PREDICTABILITY, OBSERVATIONS, AND UNCERTAINTIES IN THE GEOSCIENCES

BY MILIJA ZUPANSKI AND I. MICHAEL NAVON

Uncertainty estimation is becoming an important new research discipline, crosscutting many scientific areas. The mathematical concept of uncertainty estimation is based on probability theory and statistics, estimation theory, information theory, and control theory. Theoretical aspects of uncertainty estimation are generally well understood for linear models (operators) and Gaussian distribution. In the geosciences, however, nonlinear models are typically used; thus, the Gaussian probability assumption may not be the best option. In addition, models of geosciences systems are typically high-dimensional, with state variable dimensions of the order of 10^6–10^7. At the same time, the mathematical concept of uncertainty estimation, algorithmically defined by smoothing and/or filtering, is relatively simple, and a common mathematical framework can be applied across disciplines. These facts create a challenging problem for uncertainty estimation, requiring new scientific developments and cross-disciplinary efforts.

AFFILIATIONS: ZUPANSKI—Colorado State University, Fort Collins, Colorado; NAVON—The Florida State University, Tallahassee, Florida
CORRESPONDING AUTHOR: Milija Zupanski, Cooperative Institute for Research in the Atmosphere, Colorado State University, 1375 Campus Delivery, Fort Collins, CO 80523-1375
E-mail: ZupanskiM@cira.colostate.edu

DOI:10.1175/BAMS-88-9-xxx

In final form 27 April 2007

WORKSHOP ON PREDICTABILITY, OBSERVATIONS, AND UNCERTAINTIES IN GEOSCIENCES

WHAT: More than 50 scientists from around the world discussed methods and research challenges in the estimation of uncertainties in the geosciences and related disciplines.
WHEN: 13–15 March 2006
WHERE: Tallahassee, Florida

By creating new ties and collaborations between practitioners and theoreticians of seemingly unrelated scientific disciplines, the research on uncertainty estimation can be greatly advanced and facilitated. This workshop was a step in that direction, especially aimed at young scientists from all facets of applications and mathematics. The idea of the workshop was to learn about and to discuss common mathematical concepts behind uncertainty estimation. Predictability, data assimilation, ensemble forecasting, ensemble Kalman filters, and information theory were all seen as components of a general uncertainty estimation theory, working together toward achieving the same goal: the reduction of uncertainty in high-dimensional nonlinear systems.

The backgrounds of the workshop participants encompassed broad scientific areas, ranging from hydrology, weather, and climate to geology, nuclear sciences, and mathematics. The workshop had more than 50 attendees, many of them students from The Florida State University (FSU) and neighboring universities. At the three-day workshop (presenta-

Au: per AMS style URLs are not allowed in footnotes and are incorporated into the text. Okay as set?

tions from the workshop are available online at www.scs.fsu.edu/ws_MathGeo_Mar06_prog.php),[1] 22 papers covering uncertainty estimation in climate, weather, hydrology, nuclear science, and geology were presented. Because of the close relationship between predictability and uncertainty, many of the presented works included results in the context of chaos theory. Discussions were very helpful, fostering a better understanding of the issues and terminology used in the different disciplines and applications.

Interestingly, model error and parameter estimation were quickly established as the most relevant issues for ensemble-based uncertainty estimation. Few presenters covered this topic, but many mentioned the importance of estimating model-related uncertainty, including empirical parameter estimation as well as multimodel ensembles.

CHAOS THEORY AND UNCERTAINTY ESTIMATION. The importance of chaos theory in uncertainty estimation was underlined in an inspiring presentation by Ed Lorenz (Massachusetts Institute of Technology). Lorenz showed that chaos is abundant, but also regime dependent. Zoltan Toth [National Centers for Environmental Prediction (NCEP)] suggested that model error can be accounted for by "remapping" observations, eventually producing initial conditions consistent with the model attractor. Stephane Vannitsem (Royal Meteorological Institute, Belgium) pointed out the following two types of model error: a parametric error and a space truncation error. He found that large spatial scales of model error have a pronounced positive impact on the forecast. The consensus from Lorenz and his fellow presenters on the subject was that model error estimation needs to be addressed within the context of chaos.

METHODOLOGIES FOR UNCERTAINTY ESTIMATION. In general, uncertainty estimation methods presented at the workshop were deterministic (adjoint based) and probabilistic (ensemble based). It was also pointed out in several presentations that control theory brings an important advantage for parameter uncertainty estimation. For example, Francois LeDimet (Joseph Fourier University, France) suggested that the relevance of the second-order adjoint methodology, related to general sensitivity analysis, is a valuable tool in error covariance estimation, while Allesandro Petruzzi (University of Pisa, Italy) underlined the importance of estimating the timing and the quantity of uncertainties in applications to a simulated nuclear power plant failure using a variant of an adjoint-based method. As well, Bill Hu (FSU) discussed the important role of control theory in parameter uncertainty estimation in geological applications to heterogeneous media and tracer tests.

