MULTIVARIATE ASSIMILATION OF SEA LEVEL, SST, AND CURRENTS FOR THE TROPICAL PACIFIC

EARTH SYSTEM SCIENCE INTERDISCIPLINARY CENTER ESSIC

E.C. Hackert[1], J. Ballabrera[1], A.J. Busalacchi[1], and L. Gourdeau[2]

[1]*ESSIC, University of Maryland, College Park, Maryland* [2]*LEGOS/GRGS, Toulouse France*

INTRODUCTION

Over the past number of years great strides have been made in the field of ocean data assimilation. New data types such as those provided by the extended altimetry missions (eg. TOPEX/Poseidon and Jason 1) now provide ocean modelers with enough high-quality data to allow data assimilation on basin-wide scales. In the current study, a Reduced Order Kalman Filter (ROKF) approach is used to assimilate various data types into a general circulation model of the tropical Pacific Ocean. The ocean model is a high-resolution, reduced gravity, primitive equation, sigma-coordinate model with variable depth mixed layer (Gent and Cane, 1989). Consistent with previous studies, data include sea surface height (SL) from TOPEX/Poseidon altimeter and sea surface temperature (SST) analyses from satellite and in situ observations. In addition, this study now incorporates fields of surface currents (U+V) derived from SL and wind stress by combining the geostrophic approximation and Ekman dynamics (Bonjean and Lagerloef, 2001). Assimilation results are evaluated using RMS with observed quantities - TOPEX/Poseidon SL (top), SST (middle) and zonal currents (bottom). The impact of single-variable assimilation (next column to right) and multivariate assimilation (second column over to right) are shown.

ROKF Equations

The reduced order Kalman filter used here is based on the SEEK filter equations (Pham et al., 1998) as modified by Verron et al.(1999) and Gourdeau et al. (2000). Analysis-step equations of the SEEK filter are directly derived from Kalman filter equations by expressing the error covariance matrix, $\mathbf{P}^f$ as

$$\mathbf{P}_k^f = \mathbf{N}_k \Lambda_k \mathbf{N}_k^T \tag{1}$$

where subscript k refers to time, t_k. $\mathbf{N}_k$ is the basis of the analysis and Λ_k is the error covariance restricted to the analysis subspace. Matrices $\mathbf{P}_k^f$, $\mathbf{N}_k$, and Λ_k, are the same order as the state vector dimension. However, if the rank of $\mathbf{P}_k^f$ is r, only r columns of $\mathbf{N}_k$ and Λ_k need to be retained, reducing the numerical cost of the analysis step of the Kalman filter.

If $\mathbf{y}_k^o$ represents a set of observations at time t_k, and $\mathbf{x}_k^f$ is the current guess of the system state, the new estimate of the state of the system, $\mathbf{x}_k^a$, is obtained using equations (see Pham et al. 1998):

$$\mu = \left[\frac{1}{\rho}\Lambda_k^{-1} + (\mathbf{H}_k\mathbf{N}_k)^T\left(\frac{\mathbf{R}}{\sigma_{obs}^2}\right)^{-1}\mathbf{H}_k\mathbf{N}_k\right]^{-1}\mathbf{H}_k\mathbf{N}_k\left(\frac{\mathbf{R}}{\sigma_{obs}^2}\right)^{-1}\frac{\left(\mathbf{y}_k^o - \mathbf{H}_k\mathbf{x}_k^f\right)}{\sigma_{obs}} \tag{2}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \sigma_{mod}\mathbf{N}_k\mu \tag{3}$$

$$\mathbf{P}_k^a = \mathbf{N}_k\Lambda_{k+1}\mathbf{N}_k^T \tag{4}$$

where $\mathbf{R}$ (the observational error of $\mathbf{y}_k^o$) is a single scalar for each observational type and σ_{obs} and σ_{mod} are the spatially varying standard deviation for the observations and model data, respectively. ρ is the *"forgetting factor"* and $\mathbf{H}_k$ is the observational operator. Following Verron et al. (1999) and Gourdeau et al. (2000), the forecast step of the filter is given by:

$$\mathbf{x}_{k+1}^f = \mathbf{M}[\mathbf{x}_k^a] \tag{5}$$

$$\Lambda_{k+1} = \mathbf{C}_{k+1}\mathbf{C}_{k+1}^T \tag{6}$$

$$\mathbf{N}_{k+1} = \mathbf{N}_k\mathbf{C}_{k+1} \tag{7}$$

$$\Lambda_{k+1} = \frac{1}{\rho}\mathbf{I} \tag{8}$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{C}$ is the Cholesky decomposition, and $\mathbf{M}$ is the transition matrix associated with the model.

