The Need for a National Data Assimilation Education Program

BY T. VUKICEVIC, E. KALNAY, AND T. VONDER HAAR

The major progress made over the last two decades in numerical weather prediction (NWP) and climate change research can be attributed to advances in three areas: modeling, observing methods, and, very importantly, the improved utilization of observations, a new discipline known as data assimilation. As new technology for observing the atmosphere, ocean, and land surface—ranging from submersible buoys to multispectral satellite observations—is implemented, a significant challenge will be to optimally process large volumes of this new information. This challenge implies that data assimilation methods should advance alongside modeling and modern observing. Serafin et al. (2002) review current and future weather information technology, emphasizing that training and education are integral to technology development and transfer. We propose that education and training should also be extended into the domain of data assimilation.

SCOPE AND UTILITY DATA ASSIMILATION TECHNOLOGY. The purpose of data assimilation is the "use of all available information to determine as accurately as possible [the] state of the atmospheric or oceanic flow" (Talagrand 1997). This information includes not only the observations, but also our knowledge about physical laws, applied in a variety of environmental models. In order to maximize the extraction of information, the modeled state needs to be optimally combined with the observations. During the 1990s there was a boom in the development of optimal data assimilation methods in the atmospheric and ocean sciences, but the algorithms based on general optimality criteria are not yet fully implemented in practice. The limitations in practical applications result from a combination of a) large computational challenges and b) incomplete understanding of optimal representation of information content and uncertainties in the models and observations relative to the true state. Consequently, achievement in data assimilation requires a diversity of expertise, including meteorology, atmospheric chemistry, oceanography, hydrology, biosphere and soil dynamics, applied mathematics, statistics, observation techniques, and computer sciences. An interdisciplinary research team approach is needed, beyond atmospheric sciences, to further our ability to process the ever-increasing diversity and volume of environmental observations. This defines the need to have new educational programs at the U.S. universities that would support requirements of future data assimilation developments in the Earth system sciences.

Optimal data assimilation methods also provide powerful research tools because they explore the relationships between response and control of dynamic systems and associated predictability and stability properties (Kalnay 2002). Data assimilation methods can and should be used to test hypotheses put forward by the models by way of quantitative studies of model uncertainties; by sensitivity of retrieved unobserved quantities to representation of processes (or lack of it); or simply by way of inverse analysis of controls for predefined response functions depending not on the observations but on known integral properties.

The potential advantages of data assimilation science cannot be attained because currently there is a serious scarcity of young professionals and scientists with understanding and experience in data assimilation. The lack of formal education in the field typically results in significant time spent learning on the job, both in research and operations, and the unavoidable limited scope of experience about different methods. There are examples of existing data assimilation education efforts at several universities: the Department of Meteorology

AFFILIATIONS: VUKICEVIC AND VONDER HAAR—Cooperative Institute for Research in the Atmosphere, Colorado State University, Ft. Collins, Colorado; KALNAY—Department of Meteorology, University of Maryland, College Park, Maryland
CORRESPONDING AUTHOR: Tomislava Vukicevic, 1395 Edinboro Dr., Boulder, CO 80305
E-mail: tomi@cira.colostate.edu
DOI: 10.1175/BAMS-85-1-48

at the University of Maryland, the Meteorology Department at The Florida State University, the School of Meteorology at the University of Oklahoma, the Program in Atmospheric and Ocean Sciences at the University of Colorado, the College of Oceanic and Atmospheric Sciences at The Ohio State University, and the Atmospheric Science Department at Colorado State University. These departments offer elective graduate-level courses that contain data assimilation topics, and although these courses are good, they do not reach out to the wider pool of future professionals at the B.S., M.S., and Ph.D. levels required to support demands of the research and operational institutions for expertise in data assimilation.

WHAT SHOULD A NATIONAL PROGRAM ON DATA ASSIMILATION EDUCATION INCLUDE? We would like to propose that the data assimilation education programs should provide a solid grounding in

- optimal control theory from applied mathematics and statistics programs;
- observing technology used in the atmospheric, ocean, and land system state assessment; and
- modeling principles of the Earth systems.

Furthermore, it should provide hands-on experience in applied data assimilation methods in problems of the atmospheric, ocean, and land system state assessment.

Such a national program is ambitious, but can be realized drawing largely on already available resources such as

- existing programs in applied mathematics and statistics for basic courses on state estimation theory; and
- existing programs on observations technology and dynamical systems modeling.

Most importantly, the program would benefit from sharing resources from several universities and research centers. This can be handled in the national program, within which groups of universities and operational/research centers would share research versions of end-to-end data assimilation systems that would be similar to current operational systems. For example, the Weather Research and Forecasting (WRF) model is the existing community resource that could be used as the basis for the end-to-end regional data assimilation training system. Also, the operational global systems from NOAA, NESDIS, NCEP, or NASA could be configured for the resources available for exploratory experiments and training in collaboration with the participating universities.

Such a cooperative basis for the national data assimilation program could cooperate closely with and complement the Joint Center for Satellite Data Assimilation, a virtual center between the NCEP and the NASA Data Assimilation Office, recently created with the purpose of fostering rapid progress in this area. The center plans to make available copies of relevant data assimilation and forecast systems to investigators but without formal or guaranteed support, and will support projects selected by relevance. The center does not include a major educational component, but the scientists involved are some of the best experts in the world. Its existence, and interest in advanced satellite data assimilation, will allow it to add to basic courses on applied data assimilation methods given by experts in the field, and to direct student research projects toward problems identified as important by the operational centers.

ECONOMIC VALUE AND FUNDING. Most observations cannot be applied in environmental prediction and state assessment without being subjected to some form of data assimilation, which maximizes information content of the observations relative to the desired state representation. To continue to improve the effectiveness of rather expensive observing networks, it is necessary to both refine ways of optimizing information content of the observations, and to objectively plan additional observations. These goals are in the interest of several national agencies, such as NASA, NOAA, the National Science Foundation, the Department of Defense, the Department of Energy, and the Environmental Protection Agency. Accelerating achievement of these goals requires support for the educational program. This program would be cost efficient because it would enhance and increase the efficiency of separate efforts taking place at present, at a relatively low cost.

FOR FURTHER READING

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, 350 pp.

Serafin, R. J., A. E. Macdonald, and R. L. Gall, 2002: Transition of Weather Research to Operations: Opportunities and Challenges. *Bull. Amer. Meteor. Soc.,* **83,** 377–392.

Talagrand, O., 1997: Assimilation of observations, an introduction. *J. Meteor. Soc. Japan,* **75**, 191-209.