The GMAO's Ensemble Kalman Filter Ocean Data Assimilation System

Michele M. Rienecker[1], Christian L. Keppenne[1,2], Robin Kovach[1,2], Jossy P. Jacob[1,2], Jelena Marshak[1]

[1]Global Modeling and Assimilation Office, NASA/Goddard Space Flight Center, USA

[2]Science Applications International Corporation, USA

Correspondence: Michele.Rienecker@nasa.gov

INTRODUCTION

The ensemble Kalman filter (EnKF) was introduced to the ocean and atmospheric assimilation communities by Evensen (1994). The main attractiveness of ensemble techniques stems from their accounting for the temporal evolution of background-error covariances at small computational cost and from the ease of their implementation for complex systems. Ensemble filters have now been implemented in several real applications with state-of the art ocean (e.g., Keppenne et al., 2005; Zhang et al., 2007), atmosphere (e.g., Whitaker and Hamill, 2002; Hunt et al., 2007; Houtekamer and Mitchell, 2005) and land surface (e.g., Reichle et al., 2007) models. The Global Modeling and Assimilation Office (GMAO) developed an EnKF ocean data assimilation system to initialize a coupled model for seasonal forecasts (e.g., Keppenne et al., 2002, 2005). This paper describes the current system and its performance relative to that for the GMAO system based on optimal interpolation.

GMAO'S OCEAN DATA ASSIMILATION SYSTEM

The GMAO's ocean data assimilation system version 1 (ODAS-1) uses the quasi-isopycnal Poseidon model version 4 (e.g., Schopf and Loughe 1995) with options of a univariate optimal interpolation (UOI), a multivariate optimal interpolation (MvOI), or a multivariate EnKF. The model and the details for the ODAS-1 are summarized in Keppenne et al. (2005 and 2008). The prognostic model variables are layer thickness, h, temperature, T, salinity, S, and the zonal and meridional current components, u and v. In Poseidon V4, the sea surface height (SSH) field is diagnostic. The horizontal resolution is 1/3º meridionally by 5/8º zonally and there are 27 vertical layers. The domain is almost global, from Antarctica to 72ºN.

Both UOI and EnKF analyses are performed on 27 levels limited to the upper 1000m. The assimilation interval is 4 days for the EnKF and 5 days for the UOI. The EnKF calculates the background error covariances from the ensemble. The UOI employs time-invariant Gaussian covariances to produce separate temperature and salinity analyses. The horizontal covariance scales are reduced with latitude. Along the equator, the zonal scale is 20° (8°) and the meridional scale is 4° (3°) for temperature (salinity). The vertical scale is 100m (50m) for temperature (salinity). The prescribed background error variance is $(0.7°C)^2$ for temperature and $(0.01\ psu)^2$ for salinity.

The EnKF has been tested with 9-, 17-, 33- and 65-member ensembles. Keppenne et al. (2008) presents the details of the implementation and the spatial and temporal smoothing employed along with covariance localization to compensate for small ensemble sizes and shows that effective assimilation can be conducted with only 17 ensemble members. The EnKF assimilates SSH anomalies from the Topex/Poseidon and Jason altimeters. To avoid contamination of the anomaly signal by differences in the means used to calculate the anomalies, an online bias correction has been implemented (Keppenne et al., 2005). The UOI does not assimilate satellite altimetry data.

An important element of the application of the EnKF in ocean assimilation is the inclusion of a system noise model. This system noise represents both model error and forcing error, the latter being one of the largest sources of error for ocean state estimation. The inclusion of a system noise model obviates the need for covariance inflation as is common for ensemble methods employed for the atmosphere (e.g., Whitaker and Hamill, 2002). The GMAO's implementation has three algorithms for system noise: (i) chaotic perturbations to the ocean mixing and diffusion, (ii) perturbed forcing with realistic spatial perturbations from an ensemble of atmospheric integrations, and (iii) internal perturbations from a pre-computed ensemble of ocean integrations. The impact of the perturbed forcing is felt primarily in the upper ocean and is limited by the assimilation interval. The inclusion of the internal perturbations is meant to represent the uncertainty in the deeper ocean and is effectively additive covariance inflation.

