

FUTURE SCIENTIFIC DIRECTIONS

Views from Early Career Scientists

The first National Center for Atmospheric Research (NCAR) Early Career Scientist Association Junior Faculty Forum on Future Scientific Directions in June 2003 was intended to promote interaction between junior scientists and faculty from the University Corporation for Atmospheric Research (UCAR) and its member universities by allowing young scientists to better define their interests in a group of peers and increase awareness of resources at UCAR and NCAR. The attendees came from many disciplines—including engineering, mathematics, biology, meteorology,

oceanography, and societal impacts—and separated into small breakout sessions to discuss specific topics, including the water cycle, interactions between land ecosystems and the atmospheric hydrologic cycle, and predictability. The following group of essays follows this format and looks ahead to what attendees believe to be the major issues facing three important interdisciplinary areas in their field. We hope that these brief summary articles stimulate a continuation of the wide ranging and candid discussion that took place at the forum.

Predictability

BY JOSHUA HACKER, JAMES HANSEN, JUDITH BERNER, YANGQUAN CHEN, GIDON ESHEL, GREGORY HAKIM, STEVEN LAZARUS, SHARANYA MAJUMDAR, REBECCA MORSS, ANDREW POJE, VITALII SHEREMET, YOUMIN TANG, AND COLLEEN WEBB

The study of predictability is multifaceted and appears in diverse fields. One important goal of the NCAR Junior Faculty Forum was to identify and understand different approaches to predictability

problems. For purposes of discussion—following the lead of Joseph Tribbia¹—we adopted the definition of predictability proposed by Thompson (1957), which is “the extent to which it is possible to predict [the atmosphere] with a theoretically complete knowledge of the physical laws governing it.” More precisely, we interpreted this as the state-dependent rate of divergence² of trajectories in phase space given complete knowledge of system dynamics. Therefore, predictability is intrinsic to a system, and the atmosphere (most likely) has predictability properties distinct from those of any model. Similar statements can be made about biological and all other dynamical

AFFILIATIONS: HACKER, BERNER, AND MORSS—National Center for Atmospheric Research, Boulder, Colorado; HANSEN—Massachusetts Institute of Technology, Cambridge, Massachusetts; CHEN—Utah State University, Logan, Utah; ESHEL—The University of Chicago, Chicago, Illinois; HAKIM—University of Washington, Seattle, Washington; LAZARUS—Florida Institute of Technology, Melbourne, Florida; MAJUMDAR—University of Miami, Miami, Florida; POJE—College of Staten Island, Staten Island, New York; SHEREMET—Woods Hole Oceanographic Institution, Woods Hole, Massachusetts; TANG—Courant Institute of Mathematical Sciences, New York, New York; WEBB—Colorado State University, Ft. Collins, Colorado

CORRESPONDING AUTHOR: Joshua Hacker, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307
E-mail: hacker@ucar.edu
DOI:10.1175/BAMS-86-12-1733

©2005 American Meteorological Society

¹ Tribbia, a senior scientist at NCAR and an expert in predictability of geophysical flows, spoke to the group on the state of the science and future challenges. A summary of his talk is not part of this essay, but the papers he considers seminal in the field are noteworthy: Thompson 1957; Lorenz 1963; Epstein 1969; and Leith and Kraichnan 1972.

² The rate of trajectory divergence is actually inversely proportional to the predictability of the system.

systems. We can exactly describe and solve for the evolution of some simple systems analytically, but we are faced with the frustrating reality that we cannot precisely know the predictability of more complex systems. Thus, much of our science is the pursuit of an unknowable goal.

Using this definition of predictability, we present three basic threads that emerged at the forum. First, we consider model error and initial-condition error, given that unknowable predictability means the two may never be completely separated. This leads naturally into the second topic, a discussion about the implications of the choice of norms used to measure the results of our studies. Finally, we address the potential for generalization of concepts and results.

