An Eddy-Permitting 4DVAR Assimilation System for the State Estimation of the Tropical Pacific

Ibrahim Hoteit[1], Bruce Cornuelle[1], Patrick Heimbach[2], Armin Köhl[3], Detlef Stammer[3]

[1]Physical Oceanography Research Division, Scripps Institution of Oceanography, USA
[2]Dept. of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, USA
[3]Center for Marine and Climate Research, University of Hamburg, Germany,
Correspondence: ihoteit@ucsd.edu

INTRODUCTION

The 4DVAR tropical Pacific assimilation system is based on the MIT general circulation model (MITgcm) and its automatically generated adjoint. The model uses realistic topography with parameterizations for the surface boundary layer (KPP) and open boundaries at the south and north, as well as in the Indonesian throughflow. The adjoint method is used to adjust the model to observations in the tropical Pacific region using control parameters which include initial temperature and salinity, surface fluxes of momentum, heat and freshwater, and temperature, salinity and horizontal velocities at the open boundaries. The model is constrained with most of the available datasets in the tropical Pacific, including climatologies, TAO, Argo, XBT, and satellite SST and SSH data. Hindcast experiments were performed for one year in 2000 to test the assimilation system and to validate its outputs. It is shown that the 4DVAR method was able to significantly improve the model consistency with all multivariate data sets, providing a reasonably accurate and dynamically consistent picture of the tropical Pacific large-scale circulation.

THE MODEL

We use the MIT ocean general circulation model which solves the Navier-Stokes equations, under the Boussinesq approximation (Marshall et al., 1997). The model equations are written in z-coordinates and discretized using a staggered Arakawa C-grid. The numerical code is further designed to allow for the construction of the adjoint using the automatic differentiation tool TAF.

The model domain covers the entire tropical Pacific basin between 26o and 26oN. The maximum depth is at 6000m. In separate experiments, the model was integrated on a $1^o\times1^o$, $1/3^o\times1/3^o$ and a $1/6^o \times1/6^o$ Mercator grid, each with 39 vertical levels (with appropriate horizontal viscosities and diffusivities in each case). The vertical resolution is spaced at 10m from the surface to 300m below. The model operates in hydrostatic mode with an implicit free surface. No-slip conditions are imposed at the lateral boundaries while the friction condition is quadratic at the bottom. The sub-grid scale physics is a tracer diffusive operator of second order in the vertical, the eddy coefficients being parameterized by the KPP mixed layer model. Horizontally, diffusive and viscous operators can be either of second or fourth order. Open boundaries at 26 ^{o}S and 26 ^{o}N, as well as at four straits in the Indonesian throughflow, are implemented as in (Zhang and Marotzke, 1999). The values at the open boundaries are prescribed by the ECCO 1^o global state analysis (Köhl et al., 2007). In the runs without assimilation, the model is forced either with NCEP or ECCO forcing, and a relaxation term is included to relax surface temperature towards monthly climatology with a 30-day time-scale.

MODEL/ DATA COMPARISON AND SENSITIVITY EXPERIMENTS

Before beginning assimilation, five different 9-year forward model integrations (Table 1) were carried out between 1992 and 2001 to test the model and to study the model sensitivity to the forcing fields and to the horizontal resolution up to 1/6°. These runs were compared with the ECCO 1° global reanalysis with 23 layers.

Table 1. *Experiments for horizontal resolution and forcing sensitivity studies.*

	Resolution	***Forcing***	***Time Step***
1DEG	1°	ECCO	1h
NCEPF	1/3°	NCEP	1/2h
ECCOF	1/3°	ECCO	1/2h
1SIXTH	1/6°	ECCO	1/6h

Comparison of zonal sections of mean zonal velocity along the equator (EUC) with Johnson analysis (Johnson et al., 2002) is shown in Figure 1. 1/3° horizontal resolution seems adequate to simulate most of the tropical Pacific circulation and 1° is not enough. The use of more vertical layers improves the equatorial currents, but 39 layers is perhaps not the maximum. ECCO forcing provides a better EUC than NCEP forcing.

