden-Julian Oscillation (MJO) and northern high latitude wintertime surface air temperatures. *Geophysical Research Letters* 31(4).

Vecchi, G. A., A. T. Wittenberg, and A. Rosati. 2006. Reassessing the role of stochastic forcing in the 1997–1998 El Niño. *Geophysical Research Letters* 33:L01706, doi:10.1029/2005GL024738, 2006

Vinnikov, K. Y. and I. B. Yeserkepova. 1991. Soil moisture, empirical data and model results. *Journal of Climate* 4:66–79.

Vinnikov, K. Y., A. Robock, N. A. Speranskaya, and A. Schlosser. 1996. Scales of temporal and spatial variability of midlatitude soil moisture. *Journal of Geophysical Research* 101:7163–7174.

Vintzileos, A. and H.-L. Pan. 2007. On the importance of horizontal resolution and initial conditions to forecasting tropical intraseasonal oscillations: the maritime continent prediction barrier. NOAA CTB–COLA Joint Seminar, Sep. 19, 2007. Available at *http://www.nws.noaa.gov/ost/climate/STIP/CTB-COLA/Augustin_091907.pdf*, accessed July 26, 2010.

Vitart, F., A. Leroy, and M. C. Wheeler. 2010. A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the Southern Hemisphere. *Monthly Weather Review* (submitted).

Viterbo, P., and A. K. Betts. 1999. Impact of the ECMWF reanalysis soil water on forecasts of the July 1993 Mississippi flood. *Journal of Geophysical Research* 104:19361–19366.

Von Storch, H. and J. Xu. 1990. Principal oscillation pattern analysis of the 30- to 60-day oscillation in the tropical troposphere. *Climate Dynamics* 4:175–190.

Waliser, D. E., C. Jones, J.-K. Schemm, and N. E. Graham. 1999a. A statistical extended-range tropical forecast model based on the slow evolution of the Madden-Julian oscillation. *Journal of Climate* 12(7):1918–1939.

Waliser, D. E., K. M. Lau, and J.-H. Kim. 1999b. The influence of coupled sea surface temperatures on the Madden-Julian oscillation: A model perturbation experiment. *Journal of the Atmospheric Sciences* 56(3):333–358.

Waliser, D. E., K. Jin, I.-S. Kang, W. F. Stern, S. D. Schubert, M.L.C. Wu, K.-M. Lau, M.-I. Lee, V. Krishnamurthy, A. Kitoh, G. A. Meehl, V. Y. Galin, V. Satyan, S. K. Mandke, G. Wu, Y. Liu, and C.-K. Park. 2003a. AGCM simulations of intraseasonal variability associated with the Asian summer monsoon. *Climate Dynamics* 21(5-6):423–446.

Waliser, D. E., K. M. Lau, W. Stern, C. Jones. 2003b. Potential Predictability of the Madden-Julian Oscillation, *Bulletin of the American Meteorological Society*, *84*(1), 33-50.

Waliser, D. E., S. Schubert, A. Kumar, K. Weickmann, and R. Dole. 2003c. Proceedings from a workshop on "Modeling, Simulation and Forecasting of Subseasonal Variability." 4–5 June 2003, College Park, MD. NASA/TM 2003-104606, vol. 25, 67.

Waliser, D. E., W. Stern, S. Schubert, and K. M. Lau. 2003d. Dynamic Predictability of Intraseasonal Variability Associated with the Asian Summer Monsoon. *Quarterly Journal of the Royal Meteorological Society* 129:2897–2925.

Waliser, D. E. 2005. Predictability and Forecasting. Pp. 389–424 in *Intraseasonal Variability of the Atmosphere-Ocean Climate System*, W. K. M. Lau and D. E. Waliser, eds. Heidelberg, Germany: Springer.

Waliser, D. E., K. Weickmann, R. Dole, S. Schubert, O. Alves, C. Jones, M. Newman, H.-L. Pan, A. Roubicek, S. Saha, C. Smith, H. van den Dool, F. Vitart, M. Wheeler, and J. Whitaker. 2005. The Experimental MJO Prediction Project. *Bulletin of the American Meteorological Society* 87:425–431.

Waliser, D. E. 2006. Predictability of Tropical Intraseasaonal Variability. Pp. 275–305 in *Predictability of Weather and Climate*, T. N. Palmer and R. Hagedorn, eds. New York: Cambridge University Press.

Waliser, D. E. and M. Moncrieff. 2008. *The Year of Tropical Convection (YOTC) Science Plan: A joint WCRP-WWRP/THORPEX International Initiative*. WMO/TD No. 1452, WCRP - 130, WWRP/THORPEX - No 9. Geneva, Switzerland: World Meteorological Organization.

Waliser, D. E., J. L. F. Li, C. P. Woods, R. T. Austin, J. Bacmeister, J. Chern, A. Del Genio, J. H. Jiang, Z. M. Kuang, H. Meng, P. Minnis, S. Platnick, W. B. Rossow, G. L. Stephens, S. Sun-Mack, W. K. Tao, A. M. Tompkins, D. G. Vane, C. Walker, and D. Wu. 2009. Cloud ice: A climate model challenge with signs and expectations of progress. *Journal of Geophysical Research—Atmospheres* 114: D00A21, doi:10.1029/2008JD010015.