In addition to the SSH data assimilated by the EnKF, the input data streams are temperature profiles from the TAO and PIRATA mooring arrays and from XBTs, and temperature and salinity profiles from Argo floats. Synthetic salinity profiles, generated using the observed temperature profiles and the local T-S climatology, are also assimilated to maintain water masses and prevent a vertical circulation that is introduced if only temperature is corrected in the equatorial oceans (e.g., Sun et al., 2007). Since the multivariate covariances of the EnKF provides a mechanism for correcting salinity from temperature-only observations or from the SSH observations, the EnKF experiments are conducted without these synthetic observations.

The observation-error covariances are approximately Gaussian. For the altimetry data, an observational error-variance of 2 cm^2 is assumed and the data-error covariances vanish at 10°. For the temperature (salinity) profile observations, the assumed error-variance is 0.25°C^2 (0.0001psu^2); the covariances are spatially white in the horizontal and vanish at a vertical spacing of 500m. The synthetic salinity is given an arbitrary variance 4 times that of actual data so that information from real observations is not over-ridden by synthetic information.

EXPERIMENTS

The assimilation experiments presented here are conducted as ocean-only runs forced with daily surface wind stress derived from SSM/I and QuikSCAT, GPCP monthly mean precipitation, NCEP CDAS1 shortwave radiation (for penetrating radiation) and latent heat flux (for evaporation). Surface heat fluxes are provided by relaxation to Reynolds and Smith (1994) weekly SST. A relaxation to sea surface salinity climatology is used to compensate for biases in the freshwater flux and the omission of river runoff. The experiments, used as the basis for historical re-forecasts needed to calculate climatological forecasts drifts, are conducted from 1993 to 2006. The EnKF experiments are conducted both with SSH data and without.

PERFORMANCE

The performance of both versions of the ODAS is assessed by a comparison with the T and S profiles from CTD data and zonal current profiles from ADCP data collected on TAO servicing cruises from 1993 to 1998 (e.g., Johnson et al., 2000). These data have not been included in the assimilation. The comparison is divided into four separate regions: the Niño-3 region (150°W-90°W) with the sub-domain north (0-5°N) of the equator considered separately from that south (5°S-0) of the equator, and similarly for the Niño-4 region (160°E-150°W).

The comparison in Figure 1 shows that both assimilation schemes improve on the control temperature and salinity. The EnKF improves slightly on the UOI salinity in the near-surface Niño-4 region while the UOI is slightly better in the near-surface northern Niño-3 region. The EnKF improves the zonal currents in the upper 400m of the Niño-4 region. In the Niño-3 region, both zonal current analyses tend to be comparable to or slightly worse than the control.

