

FUTURE SCIENTIFIC DIRECTIONS: PREDICTABILITY

Joshua Hacker*

National Center for Atmospheric Research,[†] Boulder, CO

James Hansen

Massachusetts Institute of Technology, Cambridge, MA

Judith Berner

National Center for Atmospheric Research, Boulder, CO

YangQuan Chen

Utah State University, Logan, UT

Gidon Eshel

The University of Chicago, Chicago, IL

Gregory Hakim

University of Washington, Seattle, WA

Steven Lazarus

Florida Institute of Technology, Melbourne, FL

Sharanya Majumdar

University of Miami, Miami, FL

Rebecca Morss

National Center for Atmospheric Research, Boulder, CO

Andrew Poje

College of Staten Island, Staten Is., NY

Vitalii Sheremet

Woods Hole Oceanographic Institution, Woods Hole, MA

Youmin Tang

Courant Institute of Mathematical Sciences, New York, NY

Colleen Webb

Colorado State University, Ft. Collins, CO

* *Corresponding author address:* Joshua Hacker, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307. Email: hacker@ucar.edu

†The National Center for Atmospheric Research is sponsored by the National Science Foundation

Abstract

A group of junior faculty members and UCAR junior scientists convened in Boulder, CO on June 16-18, 2003 for discussion on future scientific directions. This report summarizes the goals and products of one of the three foci selected for special consideration: predictability. About 15 people, representing physical, mathematical, and biological sciences, were present for round-table discussions. The discussion sought common interpretation of the predictability problem, points of generalization, identification of major hurdles, and potential approaches to their solution. The diverse background of the participants generated a wide-ranging discussion. The participants addressed predictability generally, while supplying specific examples from their own areas of expertise. Recurring themes included the relationship between models and initial conditions, the importance of definitions and the choice of a norm for evaluation, and generalization across systems and disciplines. The group explored potential avenues for generalization through interdisciplinary networking. Short- and long-range challenges were identified related to probabilistic state estimation, verifying predictions and understanding error, and dealing with nonlinearity. In this essay we expand on these themes and challenges, and describe possible future research objectives.

1. Introduction

The first NCAR Early Career Scientist Association Junior Faculty Forum on Future Scientific Directions was intended to promote interaction between junior scientists and faculty from 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 format consisted of small break-out sessions to discuss specific topics including: the water cycle, interactions between land ecosystems and the atmospheric hydrologic cycle, and predictability. This essay summarizes the predictability discussions.

The small group of predictability session attendees (the authors of this essay) can best be described as interdisciplinary, representing the atmospheric and oceanic sciences, biology, mathematics, and engineering. Two attendees brought additional expertise in the role of science in societal decision-making. This diversity was encouraged to promote cross-disciplinary understanding and initiate thoughts on generalization, and it produced a wide-ranging, candid discussion. Prior to the forum, several discussion topics were identified and a preliminary “white paper” was circulated among the participants, who were given a chance to suggest modifications. But the sessions demonstrated little adherence to this agenda, and a free-form scientific discussion ensued, from which a few threads emerged.

The study of predictability is multifaceted and appears in diverse fields. One important goal of these sessions was to identify and understand different approaches to predictability problems. Joseph Tribbia, a Senior Scientist at NCAR and an expert in predictability of geophysical flows, was asked to address the group and provide his views 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).

The group agreed, following Tribbia's talk, to adopt the definition of predictability proposed in 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 divergence¹ 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 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 to frame the discussions, we present three basic threads that emerged among the forum participants. First, the importance of model error and initial-condition error are considered. While these topics are not new, strategies for understanding them are explored, given the fact that unknowable predictability means the two may never be completely separated. This leads naturally into 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.

¹The rate of trajectory divergence is actually *inversely* proportional to the predictability of the system.

