Theory and Practice of Data Assimilation in Ocean Modeling

Robert N. Miller
College of Oceanic and Atmospheric Sciences
Oregon State University
Oceanography Admin. Bldg. 104
Corvallis, OR 97331-5503
phone: (541)737-4555 fax: (541)737-2064 email: miller@coas.oregonstate.edu

Grant Number: N0001492J1595
http://tangaroa.oce.orst.edu/group.html

LONG-TERM GOALS

The long range goal of this project is to combine computational models with observational data to form the best picture of the ocean as an evolving system, and use that picture to understand the physical influences which govern the ocean's behavior. Oceanic observations are sparse and models are limited in accuracy, but taken together, one can form a quantitative description of the state of the ocean that is superior to any based on either models or data alone. Along with the goals of analysis and prediction, we seek reliable estimates of the errors in our results. We expect our results to have implications beyond data assimilation. In particular, we hope this research will lead to enhanced understanding of the implications of nonlinearity and randomness for predictability of the ocean and atmosphere.

It is our long range goal to develop efficient methods for estimating useful statistical measures of errors in stochastic forecast models. Since the probability density functions (PDF's) of nonlinear stochastic models are not, in general, Gaussian, we must find methods for forecast evaluation based on information about the particular PDF generated by the model.

OBJECTIVES

The principle objective of this project is the development, implementation and validation of practical data assimilation methods for synoptic ocean models. By "data assimilation" we mean the construction of a composite estimate of the state of the ocean based on a combination of observed data with computational model output. Since data assimilation methods which give the most and best information are highly resource intensive, and often not practical for use with detailed models, we are particularly interested in the price paid in terms of accuracy and confidence for using economical but suboptimal data assimilation methods.

Direct calculation of full PDF's is not feasible for practical models of the ocean or atmosphere, but useful approximations to the PDF can be calculated from Monte-Carlo experiments, by virtue of the fact that the number of truly independent degrees of freedom in practical models is very much smaller than the dimension of the state vector. This intuition is the motivation for the ensemble methods that have become popular in recent years.

Our experience with Monte-Carlo methods in simplified systems has led us to investigate the details of methods for ensemble generation that have been presented in the community. The motivation for these

Report Documentation Page	Form Approved OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
30 SEP 2002		**00-00-2002 to 00-00-2002**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Theory and Practice of Data Assimilation in Ocean Modeling	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **College of Oceanic and Atmospheric Sciences,,Oregon State University,Oceanography Admin. Bldg. 104,,Corvallis,,OR, 97331**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The long range goal of this project is to combine computational models with observational data to form the best picture of the ocean as an evolving system, and use that picture to understand the physical influences which govern the ocean’s behavior. Oceanic observations are sparse and models are limited in accuracy, but taken together, one can form a quantitative description of the state of the ocean that is superior to any based on either models or data alone. Along with the goals of analysis and prediction, we seek reliable estimates of the errors in our results. We expect our results to have implications beyond data assimilation. In particular, we hope this research will lead to enhanced understanding of the implications of nonlinearity and randomness for predictability of the ocean and atmosphere.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**6**	

specialized methods for generating ensembles is precisely the specification of the PDF of a complex model whose behavior is believed to be captured by a relatively small number of independent degrees of freedom. By detailed study of the behavior of ensembles in increasingly complex models, we hope to gain the insights necessary to generate the most efficient ensembles, which should in turn lead to the error estimates necessary for data assimilation systems and prior estimates of forecast accuracy.

Optimized methods require accurate knowledge of the statistics of the errors in the model and the data. It is therefore an objective to understand in detail the sensitivity of the data assimilation scheme to the details of the defining error estimates.

APPROACH

The basic assumptions underlying data assimilation methods in use or proposed are known to be false to some degree. We plan to study the consequences of these assumptions by constructing a hierarchy of schemes with decreasing reliance on ad hoc assumptions. It is our guiding philosophy that the best way to learn how to design and implement the most economical methods that meet our needs is to begin by implementing methods which are as close to optimal as possible. From that point, we can quantify the loss of accuracy and the saving of resources associated with each simplification of the model or the data assimilation scheme.

Work is proceeding toward a theoretical basis for the next generation of data assimilation methods in which randomness and nonlinearity must be taken into account. To this end, we are applying tools from stochastic differential equations and from dynamical systems theory. Since our model systems are characterized by high dimensional state spaces, Monte Carlo methods must be used to study the behavior of the stochastic systems.

The theory of nonlinear filtering provides a framework in which problems of data assimilation with nonlinear models and non-Gaussian noise sources can be treated (see, e.g., Miller et al., 1999). In the case of linear models and Gaussian noise sources, this theory reduces to the familiar Kalman filter. In the formal theory of nonlinear filtering, the final result is not a single model state vector or trajectory in state space, but a PDF defined as a scalar function of the state variables and time. From this PDF, the mean, median, mode, or other statistic can be computed for use as the working estimate of the state of the system, along with the desired confidence intervals. The assignment of confidence limits corresponds in the case of a group of particles in physical space to drawing contours in the spatial domain which can be expected to define a region which contains, say, 90% of the particles.