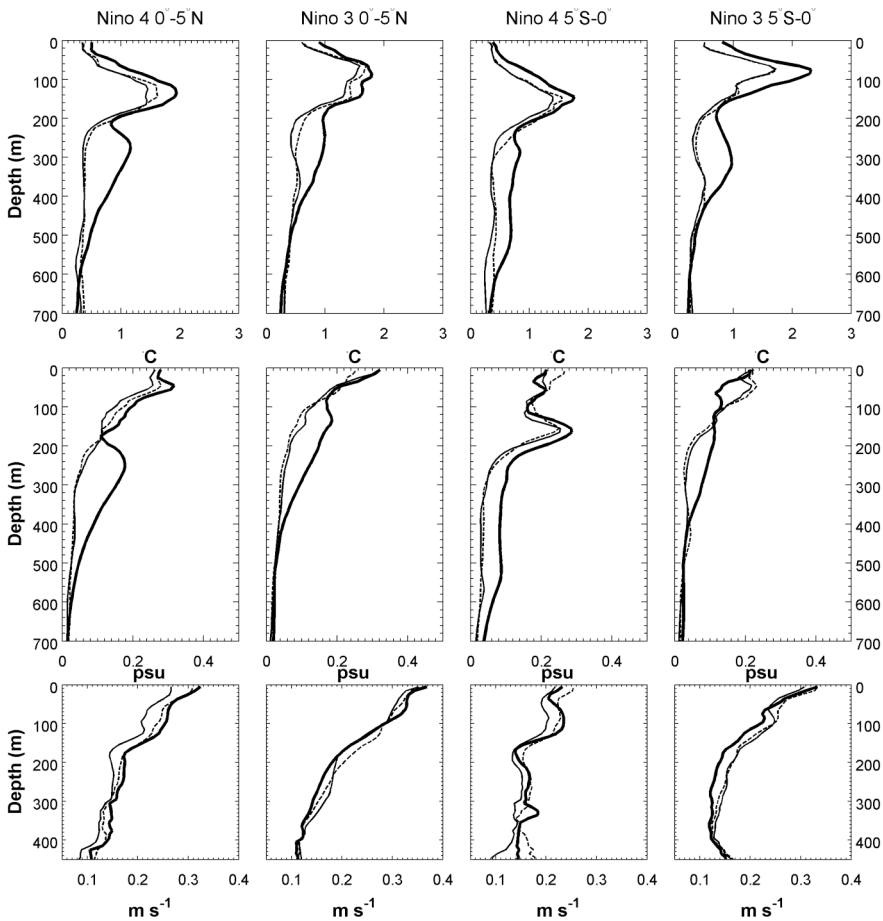

Figure 1 Comparison of UOI (dashed line) and EnKF (thin solid line) profiles of temperature (top panels), salinity (middle panels) and zonal current (bottom panels) with data from TAO servicing cruises, 1993 to 1998. Also shown is a control simulation without assimilation (thick solid line).

The performance can also be assessed through a comparison of the root mean square (RMS) innovation (observation-minus-forecast) statistics. These are shown averaged zonally and over the upper 300m for the tropical oceans in Figure 2. It is apparent that the EnKF mostly outperforms the UOI although the performance is not as good in the subtropical Atlantic.

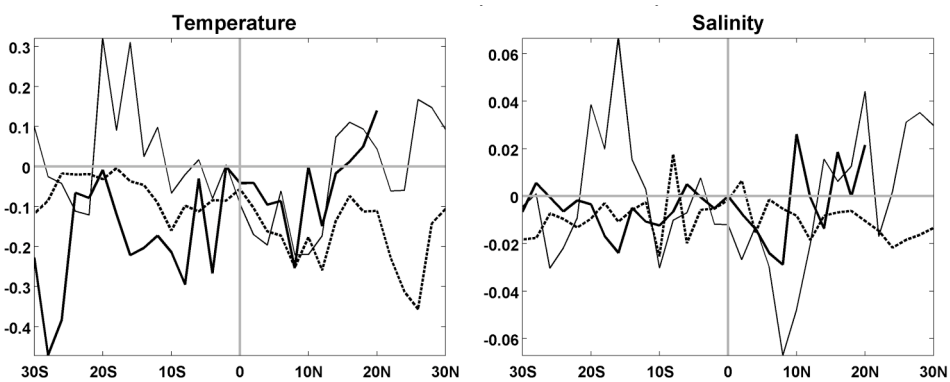

Figure 2 The difference between the RMS of innovations, averaged zonally and over the upper 300m. Temperature is in ºC, salinity in psu. The thick dashed line is for the Pacific Ocean, the thin solid line for the Atlantic Ocean and the thick solid line for the Indian ocean. Positive (negative) values indicate that the UOI (EnKF) is better.

Salinity variability can play an important role in interannual climate variations in the equatorial Pacific. Maes et al. (2002) showed that salinity accounts for as much as 30% of the dynamic height variability in the western tropical Pacific. Maes et al. (2006) suggests that the intensity of SST-wind coupling is mediated by the presence of salinity barrier layers. Yang et al. (2008) show evidence that the salinity variability near the 24.5 kgm^{-3} sigma-theta surface impacted the growth of forecast Niño-3 SST anomalies during 2006.