Walker, G.T ., 1923: Correlations in seasonal variations of weather, VIII: a preliminary study of world weather. Pp. 75–131 in *Memoirs of the India Meteorological Department*, Vol. 24. Calcutta: Government of India Press.

Wallace, J. M. and D. S. Gutzler. 1981. Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Monthly Weather Review* 109:784–812.

Wang, B. and H. Rui. 1990. Synoptic Climatology of Transient Tropical Intraseasonal Convection Anomalies: 1975–1985. *Meteorology and Atmospheric Physics* 44(1–4):43–61.

Wang, B., R. Wu, and X. Fu. 2000. Pacific-East Asia teleconnection: How does ENSO affect East Asian climate? *Journal of Climate* 13:1517–1536.

Wang, B., R.G. Wu, and T. Li. 2003. Atmosphere-warm Ocean interaction and its impacts on Asia-Australian Monsoon variation. *Journal of Climate* 15:1195–1211.

Wang, B., I. S. Kang, and J. Y. Lee. 2004. Ensemble simulations of Asian-Australian monsoon variability by 11 AGCMs. *Journal of Climate* 17:803–818.

Wang, B., Q. Ding, X. Fu, I.-S. Kang, K. Jin, J. Shukla, and F. Doblas-Reyes. 2005. Fundamental challenges in simulation and prediction of summer monsoon rainfall. *Geophysical Research Letters* 32(15):L15711, doi:10.1029/2005GL022734 12.

Wang, B., J.-Y. Lee, I.-S. Kang, J. Shukla, J.-S. Kug, A. Kumar, J. Schemm, J.-J. Luo, T.Yamagata, and C.-K. Park. 2008. How accurately do coupled climate models predict the Asian-Australian monsoon interannual variability? *Climate Dynamics* 30:605–619, doi:10.1007/s00382-007-0310-5.

Wang, B., J.-Y. Lee, I.-S. Kang, J. Shukla, C.-K. Park, A. Kumar, J. Schemm, S. Cocke, J.-S. Kug, J.-J. Luo, T. Zhou, B. Wang, X. Fu, W.-T. Yun, O. Alves, E. K. Jin, J. Kinter, B. Kirtman, T. Krishnamurti, N. C. Lau, W. Lau, P. Liu, P. Pegion, T. Rosati, S. Schubert, W. Stern, M. Suarez and T. Yamagata. 2009. Advance and Prospect of Seasonal Prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980-2004). *Climate Dynamics* 33(1):93-117. DOI: 10.1007/s00382-008-0460-0

Watson, J. S. and S. J. Colucci. 2002. Evaluation of ensemble predictions of blocking in the NCEP global spectal model. *Monthly Weather Review* 130:3008–3021.

WCRP (World Climate Research Programme). 2010. *The CLIVAR Handbook.* Southampton, U.K.: World Climate Research Programme. Available at *http://www.clivar.org/publications/other_pubs/handbook.pdf*, accessed July 26, 2010.

Weaver, A. T., J. Vialard, and D. L. T. Anderson. 2003. Three- and Four-Dimensional variational assimilation with a General Circulation Model of the Tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks. *Monthly Weather Review* 131:1360–1378.

Webster, P. J., A. M. Moore, J. P. Loschnigg, and R. R. Leben. 1999. Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997-98. *Nature* 401:356–360.

Webster, P. J. and C. Hoyas. 2004. Forecasting monsoon rainfall and river discharge variability on 20-25 day time scales. *Bulletin of the American Meteorological Society* 85(11):1745–1765.

Weickmann, K. M. 1983. Intraseasonal Circulation and Outgoing Longwave Radiation Modes During Northern Hemisphere Winter. *Monthly Weather Review* 111(9):1838–1858.

Weickmann, K. M., G. R. Lussky, and J. E. Kutzbach. 1985. Intraseasonal (30-60 Day) Fluctuations of Outgoing Longwave Radiation and 250-Mb Stream-Function During Northern Winter. *Monthly Weather Review* 113(6):941–961.

Weisheimer, A., F. J. Doblas-Reyes, T. N. Palmer, A. Alessandri, A. Arribas, M. Déqué, N. Keenlyside, M. MacVean, A. Navarra, and P. Rogel. 2009. ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions—Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophysical Research Letters* 36:L21711, doi:10.1029/2009GL040896.

Wheeler, M. and G. N. Kiladis. 1999. Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber-frequency domain. *Journal of the Atmospheric Sciences* 56(3):374–399.

Wheeler, M. and K. M. Weickmann. 2001. Real-time monitoring and prediction of modes of coherent synoptic to intraseasonal tropical variability. *Monthly Weather Review* 129:2677–2694.

Wheeler, M. and H. Hendon. 2004. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Monthly Weather Review* 132(8):1917–1932.

Wilks, D. S. 2006. *Statistical Methods in the Atmospheric Sciences*. 2nd Edition. International Geophysics Series, Vol. 59. San Diego, CA: Academic Press.

Winkler, C. R., M. Newman, and P. D. Sardeshmukh. 2001. A linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill. *Journal of Climate* 14:4474–4494.

Wittenberg, A. T. 2004. Extended wind stress analyses for ENSO. *Journal of Climate* 17:2526-2540.

