The CPTEC Ocean Data Assimilation System - CODAS

Matthew Hoffman[1], Luciano Pezzi[2], Dirceu Herdies[2], and Steven Penny[1]
[1]University of Maryland - USA
[2]National Institute for Space Research (INPE) – Brazil

ABSTRACT: This work shows very preliminary results achieved by the CPTEC Ocean Data Assimilation System - CODAS. This scheme is based on the Ensemble Transform Kalman Filter (LETKF) method. The first investigated area is the Brazil-Malvinas confluence. The analyses from the CODAS ensemble experiments using both 8 and 12 members exhibit all of the major oceanic features of the truth state. However, using twelve ensemble members, an excellent analysis, which nearly mirrors the shape of the truth, is achieved.

RESUMO: Este trabalho mostra os primeiros resultados obtidos com o esquema de assimilação de dados oceânicos do CPTEC denominados CODAS. Este esquema baseia-se na metodologia do Ensemble Transform Kalman Filter (LETKF). A primeira área analisada foi a Região da Confluência Brasil-Malvinas. Os experimentos realizados com o CODAS, tanto com 8 quanto com 12 membros do conjunto, mostraram as principais características da circulação e da termodinâmica desta região. O CODAS produziu análises que são muito parecidas com os campos que são considerados a verdade.

Key words: Assimilação de dados, LETKF, Brasil-Malvinas, Filtro de Kalman, modelagem oceânica.

1 Introduction

While data assimilation has been performed on ocean models for many years, the schemes used have usually not been as sophisticated as those used on their atmospheric counterparts. Among the reasons that ocean data assimilation has lagged behind that of the atmosphere are that observational data is much sparser in the ocean, the ocean is, in general, more stable than the atmosphere over similar time scales, and that weather prediction has a much more visible daily human impact. For these reasons, implementing an operational oceanic data assimilation scheme is still a big challenge for many operational weather and climate forecast centers around the world. This is the case at the Center for Weather Forecast and Climate Studies (CPTEC) where, up to now, the center does not have an oceanic data assimilation scheme. This is a work in progress which intends to show the very first results of the CPTEC Ocean Data Assimilation System (CODAS).

2 Methodology

Despite the problems described above, accurate forecasts can be made over reasonable time intervals and are very valuable. Improvements to forecasting skill can be made in a number of ways and serve to both improve the forecast on a particular day and to increase the length of usable forecasts. Data assimilation methods, such as the Local Ensemble Transform Kalman Filter (LETKF), attempt to improve forecasts by improving the accuracy of the current state estimate (which is the initial condition) (Hunt, 2007, Houtekamer, 1998). In a data assimilation scheme, an estimate of the current state is derived by combining current observations and a previous forecast, which is referred to as the background. This state estimate, hereafter called the "analysis", is then used as the initial condition for the model, which, in turn, creates a new forecast. Data assimilation proceeds in this iterative manner, alternating between a forecast step, where the model predicts the future state of the system, and an analysis step, where observations taken at this future time are incorporated and the analysis is created.

Both the background and the observations have errors and the analysis step consists of a statistical procedure that takes these errors into account in determining the analysis state. In the LETKF, developed at the University of Maryland, this synthesis of background and observations is accomplished using a maximum likelihood estimate (Hunt,2007). Since the estimate takes into account the observations and the background state in addition to the relative covariances of each state, approximations of the covariances must be derived. In most of the currently used data assimilation techniques, such as 3D-VAR and 4D-VAR, the background error covariance is assumed to be constant in time and is approximated using a climatological average. While this is a reasonable approximation, it does not account for the day-to-day variations in the background error covariance that naturally occurs. In contrast, the LETKF and other ensemble filters estimate the background covariance using the sample covariance of an ensemble of forecast states (Hunt, 2007, Houtekamer, 1998, Whitaker, 2002). One of the significant advantages of the ensemble methods over variational schemes is that ensemble methods account for "errors of the day" much more effectively by allowing the background covariance to change at each step (Kalnay, 2003, Hunt, 2007).

In this work, the LETKF framework is applied to a global implementation of the Modular Ocean Model, version 4 (MOM4). Details of the model numerics and physics can be found in Griffies et al (2004). The ocean model used here has a horizontal resolution of approximately 1° x 1°and 50 vertical levels, with 30 of them confined in the first 1000 meters. The ocean model was spun-up for 10 years using climatological fluxes of momentum, heat and water to generate the initial restart used.