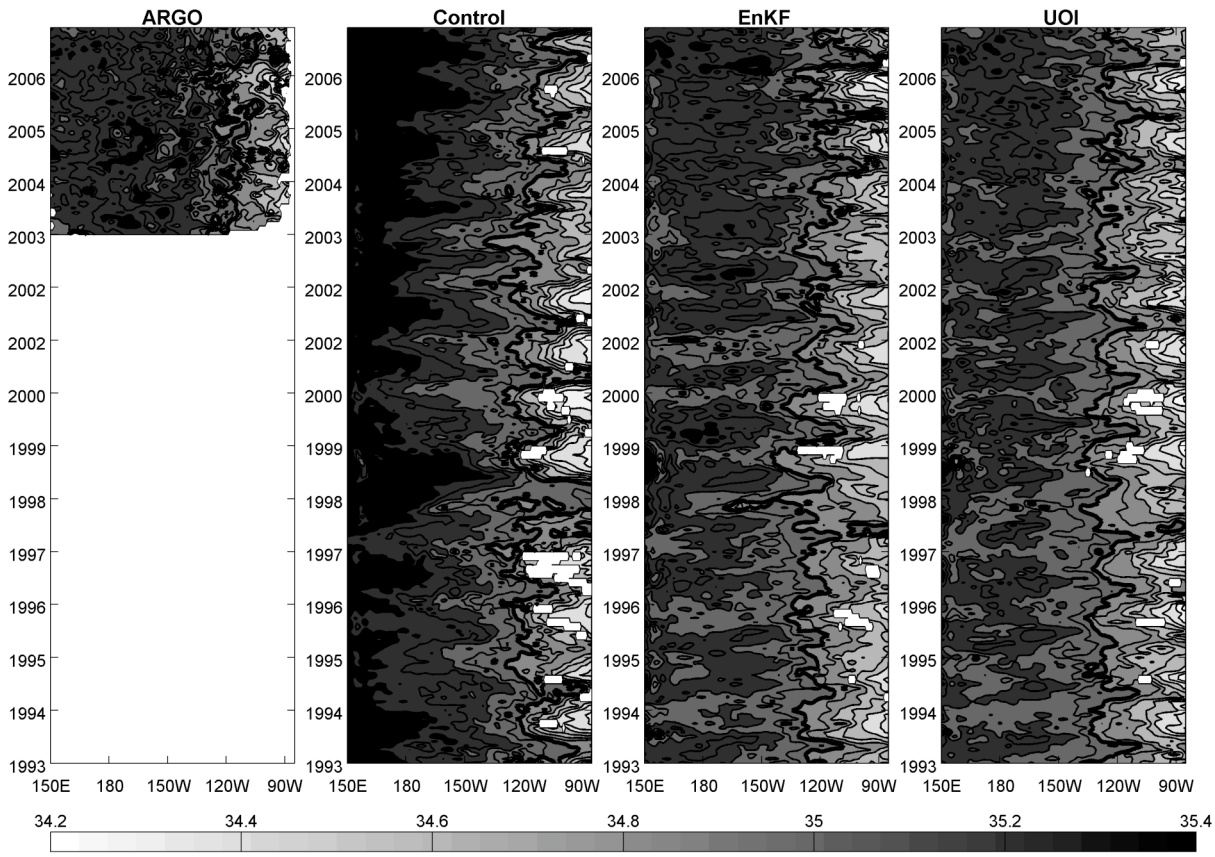

Figure 3 The time series of salinity on the 24.5 kgm^{-3} sigma-theta surface from Argo, the control (no assimilation), the EnKF, and the UOI. The 35.0 isoline is thickened.