WMO SVS-LRF. 2002 World Meteorological Organization (WMO) Standardised Verification System (SVS) for Long-Range Forecasts (LRF). New Attachment II-9 to the Manual on the GDPS (WMO-No. 485), Volume I, WMO, Geneva, Switzerland.

Wolff, J.-O., E. Maier-Reimer, and S. Legutke. 1997. *The Hamburg Ocean Primitive Equation Model: HOPE*. DKRZ Technical Report No. 13. Hamburg, Germany: Modellberatungsgruppe, DKRZ.

Wong, S., and A.E. Dessler. 2007. Regulation of H2O and CO in Tropical Tropopause Layer by the Madden-Julian Oscillation. *Journal of Geophysical Research* 112, D14305, DOI: 10.1029/2006JD007940.

Woolnough, S. J., J. M. Slingo, and B. J. Hoskins. 2000. The relationship between convection and sea surface temperature on intraseasonal timescales. *Journal of Climate* 13:2086–2104.

Woolnough, S. J., F. Vitart, and M. A. Balmaseda. 2007. The role of the ocean in the Madden-Julian Oscillation: Implications for MJO prediction. *Quarterly Journal of the Royal Meteorological Society* 133(622):117–128.

Wu, P. and C. Gordon. 2002. Oceanic Influence on North Atlantic Climate Variability. *Journal of Climate* 15:1911–1925.

Wu, R. and B. P. Kirtman. 2005. Roles of Indian-Pacific Ocean air-sea coupling in tropical atmospheric variability. *Climate Dynamics* 25:155–170.

Yang, F., A. Kumar, and K.-M. Lau. 2004. Potential predictability of U.S. summer climate with "Perfect" soil moisture. *Journal of Hydrometeorology* 5:883–895.

Yasunari, T. 1979. Cloudiness fluctuations associated with the Northern Hemisphere summer monsoon. *Journal of the Meteorological Society of Japan* 57:227–242.

Yuan, X. J. 2004. ENSO-related impacts on Antarctic sea ice: a synthesis of phenomenon and mechanisms. *Antarctic Science* 16(4):415-425.

Zebiak, S. E. and M. A. Cane. 1987. A model El Niño Southern Oscillation. *Monthly Weather Review* 115:2262-2278.

Zeng, N., J. D. Neelin, K.-M. Lau, and C. J. Tucker. 1999. Enhancement of interdecadal climate variability in the Sahel by vegetation interaction. *Science* 286:1537-1540.

Zhang, C. 1997. Intraseasonal variability of the upper-ocean temperature structure observed at 0 and 165E. *Journal of Climate* 10:3077-3092.

Zhang, C., H. H. Hendon, W. S. Kessler, and A. Rosati. 2001. A workshop on the MJO and ENSO. *Bulletin of the American Meteolological Society* 82:971–976.

Zhang, C. 2005. The Madden Julian Oscillation. *Reviews of Geophysics* 43:RG2003, doi:10.1029/2004RG000158.

Zhang, C., M. Dong, H. H. Hendon, E. D. Maloney, A. Marshall, K. R. Sperber, and W. Wang. 2006. Simulations of the Madden-Julian Oscillation in Four Pairs of Coupled and Uncoupled Global Models. *Climate Dynamics* 27(6):573-592, doi:10.1007/s00382-006-0148-2.

Zhang, H.M., R.W. Reynolds, R. Lumpkin, R. Molinari, K. Arzayus, M. Johnson, and T.M. Smith. 2009. An Integrated Global Observing System For Sea Surface Temperature Using Satellites and in Situ Data: Research to Operations. *Bulletin of the American Meteolological Society* 90:31-38.

Zhang, R.-H., A. J. Busalacchi, and D. G. DeWitt. 2008. The roles of atmospheric stochastic forcing (SF) and oceanic entrainment temperature (Te) in decadal modulation of ENSO. *Journal of Climate* 21(4):674–704.

Zhang, Y., J. M. Wallace, and D. S. Battisti. 1997. ENSO-like Interdecadal Variability: 1900–93. *Journal of Climate* 10:1004–1020.

Zheng, Y., D. E. Waliser, W. F. Stern, and C. Jones. 2004. The role of coupled sea surface temperatures in the simulation of the tropical intraseasonal oscillation. *Journal of Climate* 17:4109–4134.

Appendix A
Background Information on Statistical Techniques

Throughout this report, various statistical techniques are mentioned. In some cases, these techniques are used in model validation efforts, as a way of measuring the performance of predictions or forecasts. Many of the forecast "metrics" involve some sort of statistical algorithm. If a forecast system is deemed to be of low quality, the statistical techniques may provide a first step in identifying opportunities for improvements to the forecast models. Alternatively, if bias is detected in the predictions or forecast, statistical techniques can also be used to devise methods of bias-correction. In other cases, statistical techniques have been used to identify and characterize "patterns of variability" within the climate system, and serve as the foundation for a forecast system.

The following sections provide some background material on several statistical techniques that are frequently mentioned in the report. First, a table listing 11 commonly used statistical techniques is provided, listing some of the advantages and disadvantages in their application to model validation efforts and forecasting. Then, five specific sets of techniques (correlation, multiple regression, composites, eigentechniques, and kernel methods) are discussed in more detail.