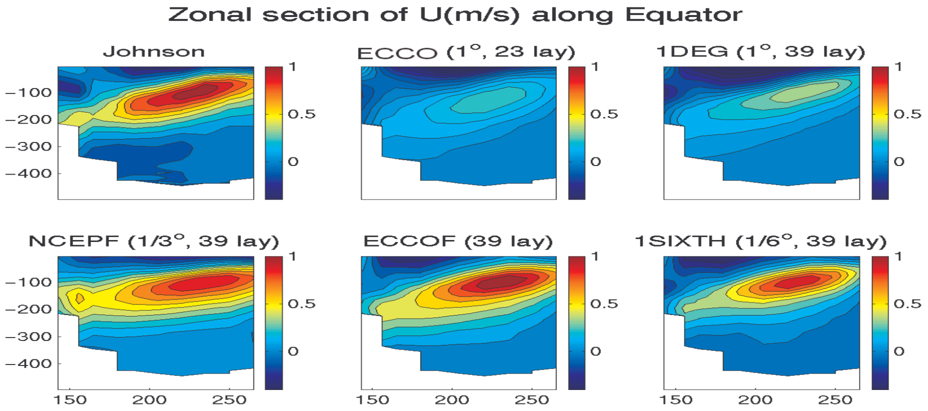

Figure 1. *Zonal section of mean zonal velocity along the equator from different model runs and Johnson analysis.*

Comparison of model RMS with TOPEX data suggests that 1° model (Figure 2) is in best agreement with the data. But looking at other field shows that ECCO forcing produces an unrealistically strong NECC with high resolution models while the SSH variability from NCEP is weak.

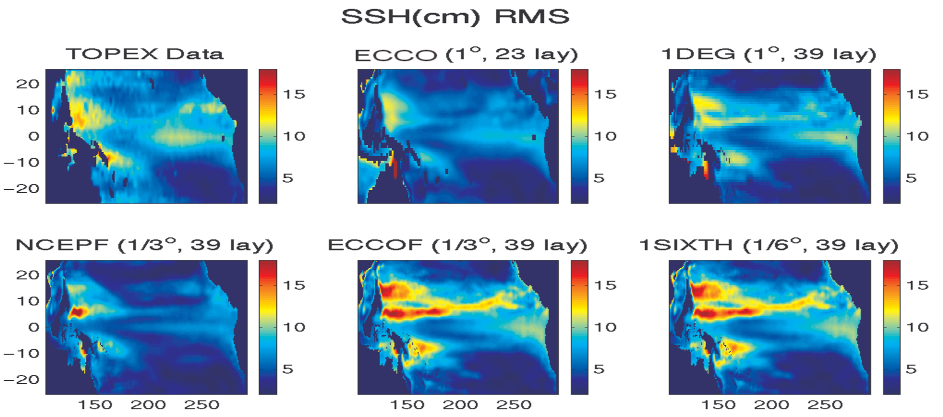

Figure 2. *SSH RMS from different model runs and TOPEX measurements.*

The main conclusions from the model/data comparison and the sensitivity experiments are: (i) The model without assimilation behaves fairly well, (ii) the optimized 1° ECCO forcing has some skill for nested higher resolution models, but can be improved, (iii) horizontal spacing could be set to 1/3° grid spacing in the tropical Pacific to accrue the benefits of enhanced resolution without paying a steep price in computer-time. Serious differences from the observations indicate that more work must be carried out to improve the quality of the forcing before increasing the horizontal and vertical resolution.

THE ASSIMILATION SYSTEM

A 4DVAR assimilation system used the 1/3° tropical Pacific model and its adjoint. The assimilation system was evaluated over a one year period starting from January 1st, 2000. Descent directions toward the minimum are determined using the Quasi-Newton M1QN3 algorithm.

Assimilated Data

Assimilated data sets are multivariate and include

- *Satellite Data*: TOPEX SSH anomalies, TOPEX mean SSH (minus Grace Geoid), and Weekly TMI SST.
- *Profiles*: 5-day averages of TAO data (S, T, U and V), Drifters (U and V), Argo profiles (S and T), Floats (S and T), XBTs, and CTDs.
- *Climatologies*: Levitus T and S, Reynolds SST, and Mean Johnson U.