The variability in salinity analyses along the equatorial ocean on the 24.5 kgm^{-3} sigma-theta surface is shown in Figure 3. The location of this surface is in the upper thermocline, through the salty intrusion, and in the core of the equatorial undercurrent in the western basin. The variability in high salinity in the control stems primarily from the western boundary. In the analyses it comes instead from variations in the central Pacific as is also apparent in the Argo time series. From the time series, it appears that the EnKF salinity analysis prior to the availability of Argo is well-behaved without the synthetic salinity. Apparently the multivariate corrections to salinity are useful. The UOI has more regular fresh salinity variations near the eastern boundary. The variations during the Argo period are more similar. The salinity in the western-central basin is saltier in the later years in the control and in the analyses.

IMPACT ON SEASONAL FORECASTS

The impact of the ocean analyses on oceanic seasonal forecast skill in the current GMAO coupled model forecasts (e.g., Vintzileos et al., 2004) have been assessed for the EnKF with SSH assimilation, and the EnKF and UOI without SSH assimilation. As for other coupled models, the forecast skill varies seasonally. It is difficult to discern significant differences in skill from the different ocean initializations for January starts. The skill for July starts is longer-lived and there are discernable differences between the performance for different ocean

initializations. The skill for SST and upper-ocean heat content in the tropical oceans is shown in Figures 4 and 5, respectively. The SST forecasts show some improvements from the use of the EnKF to initialize the ocean. The upper ocean heat content forecasts show additional improvements from the assimilation of satellite altimetry.