Table A.1. Commonly used statistical techniques and their advantages and disadvantages.

Technique	Advantages	Disadvantages
Pearson's Correlation	Well understood. Intuitive scale.	Linear, sensitive to outliers, not designed to find causal relationships.
Spearman's Correlation	Well understood. Intuitive scale. Resistant to outliers.	Linear in the ranks.
EOF/PCA	Well understood. Efficient compression of large datasets.	Linear. Sensitive to sampling errors. In most applications, requires the estimation of the dimensionality of the signal. If modes identification is desired, may require post processing with additional linear transformation.
Nonlinear (Complex) EOF/PCA	Can result in very efficient compression of data.	Sensitive to sampling errors. In most applications, requires the estimation of the dimensionality of the signal.
CCA/SVD	Well understood and applied often.	Linear. No guarantee that the cross-correlation or covariances are larger than the correlations within each variable. Often pre-processed by extracting EOFs to avoid this problem. May need post-processing with linear transformations if more than one field is desired.
Cluster Analysis	Divides data into groups based on distance.	Numerous cluster methods available that give different results when applied to a single data set using the same distance measure. Since it is an exploratory tool, does not contain rules for assigning membership to independent observations.
Compositing	Since it involves only averaging, it is well understood.	Unless careful pre-screening of data has been performed, it is possible that multiple modes may be averaged and unrepresentative results can emerge.
Discriminant Analysis	Well suited to separation of a finite number of categories if linear separability is present. Numerous variations exist to allow or outliers and unequal variance in the groups. Rules learned to classify can be applied to independent data.	Linear separability is not often present in large scale problems. Variable selection may be computationally intensive.
Regression	Well understood in basic form. Many variations exist for correlated predictors, nonlinear relationships, and when outliers are present.	Traditional multiple linear regression makes numerous assumptions that are rarely met in climate analyses.
Neural networks	Allow for fitting nonlinear relationships	Can be complicated to fit properly. Can be computationally intensive for large datasets. Does not give good guidance on the physics of a problem as there are no constant weights.
Kernel methods	Allow for fitting nonlinear relationships.	Must test for an appropriate kernel to fit. Can be computationally intensive for large datasets. Unless a linear programming approach is used, does not give good guidance on the physics of a problem as there are no constant weights.

1) Correlation Patterns

The majority of analyses that seek to establish models' teleconnections of atmospheric variability are covariance or correlation-based. Both of these statistics measure the linear relationship between a set of variables. The covariance is the cross-product of the anomalies from the mean. Owing to this definition, covariance is used to assess eddy transports in models, since the mean is used to represent climatology. This definition underscores an implicit assumption of stationarity of the mean, which is rarely present in the atmosphere. The Pearson's correlation coefficient, commonly just referred to as the correlation coefficient, is a scaled version of the covariance, where the covariance is divided by the standard deviation of the fields. This provides a convenient range. Both these coefficients measure the property of two (or more) fields co-varying. It is also related to the mean squared error between two fields as the variance of one field multiplied by the correlation between the two fields times the anomaly of the second field. This gives rises to its popularity in the form of the "anomaly correlation". However, such a relationship assumes both fields are bivariate normally distributed, which is rarely the case, and that there is a linear mapping between the fields, as both correlation and covariance measure the linear portion of the relationship between fields. Relationships that are nonlinear cannot be measured by these metrics, nor does a value of zero indicate statistical independence, despite such a statement in numerous research papers. As a rule, investigators need to insure that the distributions of the variables are valid and the relationships linear before making inferences from the correlation.

Nonlinear functional relationships that are linear in the ranks can be measured by Spearman's correlation (Grantz et al., 2007). The distribution of the ranks should be approximately normally distributed. In both types of correlation coefficients, inference will require the use of the sample size. Any serial correlation will need to be accounted for by a degree of freedom calculation or a sampling strategy to remove serial correlation.

2) Multiple Regression

A prediction model using multiple regression gives a mean of y (predictand) conditioned upon a linear combination of the various x's (predictors). The model is linear in the parameters, although any parameter can be a nonlinear function of other variables. Interpretation of the model is similar to the simple regression model, except the variance accounted for by each predictor is examined as well as a multiple R^2 statistic. The statistical significance of each model parameter can be tested with an F-test (a multivariate extension of a t-test). An additional model assumption is that each predictor is independent of the others. This assumption is rarely met in practice. Mild deviations from independence seem to have little effect on the model, whereas moderate to large correlations between predictors can lead to model instability. This can be assessed through use of a condition number statistic. If necessary, alternative models, such as ridge regression (Peña and van den Dool, 2008) or principal component regression (Tippett et al., 2008) have been shown to hold promise in additional skill and stability when applied to highly correlated predictors, although the tradeoff for the former technique is that the unbiased property of least squares is abandoned through the addition of constraints and for the latter technique, interpretation of the predictors is often difficult. Implicit in the discussion of multiple regression is model selection to obtain the m predictors. The principle of a compact model is important and there may be many more potential predictors than m. Stepwise regression is often used to reject additional predictors (Ohring, 1972; Mercer et al., 2008).