The LETKF code being used was written by Eric Kostelich of Arizona State University and was originally modified for the ocean by Ross Hoffman of AER Inc. (Hoffman, 2008). While originally written for an atmospheric model, the atmospheric assimilation variables of temperature, humidity, and zonal and meridional wind variables map one to one onto the oceanic variables of temperature, salinity, and zonal and meridional currents [2]. In addition, surface pressure in the atmosphere can be mapped to sea surface height. To test the feasibility of the LETKF on MOM4, we began by running identical twin experiments where the truth is given by a model run. In our initial experiments, the observations are simulated by sampling the truth and adding random errors based on a prescribed error covariance.

Preliminary Results

In perfect model experiments, the global RMS error is greatly reduced by the data assimilation. Here the RMS error is the difference between the given field and the "truth," which is known because we are assuming a perfect model. Model initialization is accomplished using model states from previous months. Analyses were performed every day using observations that were simulated in random locations representing a specified percentage of the entire grid. In experiments using 10% data coverage, the LETKF quickly reduces both the analysis and forecast errors below the specified observational errors, which are 0.5°C, 0.08psu, and 0.04m/s for zonal velocity and 0.02m/s for meridional velocity. This reduction below the observational errors is observed in a few days for a four member ensemble and in one day for a twelve member ensemble. Moreover, the errors remain below the observational errors for the duration of the simulation. At 1% data coverage, the analysis error converges slower than at 10% coverage, but the analysis error still drops below the observational error in couple of days [need to check] using a twelve member ensemble.