The bias correction method of Chen et al. (2000) is applied to reduce the systematic bias between the model and observations in the innovation. Monthly data from October 1992 to December 1999 for the NOASSIM experiment are used to calculate the model and error MEOFs. The number of model and error MEOFs used in this study were optimized by applying the technique to the NOASSIM output and picking the number of MEOFS that minimized the global RMS with observations. These correspond to 18, 5, 17, 17 for the model and 13, 16, 11, 19 for the number of error MEOFs for sea level, SST, zonal and meridional currents, respectively.

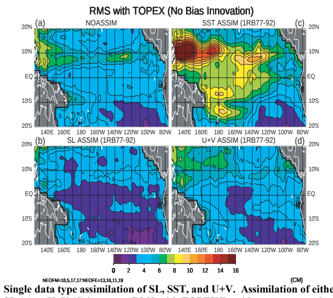
Model Description

- Reduced gravity, primitive equation, sigma coordinate model (Gent and Cane 1989) for the period October 1, 1992 - December 30, 1999.
- Model grid - Pacific basin (124°E - 284°, 30°N - 30°S). The grid is compressed to ≈ 0.5° at equator, north and south boundaries and 1° in longitude. Realistic coastline geometry and 20 sigma layers in the vertical.
- Wind forcing (COAPS), solar radiation (ERBE), cloudiness (ISCPP) and precipitation (Oberhuber, 1988) specified externally.
- Embedded variable depth oceanic mixed layer (Chen et al. 1994) simulates oceanic vertical turbulent mixing.
- Model explicitly accounts for complete upper ocean hydrology (Murtugudde et al., 1996).
- OGCM coupled to advective atmospheric mixed layer (AML) of Seager et al. 1995.
- The state vector (size of 1259167 points) for assimilation is defined as: $\mathbf{x} = (\mathbf{SL} : \mathbf{h} : \mathbf{T} : \mathbf{U} : \mathbf{V} : \mathbf{S})$ where $\mathbf{SL}$ - sea level, $\mathbf{h}$ - layer thickness, $\mathbf{T}$ - temperature, $\mathbf{U}$ and $\mathbf{V}$ - zonal and meridional currents and $\mathbf{S}$ - salinity. The contributions of all 20 layers for each variable between 20°N and 20°S are included in the state vector.
- A single reduced basis ($\mathbf{N}$) is calculated using monthly mean output of the model run from October 1977 to September 1992. 60 MEOFs, accounting for 90% of the variance, are retained for all experiments.

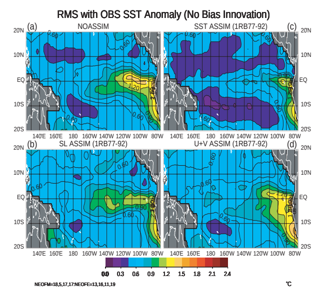
Assimilation Data Description

- Sea Level Observations - The NASA/GSFC Ocean Pathfinder TOPEX/Poseidon altimeter analysis is used to represent the observed sea level signal. All data are gridded from track data into 1°x 1°x 10 day bins and interpolated in time to the assimilation time ($\mathbf{R_k} = 36cm^2$).
- Sea Surface Temperature Observations - The blended in situ and satellite SST analysis of Reynolds and Smith (1994) is used to provide the observed SST signal. Original 1°x 1°x 7day gridded analyses are interpolated in time to the assimilation time step ($\mathbf{R_k} = 0.25°C^2$).
- Zonal and Meridional Current Observations - Total near-surface currents are computed as a sum of geostrophic (from TOPEX gradient) and Ekman parts (from SSM/I winds see Bonjean and Lagerloef, 2001). The data are regridded to 1°x 1°x 10 day bins and interpolated in time to the assimilation time step ($\mathbf{R_k} = 900\frac{cm^2}{s^2}$).

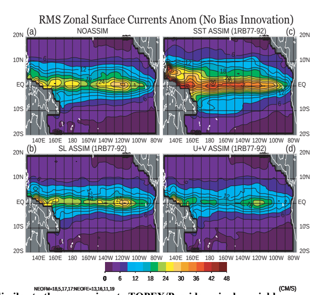
SINGLE-VARIABLE ASSIMILATION



Single data type assimilation of SL, SST, and U+V. Assimilation of either SL (a) or U+V (d) improves RMS with TOPEX/Poseidon as compared to NOASSIM (control) run (a). Assimilation of SST (c) degrades SL simulation.