3) Eigentechniques (EOF, PCA, SVD, CCA)

The use of eigentechniques was pioneered by Pearson in 1902 and formalized by Hotelling (1933) in a series of papers. In meteorology, the technique was named Empirical Orthogonal Functions (EOFs) by Lorenz (1956) who applied it to decompose a pressure data set. EOFs are unit length eigenvectors. The technique begins, implicitly or explicitly, with a correlation or covariance matrix that is decomposed into two new matrices, one of eigenvalues and one of eigenvectors. The use of these two matrices has played a central role in decomposing flows in the atmosphere into "modes of decomposition". The key ideas behind eigentechniques are to take a high dimensional problem that has structure (often defined as a high degree of correlation) and establish a lower dimensional problem where a new set of variables (e.g., eigenvectors) can form a basis set to reconstruct a large amount of the variation in the original data set. The idea is to capture as much signal as possible and omit as much noise as possible. While that is not always possible, the low dimensional representation of a problem often leads to useful results. Assuming that the correlation or covariance matrix is positive semidefinite in the real domain, the eigenvalue of that matrix can be ordered in descending value to establish the relative importance of the associated eigenvectors. Sometimes the leading eigenvector is related to some important aspect of the flow or teleconnection (Ding and Wang, 2005). In cases where the data lie in a complex domain, eigenvectors can be extracted in "complex EOFs". Such EOFs can give information on travelling waves, under certain circumstances, as can alternative EOF techniques that incorporate times lags to calculate the correlation matrix (Branstator, 1987). An alternative scaling of the eigenvectors leads to the principal component analysis (PCA) model. Both techniques are often used to filter correlated sets of times series arranged on a grid or array of stations into modes of variation. Such a decomposition requires the estimation of the number of modes that represent a geophysical signal. The techniques developed to accomplish this tend to be ad hoc (LEV test, Craddock and Flood, 1969) or based on white noise properties of the eigenvalues (North et al., 1982; Overland and Preisendorfer, 1982). Despite the widespread use of such tests, there has never been a formal linkage between the results of these tests and the true number of eigenvectors representing signal. Certain complications in such methods are the maximum variance property of the first eigenvector and the orthogonality of subsequent eigenvectors which tends to merge and smear known modes (Richman, 1986) as well as large sampling errors associated with those eigenvectors with closely spaced eigenvalues. In some cases, post-processing with an additional linear transformation of a reduced set of eigenvectors can help to draw out the modes that agree with correlation-based teleconnections. However, such analysis depends on correct determination of the number of signals in the data (Barnston and Livezey, 1987).

Canonical correlation analysis (CCA) is an extension of EOF/PCA for cases where pairs of fields are interrelated with the idea of finding couplings between the fields. The idea is to find a pair of patterns that maximize the correlation linear combinations of the eigenvectors of each field. A variation of CCA, known in meteorological research as singular value decomposition (SVD) maximizes the covariace between fields. Both techniques have been used routinely to generate medium range forecasts (Hwang et al, 2001; Shabbar and Barnston, 1996). Since CCA is an extension of PCA, the challenges of PCA, such as identification of the proper dimensionality and the effects of maximal variance and orthogonality are present in CCA (Cherry, 1996; Cheng and Dunkerton, 1995). Moreover, there is no guarantee that the desired cross correlation structure is large. In such cases, the correlations within each field may dominate. To avoid this problem, CCA/SVD is often pre-processed with EOF to extract uncorrelated vectors. This does solve the problem but makes interpretation that much more difficult.

Discriminant analysis is used to divide data into a number of linearly separable groups with similar membership among the variables within each group. Hence it is a discrimination and classification methodology. The function that discriminates is essentially a linear combination of the variables that is an eigenfunction. Skillful forecasts of temperature have been identified using this approach (DelSole and Shukla, 2006). Variable selection is crucial for multiple discriminant analysis (Lehmiller et al., 1997).

Another technique that is used often to find patterns of variability is cluster analysis. There are two broad families of clustering, hierarchical and non-hierarchical. These techniques have been found useful to group members of forecast ensembles (Mo and Ghil, 1988; Tracton and Kalnay, 1993). Gong and Richman (1995) present modes of rainfall variability for numerous cluster methods and distance measures to illustrate the differences between the techniques.

4) Composites

Modes of variability can be determined through composite analysis (averaging patterns with similar features). The technique is used often by synoptic meteorologists to determine dynamic fields of interest (Chen and Bosart, 1977). The idea of compositing has been extended to multivariate analyses for modes associated with the MJO (Weare, 2003). Climatological applications of composites would benefit from lessons learned from synoptic meteorology: all members of the composite pool need to be checked for consistency prior to averaging. This insures a unimodal distribution.

5) Kernel Techniques

Kernel techniques use a process that replaces an inner product with a kernel and then the solution is made in high dimensional "feature" space. In comparing kernel techniques to traditional linear approaches, Lima et al. (2009) report that the kernel technique offers significant skill improvements over traditional linear methods.

One example of the use of a kernel technique is shown in Figure A.1 where two classes exist (+, 0) and cannot be separated in two-space. The kernel projects the data into three-space and a linear separation is possible. Kernel techniques have a high potential for mode identification where linear low level modes provide ambiguous separability (e.g., the Arctic Oscillation versus the North Atlantic Oscillation). The most challenging aspect of kernel methods is finding the appropriate kernel to fit. Most often experiments using linear, polynomial and radial basis functions are fit and the method that generalizes best (highest skill on independent data) is selected.