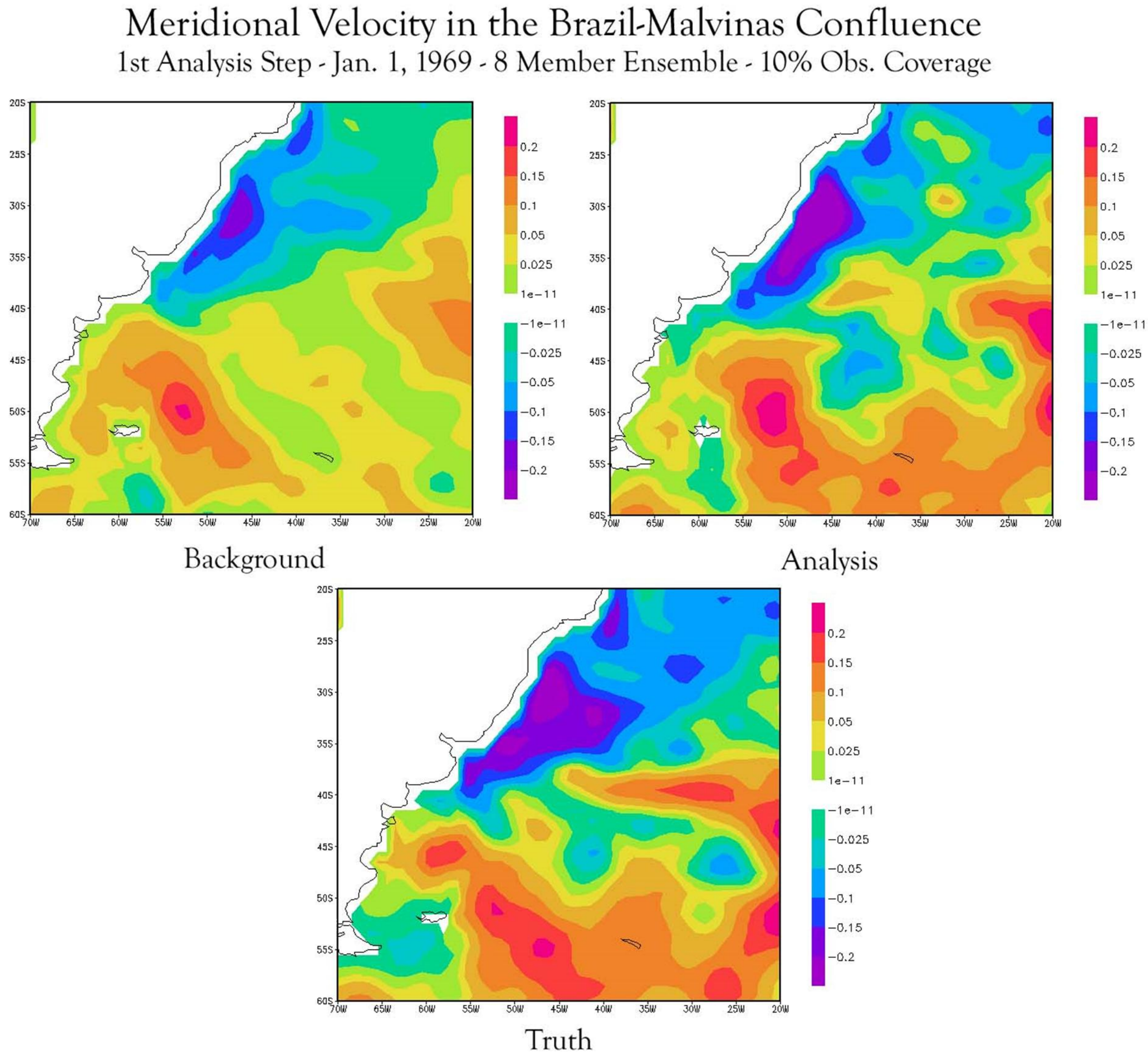

Figure 1- Meridional velocity at Brazil-Malvinas Confluence. Panels showing the Background state, Analysis and Truth from 8 members simulations. All simulations refer to a climatological January.

In addition to the global improvement, local improvements can also be seen. The LETKF not only reduces the global error, but it allows the analysis to capture local shapes as well. The

Figures 1 and 2 show the analysis, background, and truth of meridional sea surface currents in the Brazil-Malvinas confluence.