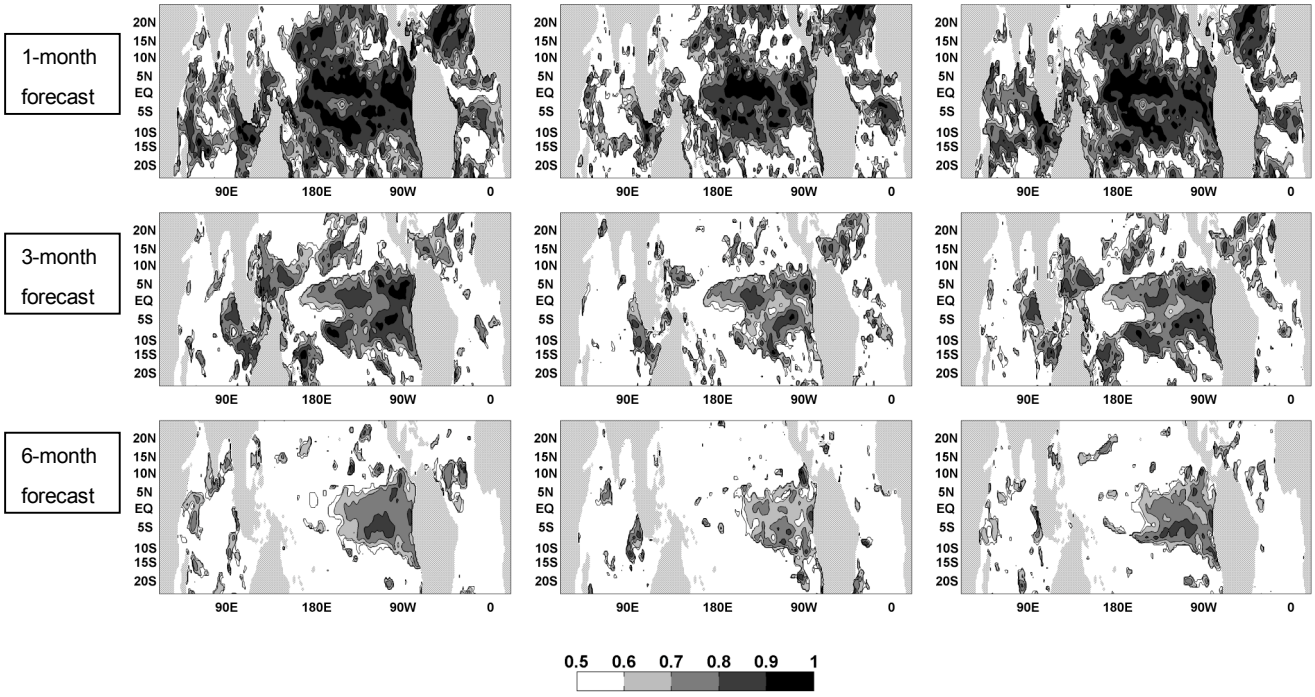

Figure 4 The anomaly correlation skill score for SST from the GMAO CGCMv1 for 1-, 3-, and 6-month forecasts from 1 July initial conditions. The ocean is initialized from the EnKF assimilating T, S, and SSH (left-hand column), from the EnKF without SSH (middle column), and from the OI (right-hand column). Only correlations higher than 0.6 are shown.

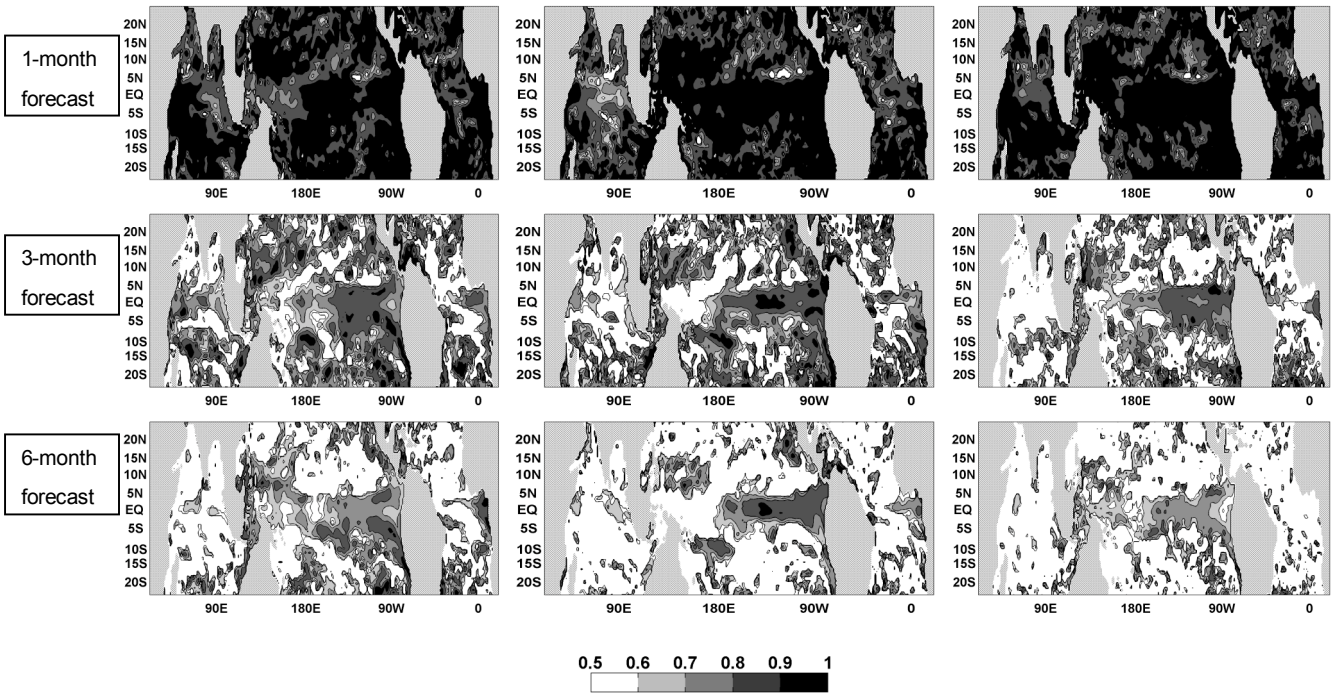

Figure 5 As for Figure 4, but for heat content in the upper 300m.

SUMMARY AND FUTURE DEVELOPMENTS

The evaluation of the GMAO assimilation systems to date show that the multivariate EnKF improves the ocean state analysis compared to that using independent, univariate UOI for temperature and salinity. The EnKF also improves the 6-month forecast skill score for forecasts initialized in July.