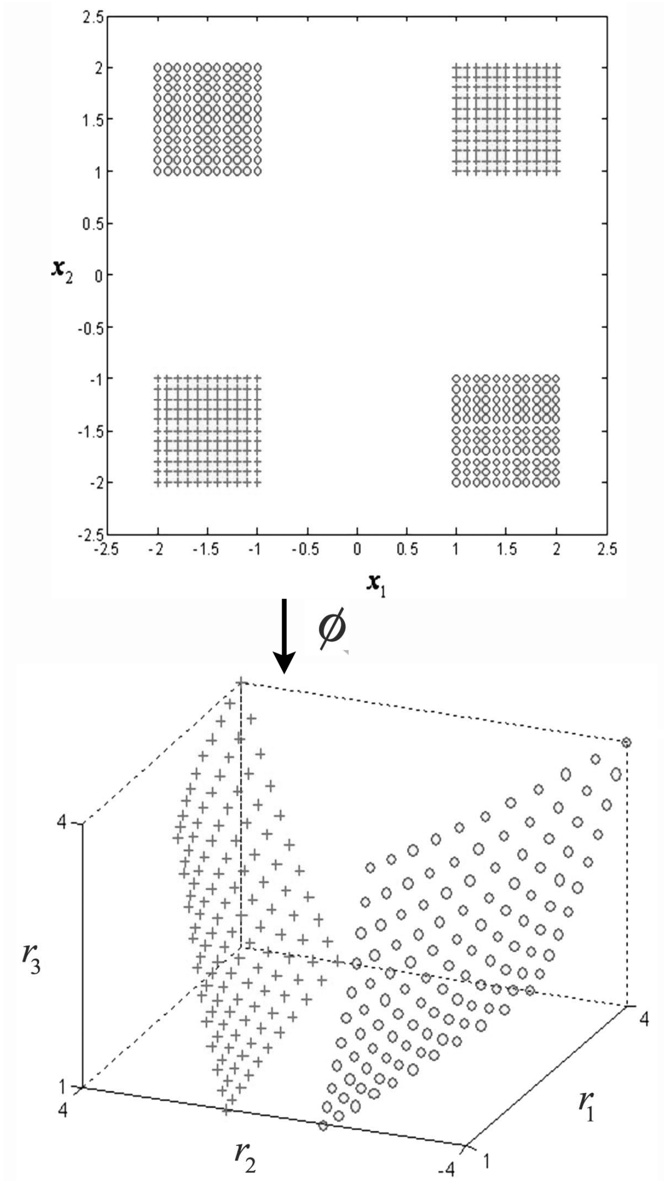

FIGURE A.1 A kernel map, ϕ, converts a nonlinear problem into a linear problem in the feature space. "+" belongs to positive class and "o" belongs to negative class. SOURCE: Richman and Adrianto (2010).

Appendix B

Committee Members' Biographical Information

Robert A. Weller (Chair)
Woods Hole Oceanographic Institution

Dr. Robert A. Weller is the director of the Cooperative Institute for Climate and Ocean Research and Chair of the Physical Oceanography Department at the Woods Hole Oceanographic Institution. His research focuses on atmospheric forcing, surface waves on the upper ocean, prediction of upper ocean variability, and the ocean's role in climate. Dr. Weller has been a pioneer in developing tools and technologies that enable scientists to investigate upper ocean processes on scales from meters to tens of kilometers and with accuracy never before available. In recognition of Dr. Weller's distinguished contributions to ocean science, he was named Secretary of the Navy/Chief of Naval Operations Oceanographic Research Chair by the Office of Naval Research. Dr. Weller has been on multiple mooring deployment cruises and has practical experience with ocean observation instruments. He served as co-chair of the U.S. Climate Variability and Change (CLIVAR) Scientific Steering Group and a member of the international CLIVAR Scientific Steering Group. He serves on the international Ocean Observing Panel for Climate and the NOAA Climate Observing System Council and Climate Working Group. He has served on several NRC committees, including the Committee to Review the U.S. Climate Change Science Program Strategic Plan, the Committee on Implementation of a Seafloor Observatory Network for Oceanographic Research, and the Committee on Utilization of Environmental Satellite Data: A Vision for 2010 and Beyond. He also served on the NRC Board on Atmospheric Sciences and Climate. Weller received his AB in Engineering and Applied Physics from Harvard and his Ph.D. in Physical Oceanography from Scripps Institution of Oceanography.

Jeffrey Anderson
National Center for Atmospheric Research

Dr. Jeffrey Anderson is a senior scientist at the National Center for Atmospheric Research where he heads the Data Assimilation Research Section. From 1992 to 2000 he was a scientist at NOAA's Geophysical Fluid Dynamics Laboratory; where he led the experimental prediction group. He has made research contributions in theoretical geophysical fluid dynamics, seasonal prediction, predictability, and ensemble data assimilation. His work is focused by a goal to improve geophysical prediction. He has an M.S. in computer science from the University of California, Berkeley and a Ph.D. in atmospheric and oceanic sciences from Princeton University.