INITIAL-CONDITION AND MODEL ERROR. Although model and initial-condition error have been addressed extensively in the literature (e.g., Tribbia and Baumhelfner 1988), the synergy between error sources has kept quantification of their respective importance elusive. Thompson's (1957) definition of predictability implies that we can measure the predictability of a model, but the predictability of a physical system cannot be precisely known without a perfect model or a very long and precise observational record. Unfortunately, it is impossible to perfectly observe geophysical systems, biological systems, and many engineering processes; thus, we eternally face initial-condition uncertainty. Model error is equally unavoidable for complex systems. It inhibits both our ability to forecast and our physical understanding of a system because models are indispensable tools for studying a physical system. By improving our models using a combination of empiricism, physical understanding, and computational power, we hope to make better estimates of the predictability of the physical system. Thus, attempting to forecast the physical system in the face of initial-condition and model uncertainty is tantamount to seeking a fundamental property of the system—its predictability.

A model M propagates its state, x , from time 0 to t following

$$x_t = M(x_0) + f,$$

where f includes all external forcing and may include spatial and temporal dependence. Initial-condition and model error is typically found in x_0 and M , respectively. The problem at hand determines whether boundary-condition error should be considered as

RECURRING THEMES


- Model error and initial-condition error
- The choice of norms to quantify results
- Generalization across disciplines

part of model error or as a third, independent source of uncertainty. In discretized models, boundary conditions can be included in either M or f , both of which include errors. Incorporating them into the model M implies that boundary-condition error is part of model error. But if boundary conditions are specified in f , then boundary-condition error must also be considered elsewhere, and one might choose to conceptually include it as an independent source of uncertainty.

Approaches to understanding the predictability of a physical system or a model need not coincide with efforts designed to improve forecasts. Predictability is a system property that depends on intrinsic dynamics. Fortunately, complete understanding of the system is not a prerequisite for forecasting, as demonstrated by the success of operational weather prediction. Because many forecasting applications rely heavily on statistical parameterizations of poorly understood processes, research results may lead to forecast improvement without a concomitant improvement in physical understanding. Carefully defining research goals as either forecast improvement or physical understanding is an important step toward interpreting results and designing studies.

Although strict separation of model and initial-condition error is likely impossible, acknowledging forecast improvement and physical understanding as separate pursuits motivates further attempts at estimating model and initial-condition error. The relative importance of each may shift depending on the goal and the type of model considered. In the atmospheric sciences, we typically focus on hydrodynamic models expressed as systems of differential equations, but other classes may be considered. Statistical and empirical models can be excellent tools for improving forecasts and should not be dismissed for being an “engineering” approach or “not derived from basic physical principles.” Separately, they may play an important role in fundamental studies to identify basic system properties.

A theory for the statistics of model error is currently unavailable, but the theory of state estimation



has a long history as an inherently probabilistic problem. To approach an optimal estimate, we must seek ways to reduce model error, the impact of simplifying assumptions, and sensitivity to the chosen norm. At best, we can hope to gain an estimate of model error by trying to understand the impact of an inadequate model on probabilistic state estimation.

Probabilistic state estimation and forecasting may be a useful tool for attempting to disentangle initial condition and model error. Error can be systematic or appear random. A distribution that appears random does not ensure that the underlying process is truly random, but it does allow access to probability theory for describing it. From large samples we can estimate an initial-condition error distribution and a forecast error distribution. Due to the Bayesian nature of state estimation, the initial-condition error distribution can never be fully separated from the model, but we can make choices in a state-estimation algorithm to emphasize the observations (i.e., minimize the error at $t = 0$). The forecast error distribution includes both effects, and we might begin to estimate model error distributions by removing the estimated contribution of initial-condition error.

In state estimation the difficulty arises from the fact that distributions drawn from a model may never be the same as those from the true physical system. Because an imperfect model may not have access to the correct distributions, it cannot produce correct probabilistic forecasts to be used as a first-guess for state estimation. The scientific community will no doubt continue improving models by reducing deterministic model error, but gaps will always exist in our understanding and computational capabilities. Statistics may help fill those gaps by accounting for model error such that the deterministic components of first-guess forecasts do not destroy a probabilistic state estimate derived from a good observing system.

A few topics are relevant for any approach to characterizing initial condition and model error, and to understanding their relationship. Observation networks may be designed for the purpose of identifying model error, and preliminary work is needed to identify potential designs. Fundamental research is needed to understand the impact of spatially and temporally correlated observational errors on state estimation and forecast assessment. The impact of strong nonlinearities (e.g., a threshold) on different error sources should be better understood. Finally, a practical consideration is how to deal with the vast

quantities of data as modern observation platforms continue to be deployed.