The problem is that for even schematic models of the ocean or atmosphere, an unrealistically large number of particle trajectories in phase space must be calculated in order to represent the PDF faithfully. Useful ensemble analysis therefore requires judicious choice of ensemble members. We have concentrated our recent efforts on evaluation of ensemble methods, which we see as facilitating the generation of the forecast error estimates necessary for data assimilation. These forecast error estimates are of interest in and of themselves, since they have the potential of providing a priori estimates of the reliability of a given forecast.

With our work with a simple spectral approximation of a beta-plane channel with topography (Gravel and Derome, 1993) complete, we have begun work with more complex and relevant systems. Much of our work is focused on a model of the Kuroshio. This is a comparatively simple model which reproduces the observed bimodality. It operates on a state space with dimension two orders of magnitude greater than that of our beta-plane channel model, and, for this reason alone, presents significant technical challenges.

Once we are confident that we understand the qualitative behavior of the Kuroshio model, we will apply our stochastic techniques to it, in order to understand propagation of errors and the evolution of the PDF arising from random initial and boundary conditions. This should allow us to construct reliable data assimilation systems for use with simulated and real data from the Kuroshio.

Many different models, based on fundamentally different physical assumptions, exhibit the observed bimodality of the Kuroshio in some form. We are now in the process of comparing our model to different models and to observed data in order to determine a basis for distinction between the physical mechanisms in the different models.

We have also begun to investigate the generation of bred vectors in the context of the operational limited area model Aladin at Meteo-France. Through this work, we have gained valuable experience in the operational setting, and have begun to accumulate experience with complex limited area models. Dr. Francois Bouttier of Meteo-France worked with us on this phase of the project.

Professor Michael Ghil of UCLA and Professor James G. Richman of the College of Oceanic and Atmospheric Sciences at Oregon State are working with us on the dynamical analysis of the Kuroshio. Professor Ghil and his group are working with analysis of simplified basin scale models of the north Pacific and with analysis of modeled and observed time series. Professor Richman is also working on analysis of time series, and on comparison of ocean general circulation model results to data and to other models. Technical support for this project is provided by Ms. Laura Ehret.

WORK COMPLETED

An article containing a detailed description of our results of ensemble generation experiments with a simple beta-plane channel model has appeared in the Monthly Weather Review (Miller and Ehret, 2002).

RESULTS

In first experiments with the operational limited area forecast model Aladin at Meteo-France, we have learned the importance of using the correct normalization factors for generation of ensembles by the breeding method. Figures 1 and 2 show two predicted cumulative convective rainfall predictions generated by Aladin for the 6 hour period ending 0000Z, 31 August, 2001. The operational forecast for that time is depicted in figure 1. Figure 2 shows the results of the fourth 6 hour cycle following the imposition of a small perturbation in the initial condition at 0000Z, 30 August, 2001; the forecasts are otherwise identical. The mean square difference between the two runs decays steadily, and one might be led to conclude that the effect of the perturbation is insignificant, but, as the figures show, the rainfall patterns are distinctly different 24 hours into the experiment. Evidently the simple mean square differences are not the right normalization criteria for generation of bred vectors with this model. The relevant dynamics are quite significantly nonlinear.

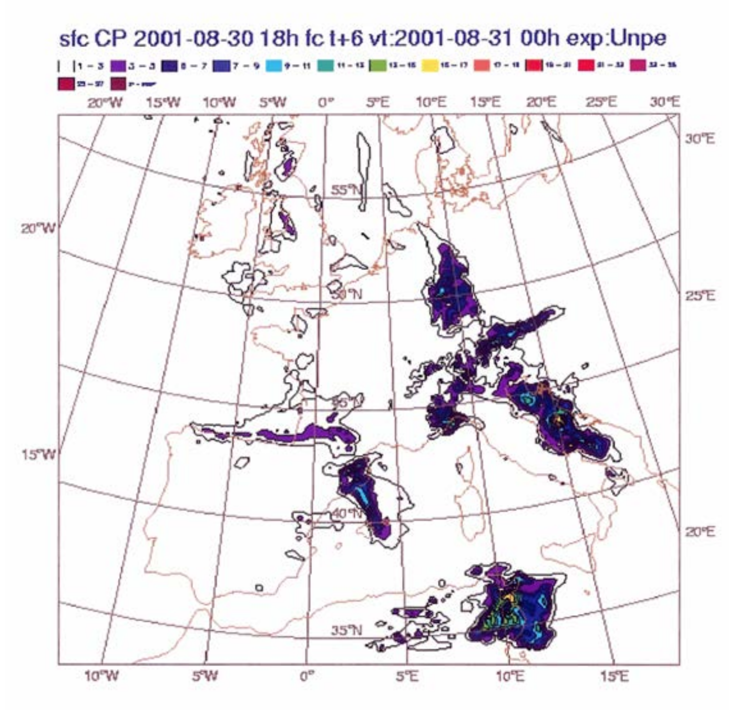

Figure 1. Predicted 6 hour accumulation of convective rainfall for 0000Z, 31 Aug, 2001. Operational forecast by Meteo-France