Alberto Arribas
Met Office Hadley Centre

Dr. Alberto Arribas is the Manager of the Seasonal Forecasting group at the Met Office Hadley Centre where he is responsible for the research and development of new operational forecasting systems for intraseasonal-to-interannual timescales. Prior to this, Alberto has been heavily involved in the development of ensemble prediction systems for short- and medium-range forecasting, particularly in the area of representing model uncertainties. Other research interests include the application and use of probabilistic weather and climate forecast information. He received his BSc in Physics and PhD in Atmospheric Physics from the University Complutense (Madrid, Spain). He is a fellow of the Royal Meteorological Society and a lecturer for the World Meteorological Organization.

Robert E. Dickinson
University of Texas at Austin

Dr. Robert E. Dickinson is a Professor at the University of Texas at Austin, Jackson School of Geosciences. His research interests are in terrestrial and atmospheric interaction, terrestrial remote sensing. He has been contributing to the fields of climate modeling and global change research for over 40 years. Dr. Dickinson joined the staff of NCAR (National Center for Atmospheric Research) in 1968. In 1975, he became Head of the Climate Section and in 1981, Deputy Director of the Climate and Global Dynamics Division. During 1990–1999, Dr. Dickinson was Regents Professor at the University of Arizona in Tucson, where he held joint appointments in the Department of Atmospheric Sciences, the Institute of Atmospheric Physics, the Department of Hydrology and Water Resources, and the Laboratory of Tree-Ring Research. From 1999–2009, he was the Georgia Power Georgia Research Alliance Chair Professor of Atmospheric Sciences at the Georgia Institute of Technology. He has been active in committees, panels, and working groups of the NRC, IGBP, WCRP, and IPCC, and is a member of both the National Academy of Sciences and the National Academy of Engineering. His current research aims to prove the understanding of global and regional climate and earth system through the modeling of land, vegetation and radiative processes. He received his PhD in Meteorology from Massachusetts Institute of Technology in 1966.

Lisa Goddard
Columbia University

Dr. Lisa Goddard is a research scientist at the International Research Institute for Climate and Society (IRI) and an adjunct associate professor within the Department of Earth and Environmental Sciences of Columbia University. She has been involved in El Niño and climate forecasting research and operations since the mid 1990s. She has extensive experience in forecasting methodology and has published papers on El Niño, seasonal climate forecasting and verification, and probabilistic climate change projections. Currently leading the IRI's effort on Near-Term Climate Change, Dr. Goddard oversees research and product development aimed at providing climate information at the 10–20 year horizon and how that low frequency variability and change interacts with the probabilistic risks and benefits of seasonal-to-interannual variability. Most of Dr. Goddard's research focuses on diagnosing and extracting meaningful information from climate models and available observations. She also developed and oversees a new national post-doctoral program, the Climate Prediction Applications Postdoctoral Program (CPAPP), which explicitly links recent climate PhDs with decision making institutions. In addition, Dr. Goddard sits on five scientific advisory panels and co-chairs two working groups. Dr. Goddard holds a Ph.D. in atmospheric and oceanic sciences from Princeton University and a B.A. in physics from the University of California at Berkeley.

Eugenia Kalnay
University of Maryland

Dr. Eugenia Kalnay (NAE) is Distinguished University Professor for the Department of Meteorology, University of Maryland. Dr. Kalnay is a former holder of the Robert E. Lowry Chair, School of Meteorology, University of Oklahoma, and former director of the NOAA Environmental Modeling Center at the National Center for Environmental Prediction (NCEP). During Dr. Kalnay's tenure at NCEP, major improvements were made in the National Weather Service models' forecast skill. Her current research interests are in predictability and ensemble forecasting, numerical weather prediction, data assimilation and coupled ocean-atmosphere modeling. Dr. Kalnay is the recipient of several major awards including the AMS Jule G. Charney Award, the NASA medal for Exceptional Scientific Achievement, Dept of Commerce gold medals, and the Senior Executive Service Presidential Rank Award. Dr. Kalnay has served on several NRC committees, including the Panel on Digitization and Communications Science (2000–2002), the Committee on Weather Radar Technology Beyond NEXRAD (2001–2002), and the Board on Atmospheric Sciences and Climate (1988–1991). She also served as a member of the Commission on Geosciences, Environment, and Resources (1999–2000) and Head of NASA Goddard Global Modeling and Simulation Branch (1983–1986). Dr. Kalnay received the 2009 World Meteorology Organization IMO prize, which is the highest distinction of the World Meteorology Organization.

Benjamin Kirtman
University of Miami

Dr. Benjamin Kirtman has been a full Professor at the University of Miami—Rosenstiel School for Marine and Atmospheric Science since 2007. From 1993–2002 Dr. Kirtman was a research scientist with the Center for Ocean-Land-Atmosphere Studies and in 2002 joined the faculty of George Mason University as an Associate Professor. Dr. Kirtman uses complex coupled ocean atmosphere general circulation models to investigate the predictability of the climate system on timescales from days-to-decades and to study the influence of tropical variability on mid-latitude predictability and to assess how the annual cycle affects intraseasonal and interannual predictability. Current areas of interest include: El Niño prediction, dynamics and low frequency variations; impact of atmospheric stochastic forcing on coupled climate variability; El Niño Monsoon interactions; and the maintenance of the inter-tropical convergence zone. Currently, Dr. Kirtman is co-Chair of the US Clivar Prediction, Predictability and Applications Interface (PPAI) panel, co-Chair of the NOAA Climate Test Bed—Climate Science Team and co-Chair of the International Clivar Working Group on Seasonal to Interannual Prediction (WGSIP). Professor Kirtman is also an Executive Editor of Climate Dynamics. Professor Kirtman received is BS in Applied Mathematics from the University of California, San Diego in 1987, and his MS and Ph.D. in 1992 from the University of Maryland, College Park.