Here we have expressed that we want to measure initial-condition and model error. We next explore the implications of choosing how to measure it, and suggest some alternative methods.

THE IMPORTANCE OF THE NORM. In any quantitative study we must choose a norm or metric by which we evaluate results. For our purposes, the norm determines a magnitude (such as an error-vector length) in phase space. Though infinite for many physical systems, the phase space of a model can be large, but is finite, and computational and observational constraints have led to norms in a subspace of the model phase space. In the atmospheric sciences, norms have traditionally been observable quantities evaluated on critical levels, such as 500-hPa geopotential heights. More recently, norms based on energy have become favorable because they include a larger subspace that represents more degrees of freedom in the system dynamics. Whether the goal is understanding system dynamics or producing the greatest forecast skill, we are free to choose the norm. The norm chosen to answer questions about system dynamics may be different from one chosen to maximize forecast utility—and interpretation becomes more difficult when considering forecast value.

In the context of a forecast cycle, the choice of norm and subspace has profound implications for both state estimation and verification. The relationship between state estimation and verification is complex, and choosing the same norm for both is satisfying because it is consistent, but it is certainly not a requirement. One example is the practice at the European Centre for Medium-Range Weather Forecasts (ECMWF) of using a 48-h energy norm to determine the initial perturbations in their ensemble forecasting system and later verifying quantitative precipitation forecasts. If the goal is a skillful precipitation forecast, then why not design perturbations based on some “precipitation norm” instead? While compelling arguments may lead to a variety of answers, the best choice may still be the norm most closely related to the problem at hand.

When the goal is to understand system behavior, the results can be sensitive to the choice of norm and subspace, and a thorough treatment should include evaluation with more than one. In the long term we should seek norms, or groups of norms, that expose similar dynamics. Basic research is

required to establish whether such groups exist. Assessing the usefulness of norms based on information theory, such as relative entropy or mutual information (Schneider and Griffies 1999), may be a starting point.

Inasmuch as end users define the value of a forecast, state estimation norms based on user needs may lead to a more socioeconomically valuable forecast. The possibility of user-defined norms to estimate or measure value admits that the most skillful forecast in physical terms may not be the most valuable. Even a perfect forecast with respect to a norm defined by a class of users, however, need not result in maximum value for all in that class. For example, a small farmer may not be able to adjust crops to account for a forecast of freezing weather because he lacks access to the forecast information or the resources to make an adjustment, and thus is at a disadvantage relative to a modern agricultural company. Defining classes of users that can benefit similarly from a forecast is a complex problem in itself, and it combines physical and socioeconomic issues.

User-defined norms could also rely on information external to the physics of the system we are modeling. If we care about increasing forecast value, then the established practice of verifying state variables, or even diagnostic variables that are primary to the modeled system, may not be the best approach. Rather, defining norms with variables that are not immediately part of our physical system may lead to more useful forecasts and even physical insight in the absence of rigorous physical understanding. For example, a biological problem might include a norm for animal health when producing a state estimate for a vegetation model. In ocean forecasting, shipping efficiency may be considered. Again, specific goals determine the norm or subspace by which one should evaluate success. Regardless of one's definition of value, deduction from quantitative results that are extremely sensitive to the choice of norm may be insignificant, and seeking norm-insensitive results seems prudent.

TOWARD GENERALIZATION. So far it would seem that widely generalizable results may be impossible to achieve, but in fact generalization may be both possible and profitable. Scientific studies with broad implications potentially contribute the most to our understanding of the natural world. They may cross disciplinary boundaries or cover a wide range of problems within one discipline, accelerating learning

with a greater exchange of ideas. Generalization of predictability studies is also likely to promote rapid progress. Because of norm-dependence, initial-condition and model error, and our lack of understanding of some physical systems, generalization has proven difficult. But generalization may be facilitated by seeking different bases for system classification and cross-disciplinary communication.

The geophysical sciences lend themselves to a specific type of generalization based on mathematical representation of a dynamical system with differential equations. In the context of predictability this has led to hierarchical studies where certain characteristics of simple systems, which are arguably similar to more complex systems, are generalized.