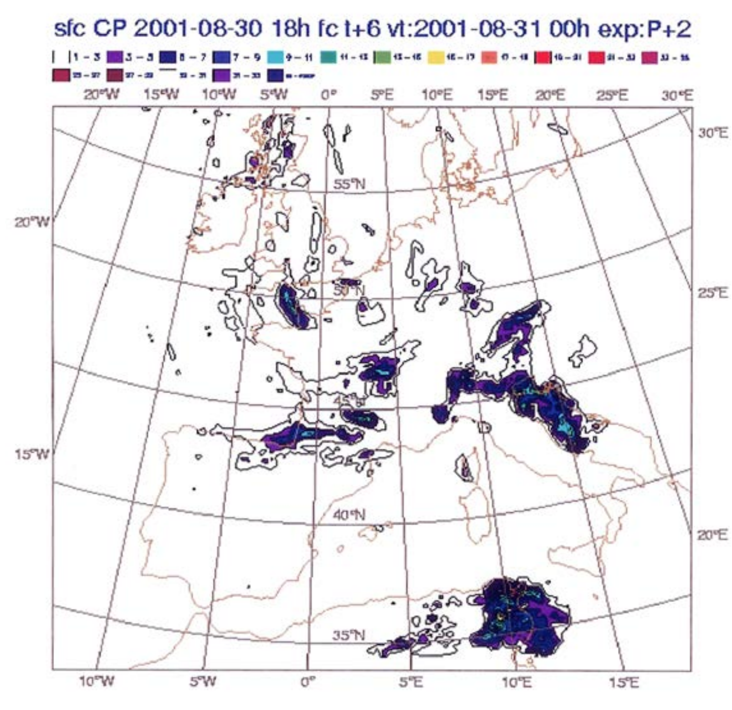

Figure 2. Similar to figure 1, but for slightly perturbed initial condition at 0000Z, 30 August, 2001. Note the differences in rainfall patterns, despite the small size of the initial perturbation, and the decay of the mean square difference between the operational and perturbed forecasts over four 6 hour prediction cycles.

IMPACT/APPLICATIONS

Major weather centers, including the US National Center for Environmental Prediction (NCEP) and the European Center for Medium-Range Weather Forecasting (ECMWF) now use ensemble methods for operational forecast validation; see Molteni et al. (1996), Toth and Kalnay (1993). Our work on Monte-Carlo methods should provide enhanced capability for validation of forecasts of the ocean and atmosphere, in addition to application to data assimilation. Our work on breeding modes and planned work on other schemes for ensemble generation should provide significant guidance in optimizing methods for generation of ensembles. Our work with the French limited area model should form a part of the framework for predictability studies in limited area models.

We expect that our work with comparisons among models and data for the Kuroshio will lead to greater insight into the intrinsic variability of basin-scale ocean circulation.

TRANSITIONS

RELATED PROJECTS

"Assimilation of Coastal Radar Surface Current Measurements in Shelf Circulation Models." Work is in progress on the investigation of data assimilation systems for use with surface velocity data from coastal radar. This project is in collaboration with Professor John Allen.

"Uncertainties and Interdisciplinary Transfers through the End-to-End System (UNITES)." This is a cooperative project under the "Capturing Uncertainty in the Common Tactical Environment" DRI. The leaders of the UNITES team are Philip Abbot of OASIS, Inc. and Professor Allan Robinson of Harvard University.

REFERENCES

Gravel, S. and J. Derome, 1993: A study of multiple equilibria in a beta plane and hemispheric model of a barotropic atmosphere. *Tellus*, **45A**, 81-98.

Kawabe, M., 1995: Variations of current path, velocity and volume transport of the Kuroshio in relation with the large meander. *J. Phys. Oceanogr*., **25**, 3103-3117.

Miller, R. N., E. F. Carter and S. T. Blue, 1999: Data assimilation into nonlinear stochastic models. *Tellus*, **51A**, 167-194.

Molteni, F., R. Buizza, T. N. Palmer and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Q. J. R. Meteorol. Soc*., **122**, 73-119.

PUBLICATIONS

Miller, R. N. and L. L. Ehret, 2001: Ensemble generation for models of multi-modal systems. *Mon. Wea. Rev*., **130**, 2313-2333.