The next GMAO ODAS, version 2, has been developed for MOM4 and is implemented under the Earth System Modeling Framework (ESMF) used for GEOS-5. This system will be used to interface with the MERRA atmospheric analyses under a coupled model implementation. The goal is to reduce initialization shocks encountered in coupled model forecasts when the ocean and atmosphere are initialized independently.

REFERENCES

Evensen, G. 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res., C**99**, 10,143-10,162.

Houtekamer, P.L., and H.L. Mitchell 2005: Ensemble Kalman filtering. Quart. J. Roy. Meteor. Soc., **131C**, 3269-3289.

Hunt, B. R., E. Kostelich, I. Szunyogh 2007: Efficient data assimilation for spatiotemporal chaos: a Local Ensemble Transform Kalman Filter. Physica D, **230**, 112-126.

Johnson, G.C., M.J. McPhaden, G.D. Rowe, and K.E. McTaggart 2000: Upper equatorial Pacific ocean current and salinity variability during the 1996-1998 El Niño-La Niña cycle. J. Geophys. Res., **105**, 1037-1053.

Keppenne, CL., and M. Rienecker 2002: Development and initial testing of a parallel Ensemble Kalman Filter for the Poseidon isopycnal ocean general circulation model. Mon. Wea. Rev., **130**, 2951-2965.

Keppenne, CL., M. Rienecker, N.P. Kurkowski, and D.A. Adamec 2005: Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. Nonlinear Processes in Geophysics, **12**, 491-503.

Keppenne, C.L., M.M. Rienecker, J.P. Jacob, and R. Kovach 2008: Error Covariance Modeling in the GMAO Ocean Assimilation Ensemble Kalman Filter. Mon. Wea. Rev., (in press).

Maes, C., M.J. McPhaden, and D. Behringer 2002a: Signatures of salinity variability in tropical Pacific Ocean dynamic height anomalies. J. Geophys. Res., **107**, 8012, doi:10.1029/2000JC000737.

Maes, C., K. Ando, T. Delcroix, W.S. Kessler, M.J. McPhaden, and D. Roemmich, 2006: Observed correlation of surface salinity, temperature and barrier layer at the eastern edge of the western Pacific warm pool. Geophys. Res. Lett., **33**, L06601, doi:10.1029/2005GL024772.

Reichle, R.H., R.D. Koster, P. Liu, S.P.P. Mahanama, E.G. Njoku, and M. Owe 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). J. Geophys. Res., **112**, D09108, doi:10.1029/2006JD008033.

Reynolds, R. W. and T. M. Smith 1994: Improved global sea surface temperature analyses using optimum interpolation. J. Clim., *7,* 929-948.

Schopf, P.S., and A. Loughe 1995: A reduced-gravity isopycnic ocean model–hindcasts of El-Niño. Mon. Wea. Rev., **123**, 2839-2863.

Sun, C., M.M. Rienecker, A. Rosati, M. Harrison, A. Wittenberg, C.L. Keppenne, J.P. Jacob, and R.M. Kovach 2007: Comparison and sensitivity of ODASI ocean analyses in the tropical Pacific. Mon. Wea. Rev., **135**, 2242–2264.

Vintzileos, A., M.M. Rienecker, M.J. Suarez, S.K. Miller, P.J. Pegion, and J.T. Bacmeister 2003: Simulation of the El Niño – Southern Oscillation phenomenon with NASA's Seasonal-to-Interannual Prediction Project coupled general circulation model. CLIVAR Exchanges*,* **8**, 25-27.

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

Yang, S.-C., M.M. Rienecker, C.L. Keppenne, and E. Kalnay 2008: Impact on the seasonal-to-interannual forecasting from the ocean data assimilation: A case study of the 2006 El Niño event (draft ms).

Zhang, S., M.J. Harrison, A. Rosati, and A. Wittenberg 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. Mon. Wea. Rev., **135**, 3541-3564.