The ensemble-based uncertainty estimation algorithms discussed at the workshop were the ensemble adjustment Kalman filter (EAKF), presented by Jeffrey Anderson (National Center for Atmospheric Research), the localized ensemble transform Kalman filter (LETKF), presented by Eugenia Kalnay (University of Maryland), and the maximum likelihood ensemble filter (MLEF), described by Milija Zupanski (Cooperative Institute for Research in the Atmosphere at Colorado State University).

Similarities and differences between variational and ensemble-based methodologies were of special interest. It was suggested that both the four-dimensional variational data assimilation and ensemble Kalman filter (EnKF) give comparable results, although there is an overall practical advantage of the EnKF. The use of an iterative minimization with improved Hessian preconditioning within the MLEF was suggested as a way to address the nonlinearity of observation operators. The hierarchical Bayesian methodology was additionally suggested as a way to improve the EAKF capability to simultaneously adjust parameters and other specifications of the system for adaptive error correction.

Improvements to the above-mentioned methodologies were also discussed, including the need for improving existing variational and ensemble data assimilation methods by adding a capability for non-Gaussian state and observation errors. Based on control theory and these ideas, a non-Gaussian generalization of the MLEF algorithm was presented. The use of the proper orthogonal decomposition method was suggested as a possibility to address the efficiency of the EnKF methods, by inflating/deflating the ensemble size.

In addition to data assimilation, the mentioned methodology improvements were also related to ensemble forecasting and its potential for computational savings. The mentioned applications were related to the use of the NCEP operational short-range ensemble forecasting (SREF), and to the development of a hybrid dual-resolution EnKF method. The means to convey the probabilistic information from an ensemble forecast to the public brought forth the

---

[1] This workshop was supported by the School of Computational Science at The Florida State University and by the National Science Foundation.

mention that NCEP plans to issue a forecast uncertainty product for the upcoming 2008 Beijing, China, Olympics.

CLIMATE UNCERTAINTY ESTIMATION. The importance of parameter estimation in climate and the advantage of using the EnKF in particular were stressed by James Annan (Frontier Research Center for Global Change, Japan). He said the EnKF can successfully perform the task, and can do it a few orders of magnitude more efficiently than other existing parameter estimation methods. There is still a need, however, to improve the nonlinear aspects of the problem. T. N. Krishnamurti (FSU) discussed the challenges of modeling seasonal climate, in particular a diurnal cycle, and suggested using a multimodel ensemble to predict it. In applications to seasonal climate predictability associated with land surface, Rolf Reichle (National Aeronautics and Space Administration, Global Modeling and Assimilation Office) pointed out the need to use remote sensing observations of soil moisture and temperature. He suggested that a scaling approach may help in reducing large biases between satellite and model estimates. He also mentioned the need for more work in snow assimilation and coupled land–atmosphere systems. Alicia Karspeck (NCAR) examined the problem of error covariance inflation and localization in applications to El Niño–Southern Oscillation. She suggested that localization may not be always beneficial, probably because of the system's low dimensionality and the need to maintain balance in the system.

OTHER APPLICATIONS. There were other novel applications of the ensemble and related methodologies for uncertainty estimation. One such application was to the estimation of uncertainty of image restoration in geostatistical techniques using remote sensing. Multiple techniques to reduce uncertainty estimation produced similar results, yielding the suggestion to use the easiest and most efficient technique from now on. A novel technique for sensitivity analysis using an ensemble was also promoted. Because assumed infinitesimal perturbations make using the adjoint-based sensitivity technique difficult, it was suggested that ensembles be used for both the forward and backward sensitivity analysis, because they better cope with larger physical perturbations. Another technique suggested using an archive of retrospective analyses for the statistical correction of biases in the NCEP ensemble precipitation forecasts.

OUTCOME. This workshop indicated the following important research avenues relevant to uncertainty estimation in the geosciences: (i) model error uncertainty estimation; (ii) parameter uncertainty estimation, especially in climate and hydrology applications; (iii) development of non-Gaussian data assimilation techniques; and (iv) further exploration of the link between uncertainty estimation, control theory, and chaos theory. The complexity of these issues clearly indicates the need to foster this kind of interdisciplinary meeting, eventually leading to new ideas and collaborative work. Our hope is that such meetings will become more frequent, eventually providing a mechanism for the exchange of new ideas at all levels.

ACKNOWLEDGMENTS. We are grateful to the SCS at FSU and its director, Professor Max Gunzburger, for their financial support and encouragement, essentially for making the workshop possible. We are indebted to Stephane Vannitesem and Zoltan Toth for carefully reviewing the manuscript. Special thanks go to Ms. Anne Johnson and Ms. Eva Ronquist for making the workshop organization impeccable and transparent, and to Ms. Lynn LaCombe for invaluable administrative help. We are also grateful to the National Science Foundation Collaboration in Mathematical Geosciences for its support under Grants ATM-0327818 and ATM-0327651.