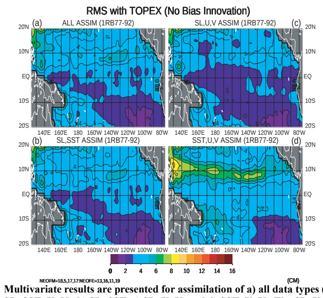


Assimilation of SST (c) improves simulation while SL (b) and U+V (d) assimilation improves slightly compared to NOASSIM control run (a) in the cold tongue region.

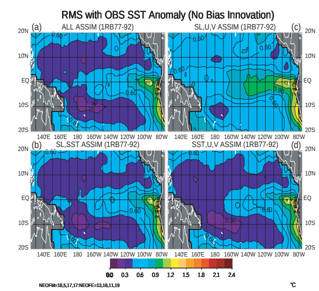


Similar to the comparison to TOPEX/Poseidon, single variable assimilation of both SL (b) and especially U+V (d) improves simulation of zonal currents. Assimilation of SST (c) degrades results significantly.

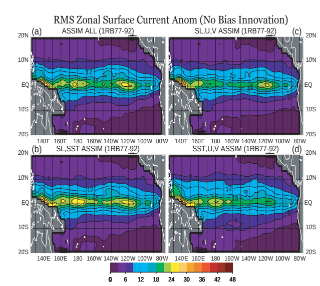
MULTIVARIATE ASSIMILATION



Multivariate results are presented for assimilation of a) all data types (ie. SL, SST, U+V), b) SL, SST, c) SL, U+V, and d) SST, U+V. The SL, U+V experiment is comparable to the SLASSIM run. However, all experiments which include SST increase RMS in the NECC region but still improve upon the NOASSIM case.

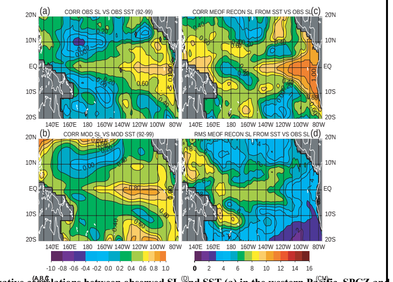


All experiments which include SST (a, b, d) significantly improve RMS. The SL, U+V experiment which excludes SST improves RMS between 130W-80W along the equator.

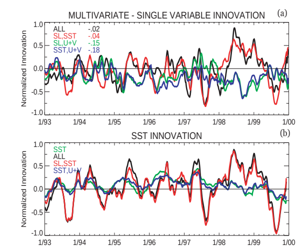


For zonal current, all experiments show that the RMS is improved with respect to the NOASSIM case. RMS is especially small when the surface variables (ie. SST, U+V) are combined (d).

SUMMARY



Negative correlations between observed SL and SST (a) in the western Pacific, SPCZ and NECC regions are not reproduced by the model (b). This leads to an unrealistic relationship between SL and SST in the MEOF basis. The low correlation between TOPEX/Poseidon and the SL obtained by an observed SST-MEOF fitting illustrates the consequences of using erroneous MEOFs (c).



Negative values for the spatial mean difference between the multivariate and single-variable assimilation experiments (a) indicate that the multivariate assimilation is converging to the observations better than the single-variable assimilation. The large positive values in 98-99 for ALL and SL, SST experiments are associated with large SST innovations when both SL and SST are assimilated (b).

Conclusions

- **Assimilation experiments show the zero-th order success of the technique – as expected each parameter (SL, SST, U+V) is improved after its individual assimilation.**
- **Comparison between observed quantities and experiments which assimilate a different data type show mixed results:**
  - **Observed SL: U+V improves SL of the model results, whereas assimilation of SST degrades the simulation**
  - **Observed SST: SL and U+V assimilation improves the model SST in the cold tongue region only.**
  - **Observed zonal current: SL assimilation improves the zonal current of the model but SST assimilation degrades the simulation.**
- **When SL and SST are assimilated using a multivariate approach the results are degraded due to deficiencies in the basis function used in equation (1), especially in the western Pacific. However, multivariate assimilation of SL, U+V and SST, U+V improve results with respect to the single-variable assimilation and NOASSIM experiments, even for the non-assimilated validation observation.**
- **This leads to the conclusion that the multivariate approach is successful at constraining non-assimilated variables when the reduced basis is formulated properly.**
- **In the future, we will attempt to solve the problem of the poor basis by exploring alternative function basis, using regional strategies, and by allowing the dynamical evolution of the basis.**

Contact: Eric Hackert (ehackert@essic.umd.edu) or http://www.essic.umd.edu/~ehackert

October 21-23, 2002, Jason Science Working Team Meeting, New Orleans, LA