Control Variables

The adjoint method was used to improve model/data consistency by adjusting:

- *Initial Conditions*: S and T.
- *Atmospheric Forcing*: Wind stress, heat flux and Salinity flux adjusted every two days.
- *Open Boundaries*: S, T, U and V adjusted every week.

Background States and First Guesses

The following background states were taken as starting points for optimization.

- *Initial Conditions*: Levitus climatologies.
- *Atmospheric Forcing*: NCEP analysis (or QSCAT for TAUU and TAUV).
- *Open Boundaries*: ECCO state analysis.

Errors and Enforcing Smoothness of the Control

Data and model errors were prescribed only on the diagonal of the error covariance matrices. S and T errors were parameterized by displacements of order 20m acting on potential T and S gradients from the Levitus climatology. Errors for SSH anomalies, SST, U, V data and forcing fields were estimated as a fraction of the data standard deviations. Mean SSH was assimilated with an accuracy of 4.5cm. Errors in Johnson mean U are set as 20% of the total signal.

Smoothness of the control variables is achieved by penalizing the Laplacian (for the horizontal correlations) and the first derivatives (for the vertical and time correlations) of the control fields in the cost function. Combining the Laplacian and first derivatives penalties terms with diagonal weights respectively approximates a Gaussian and an exponential covariance for the control fields. In addition, the time-mean vertical velocity over the first 3 months was constrained to ``normal conditions" in order to reduce the spurious effect of model adjustments to changes in the initial S and T.

ASSIMILATION RESULTS

The results presented below were obtained after 100 optimization iterations.

Individual cost function terms for the data and the control terms before and after assimilation are plotted in Figure 3. The assimilation reduces the data misfits and ageostrophic penalties. The control cost terms show that wind stress dominates the adjustments.

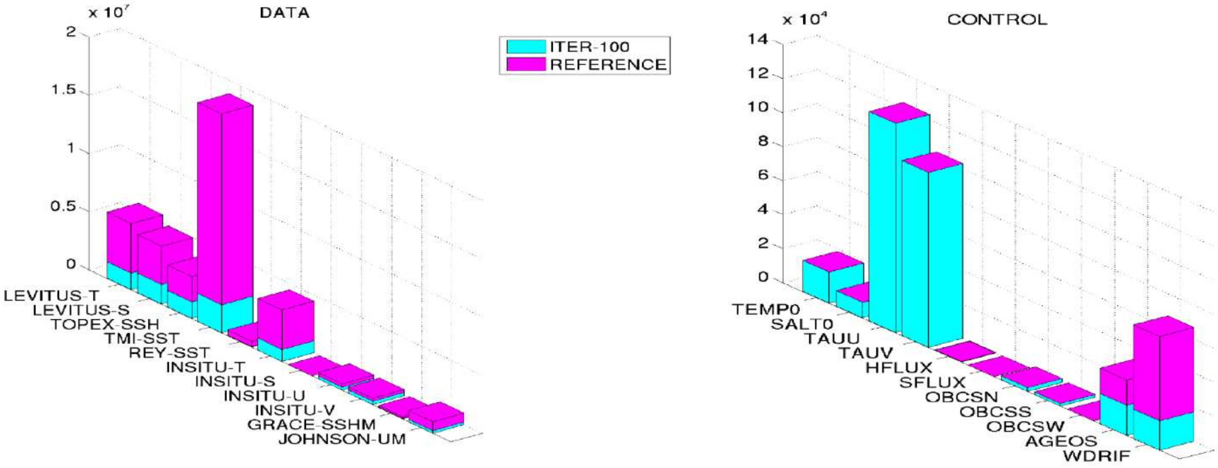

Figure 3. *Individual cost function terms before and after assimilation for data and control.*

Figure 3 shows that the adjusted control parameters produce a statistically better fit to the observations, but it is interesting to look at a few snapshot comparisons. Figure 4 shows one-week-averaged SSH and SST for the last week in July, 2000. The reference run (left panels) was the starting point for the assimilation, and the agreement with the data (center panels) is obviously much better for the adjusted solution (left panels). The snapshots show the details of the variability allowed in the 1/3 resolution model.