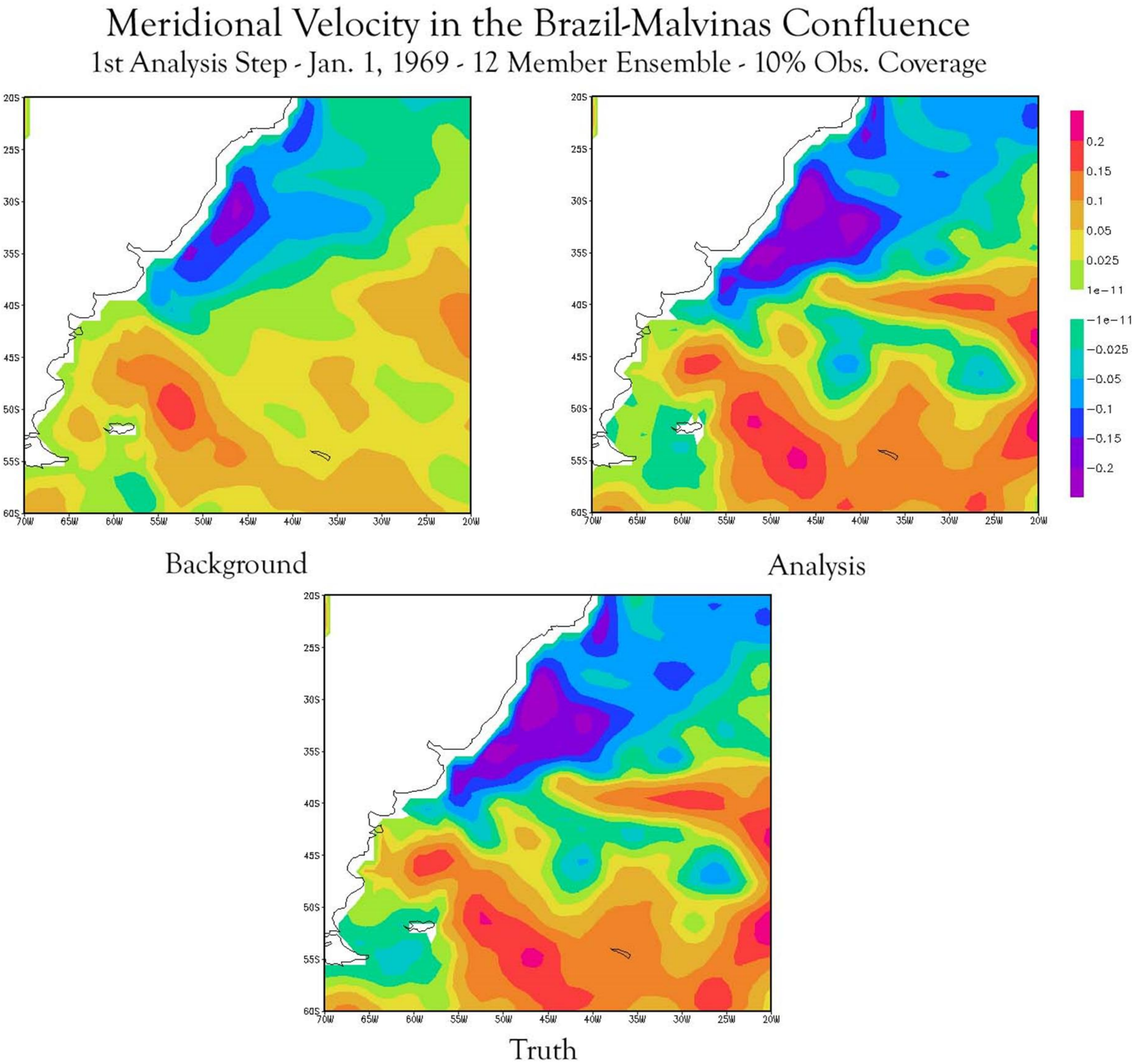

Figure 2- Meridional velocity at Brazil-Malvinas Confluence. Panels showing the Background state, Analysis and Truth from 12 members simulations. All simulations refer to a climatological January.

In the first analysis step, the improvement from the LETKF is evident even with only eight ensemble members. The truth shows a strong, narrow tongue of northward velocity around 40°S latitude, while the background shows only weak, more dispersed velocity in the same area. After one LETKF step, the analysis exhibits a narrower area of strong velocity around 40°S that more closely resembles the truth. There is also increased southward velocity right off the coast between 30°S and 40°S which is in agreement with the truth. Moreover, after a few LETKF steps, the analysis very accurately represents the shape of the meridional velocity field in the Brazil-Malvinas confluence. Even using only four ensemble members, the analysis after seven steps shows all of the major features of the truth. Using twelve ensemble members, an excellent analysis, which nearly mirrors the shape of the truth, is achieved. Similar results are seen in all of the analysis fields.

Final Remarks

This work shows the very first CODAS results, which is an implementation of an oceanic data assimilation framework based on the LETKF at CPTEC. Preliminary results show that, even for relatively small ensembles, the LETKF reduces global errors and also accurately recreates the shape of the true ocean state in the Brazil-Malvinas confluence region. In the first stage of this initiative, the implementation is being conducted as a research topic. In the next step, however, CODAS will be implemented as an operational routine at CPTEC.

Acknowledgments

We would like to thank the Brazilian Ministry of Science and Technology (MCT) and National Brazilian Research Council (CNPq) for funding Mr. Matthew Hoffman's visit to CPTEC. We'd also like to thank Dr. Eugenia Kalnay and Dr. James A. Carton for their guidance.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly Weather Review*, **129**, 420-436.

Griffies, S.M., M.J. Harrison, R.C. Pacnowsky and A. Rosati., 2004 A Technical Guide to MOM4. GFDL Ocean Group Technical Report No. 5. NOAA/Geophysical Fluid Dynamics Laboratory.

Hoffman, R. N., R. M. Ponte, E. J. Kostelich, A. Blumberg, and I. Szunyogh, S. V. Vinogradov, J. M. Henderson 2008: A Simulation Study Using a Local Ensemble Transform Kalman Filter for Data Assimilation in New York Harbor. *JAOTO*, Accepted.

Houtekamer, P.L. and H. L. Mitchell, 1998: Data assimilation using an Ensemble Kalman Filter Technique, 1998. *Monthly Weather Review*, **126**, 796-811.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter. *Physica D*

Kalnay, E.: Atmospheric Modelling, Data Assimilation and Predictability, Cambridge University Press, 2003.

Kalnay, E., H. Li, T. Miyoshi, S.C. Yang, and J. Ballabrera-Poy, 2007: 4D-Var or Ensemble Kalman Filter. *Tellus*

Ott, E., B.R. Hunt, I. Szunyogh, A.V. Zimin, E.J. Kostelich, M. Kostelich, M. Corazza, T. Sauer, E. Kalnay, D.J.

Patil, and J.A. Yorke, 2004: A Local Ensemble Kalman Filter for Atmospheric Data Assimilation. *Tellus*, **Vol. 56A**, 415-428.

Patil, D. J., B.R. Hunt, E. Kalnay, J.A. Yorke, and E. Ott, 2001: Local Low Dimensionality of Atmospheric Dynamics. *Physical Review Letters*, **86**(26), 5878-5881.

Szunyogh, I., E.J. Kostelich, G. Gyarmati, D.J. Patil, B.R. Hunt,E. Kalnay, E. Ott and J.A. Yorke, 2005: Assessing a Local Ensemble Kalman Filter: Perfect Model Experiments with the NCEP Global Model. *Tellus*, **Vol. 57A**, 528-545.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913-1924.