Although hierarchical approaches are satisfying because simple systems are computationally inexpensive, can sometimes be analytically tractable, and have results that can often be unambiguously interpreted, the results may not always withstand experimentation with more complex systems. One alternative approach to generalization is to seek different bases for system classification, and a natural place to begin is by grouping nonlinear systems, since nonlinearity is ubiquitous. Examples of characteristics appearing in many types of nonlinear systems are multiple states, bifurcation, thresholds, transition to chaos, hysteresis, scale cascades, and the existence of coherent structures. Atmospheric, oceanic, ecological, biological, and engineering control systems all demonstrate one or more of these attributes, though they may not be modeled with similar sets of equations. Each of those sciences may stand to benefit from methods and results already established in the others. We may similarly benefit from looking outside of geophysics, biology, or engineering to find other systems that display threshold phenomena, coherent structures, and nonlinear error growth.

Systems may also be grouped by other characteristics that are not typically considered in the geophysical sciences, thereby engendering interaction with other scientists. One characteristic, for example, is robustness, which essentially means insensitivity to perturbations. Biological, control, and computer systems are classified as robust when they have certain levels of diversity, redundancy, modularity, and control (Carlson and Doyle 2002). This particular set of classification criteria may or may not have applications to geophysical systems, but the process that led to identification of a robust class of systems may be useful.

A significant barrier to successful cross-disciplinary interaction is the use of incompatible terminology (this is even evident among the geophysical scientists). Despite the hurdles, we believe that expanding informal discussions and perhaps organizing a larger, more inclusive predictability workshop would likely prove beneficial.

To summarize, we are optimistic that generalization is possible, despite the difficulties, by seeking different bases for system classification and fostering cross-disciplinary interaction. Seeking classification bases outside our own fields may lead to new bases and ultimately accelerate the learning process. It will take effort to expand our experience, but the return on scientific progress may prove substantial.

ACKNOWLEDGEMENTS. We thank Joe Tribbia for his contribution to our discussions at the forum. NCAR and UCAR are thanked for the sponsorship of the forum. NCAR is operated by UCAR under the sponsorship of the National Science Foundation.

FOR FURTHER READING

- Carlson, J., and J. Doyle, 2002: Complexity and robustness. *Proc. National Acad. Science, Suppl. 1*, **99**, 2538–2434.
- Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus*, **21**, 739–759.
- Leith, C., and R. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.*, **29**, 1041–1052.
- Lorenz, E., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.
- Schneider, T., and S. Griffies, 1999: A conceptual framework for predictability studies. *J. Climate*, **12**, 3133–3155.
- Thompson, P., 1957: Uncertainty in the initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus*, **9**, 275–295.
- Tribbia, J., and D. Baumhefner, 1988: The reliability of improvements in deterministic short-range forecasts in the presence of initial state and modeling deficiencies. *Mon. Wea. Rev.*, **116**, 2276–2288.

The Father James B. Macelwane Annual Award

AMS UNDERGRADUATE AWARD

The Father James B. Macelwane Annual Award was established by the American Meteorological Society to honor the late Rev. James B. Macelwane, S.J., a world-renowned authority of seismology, who was a geophysicist and Dean of the Institute of Technology, Saint Louis University, until his death in 1956. The recipient of the Father James B. Macelwane award will receive a stipend of \$300.

The purpose of this award is to stimulate interest in meteorology among college students through the encouragement of original student papers concerned with some phase of the atmospheric sciences. The student must be enrolled as an undergraduate at the time the paper is written, and no more than two students from any one institution may enter papers in any one contest.

Submission of Papers:

To consider papers for the Macelwane Award, the AMS Committee of Judges must receive the following: 1) an original copy of the paper; 2) a letter of application from the author, including contact information, stating the title of the paper and the name of the university at which the paper was written; 3) a letter from the department head or other faculty member of the major department, confirming that the author was an undergraduate student at the time the paper was written, and indicating the elements of the paper that represent original contributions by the student; and 4) an abstract of no more than 250 words.

The above information must be received by **9 June 2006**. Mail to American Meteorological Society, Father James B. Macelwane Award, 45 Beacon Street, Boston, MA 02108-3693 The evaluation of papers occurs during the summer. Announcement of the award recipient will be made in October of 2006.