2. Summary of discussion

a. Initial-condition and model error

Although model and initial-condition error have been addressed extensively in the literature (e.g. Tribbia and Baumhefner 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, \mathbf{x} , in time t following

$$\mathbf{x}_t = M(\mathbf{x}_0) + \mathbf{f}(t) ,$$

where \mathbf{f} includes all external forcing. Initial-condition and model error is typically found in \mathbf{x}_0 and M , respectively. The problem at hand determines whether boundary-condition error should be considered as part of model error, or as a third, independent source of uncertainty. In discretized models boundary conditions can be included in either M or \mathbf{f} , both of which include errors. In-

incorporating f 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.

For practical research purposes, 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.

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.” They may also play an important role in fundamental studies to identify basic system properties.

Although strict separation of model and initial-condition error is likely impossible, probabilistic state estimation and forecasting may be a useful tool for attempting to disentangle the two. 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 model 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.

A few topics are relevant for any approach to characterizing initial condition and model error, and 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.

b. 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. 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 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. Inasmuch as end users define the *value* of a forecast, state-estimation norms based on user needs may lead to a more socio-economically valuable forecast. Conversely, a perfect forecast with respect to a norm need not result in maximum value. For example, a small farmer may not be able to adjust crops to account for a weather forecast 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.

To address forecast value we might consider norms (possibly user-defined) that contain more than the physics of the system we are modeling. For example, in biology it 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.

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.

So far it would seem that widely generalizable results may be impossible to achieve, but the discussion summarized in the next subsection identifies places where generalization may be both possible and profitable.

c. Towards generalization

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 these 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 approach to generalization is to seek different bases for system classification, and a natural place to begin is by grouping nonlinear systems. 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.

Systems may also be grouped by other characteristics that are typically not 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 was even evident amongst the geophysical scientists!). Despite the hurdles, it was agreed that expanding informal discussions and perhaps organizing a larger, more inclusive, predictability workshop would likely prove beneficial.

To summarize, the group was optimistic that generalization is possible, despite the difficulties, by seeking different bases for system classification and fostering cross-disciplinary interaction. It will take effort to expand our experience, but the return on scientific progress may prove substantial.

3. Conclusions

The study of predictability is diverse and interdisciplinary. Attendees of the predictability sessions at the UCAR Junior Faculty Forum on Future Scientific Directions came with diverse backgrounds that led to a wide-ranging, candid discussion. After agreeing to frame discussion topics with the definition of predictability first proposed by Thompson (1957), three threads emerged: initial-condition and model error, the importance of the norm, and generalization. By cutting across these three topics, we can identify some immediate challenges (perhaps without immediate solutions) related to state estimation, verification, and nonlinearity.

Probabilistic approaches to state estimation and forecasting offer hope for beginning to disentangle model and initial-condition error. To characterize uncertainty, which is here to stay in our observations and treatment of most physical systems, we can use the language of statistics and probability. 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.

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.

The relationship between verification and state estimation is complex. The possibility of adding user-defined norms to determine optimal estimates or measure value as opposed to skill, is intriguing. It admits that the most skillful forecast may not be the most valuable, implying that we must understand and carefully define goals before choosing a norm. 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. 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.

Finally, the ubiquity of nonlinearity suggests that it is a natural system trait where one might begin classifying systems for the purpose of generalization. We can look outside geophysics, biology, or engineering to find other systems that display threshold phenomena, coherent structures, and nonlinear error growth in general. This may lead to new bases on which to classify systems, and accelerate the learning process.

Acknowledgements. We thank Joe Tribbia his contribution to our discussions at the forum. NCAR/UCAR are thanked for the sponsorship of the forum. The National Center for Atmospheric Research is operated by the University Corporation for Atmospheric Research under the sponsorship of the National Science Foundation.

Sidebar 1: Meeting Summary

- UCAR/NCAR Junior Faculty Forum on Future Scientific Directions
- June 16-18, 2003
- NCAR, Boulder, CO
- Organized to promote interaction between junior scientists and faculty by allowing young scientists to better define their interests in a group of peers.

Sidebar 2: Recurring Themes

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

References

- 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.