Randal D. Koster
National Aeronautics and Space Administration

Dr. Randal Koster has worked at NASA/GSFC since September of 1987, first as a member of the Hydrological Sciences Branch, and currently as a member of the GMAO. His early work focused on the analysis of global water isotope geochemistry. Most of his tenure at GSFC, though, has been dedicated to two research thrusts: (i) the development of improved treatments of land surface physics for atmospheric general circulation models, and (ii) the analysis of interactions between the land and atmosphere, using these models. He has examined many questions regarding land-atmosphere feedback, including: Can knowledge of soil moisture conditions at the beginning of a seasonal weather forecast improve the forecast? Can we find evidence in the observational record that variability in land surface states has an effect on rainfall, air temperature, and other atmospheric variables? Dr. Koster currently leads the development and maintenance of the land surface model component of the GMAO's Earth system model, a resource for the research community at large. Dr. Koster is a fellow of the American Meteorological Society. He has served on GEWEX and CLIVAR panels focused on land modeling and seasonal prediction. He served for many years as a lecturer for the climate program at George Mason University, teaching a course on land-climate interactions. He received his Sc.D from the Massachusetts Institute of Technology and his B.S. from California Institute of Technology.

Michael B. Richman
University of Oklahoma

Dr. Michael B. Richman has a wide range of interests, including analysis of global climate models, examination of the climate dynamics associated with El Niño/Southern Oscillation (ENSO), interaction of planetary- and synoptic-scale features, analysis of climate variability on both the intra-seasonal and interannual time scales, application of data mining to different radar platforms and statistical methodology. His work has involved analysis of four-dimensional climate models on supercomputers, using high-performance and massively parallel algorithms. Additionally, his expertise in statistical meteorology has led to development of multivariate techniques that summarize very large data sets, identifying their modal patterns, as well as eigentechniques that search for theoretical patterns in observed and modeled data. He has served several terms on both the American Meteorological Society's Committee on Probability and

Statistics and Committee on Artificial Intelligence Applications to Environmental Science PhD, University of Illinois.

R. Saravanan
Texas A&M University

Dr. R. Saravanan is a Professor at Texas A&M University. His research interests include the variability and predictability of climate on seasonal to millennial timescales, air-sea coupled interaction in both the tropical Atlantic and Pacific Oceans, and large-scale dynamics of the atmosphere and the oceans. His work has also addressed climate theory, hierarchical climate modeling, stochastic dynamics, and short-term climate prediction. Dr. Saravanan received his Master of Science in Physics from the Indian Institute of Technology, Kanpur in 1986 and his Ph.D. in Atmospheric and Oceanic Sciences from Princeton University in 1990.

Duane Waliser
Jet Propulsion Laboratory

Dr. Duane Waliser is a Senior Research Scientist in the Water and Carbon Cycles Group, in the Earth Sciences Section at the Jet Propulsion Laboratory in Pasadena, CA, a Visiting Associate in the Geological and Planetary Sciences Division at Caltech and an Adjunct Professor in the Atmospheric and Oceanic Sciences Department at UCLA. His principle research interests lie in climate dynamics and in global atmosphere-ocean modeling, prediction and predictability, with emphasis on the Tropics. His recent work at JPL involves utilizing new and emerging satellite data sets to study weather and climate as well as advance our model simulation and forecast capabilities, particularly for long-range weather and short-term climate applications. He received a B.S. in Physics and a B.S. in Computer Science from Oregon State University in 1985, a Masters degree in Physics from U.C. San Diego in 1987, and his Ph.D. in Physical Oceanography from the Scripps Institution of Oceanography at U.C. San Diego in 1992. He is presently Co-Chair, WCRP/WWRP-THORPEX Madden-Julian Task Force, Co-Chair of the Center for Multi-scale Modeling and Applications (CMMAP) MJO Working Group, Co-chair of the WCRP/WWRP-THORPEX Year of Tropical Convection (YOTC) Activity.

Bin Wang
University of Hawaii

Dr. Bin Wang is a Professor of Meteorology at the University of Hawaii. His current research themes include tropical intraseasonal oscillation, monsoons, ENSO, climate predictability and prediction, tropical cyclones, climate change, wave and instability, large-scale air-sea interaction, intermediate modeling of tropical climate. Dr. Wang's research approaches involve theoretical, numerical modeling, and observational analyses. His research efforts focus on understanding of the fundamental physics governing variations of weather and climate. Dr. Wang served as Co-Chair of the Asian-Australian Monsoon Panel (AAMP)CLIVAR/WCRP, member, CLIVAR/WCRP Science Steering Group and the American Meteorological Society

(AMS)/Committee on Interaction of the Sea and Atmosphere. Dr. Bin Wang received his MS in Meteorology from University of Science and Technology of China, Beijing and a Ph.D. in 1984 in Geophysical Fluid Dynamics from Florida State University.