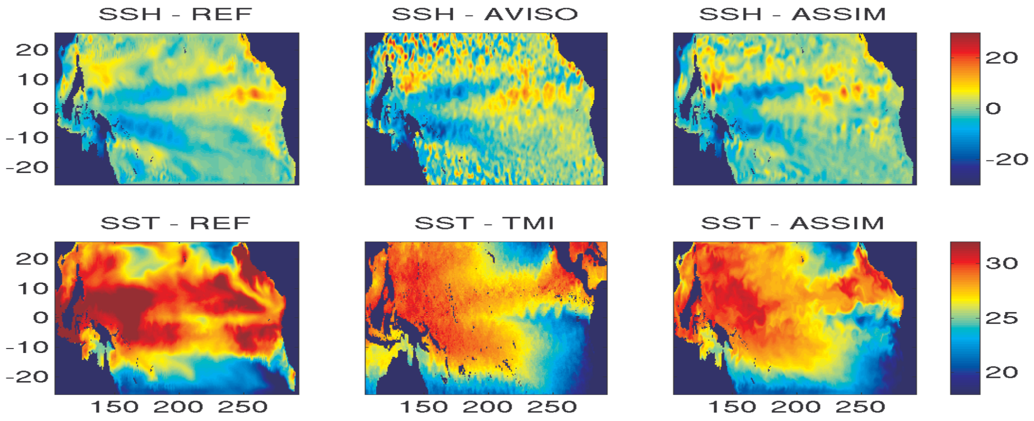

Figure 4. *Weekly fields (end of July) from the reference, data, and assimilation*.

One of the questions that arises in assimilation is how much the choice of starting point affects the final solution. To address this, several assimilation experiments were started from different initial conditions and forcing. The starting and ending zonal velocity fields are shown in Figure 5 after 30 iterations. TAUU solution seems to converge toward similar solutions.

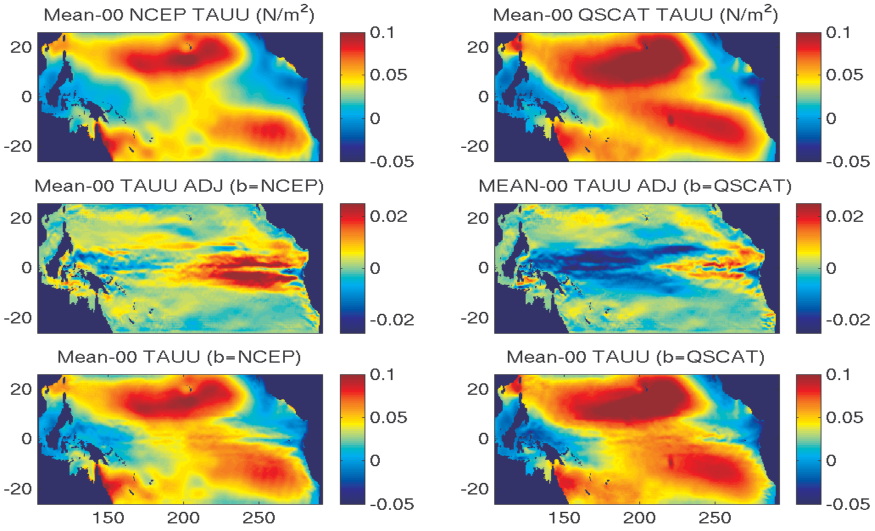

Figure 5. *Mean adjustments to TAUU starting the optimization from NCEP or QSCAT winds.*

DOWNSCALING TO 1/6^o RESOLUTION

The estimated forcing, initial conditions, and boundary conditions constitute a "recipe" for a model run that reproduces the observations. In principle, this recipe should be useable by any model. As a preliminary test of this concept, the control parameters (except open boundaries) estimated by the 1/3^o assimilation were used to run the 1/6^o model over the same time range. The only difference was the horizontal resolution and the mixing levels.

Mean zonal velocities in a vertical section on the equator for the 1/3 ^{o} and 1/6 ^{o} model runs with adjusted controls. The wind adjustments alone improve the model but the best 1/6 ^{o} run is obtained using all adjusted controls.

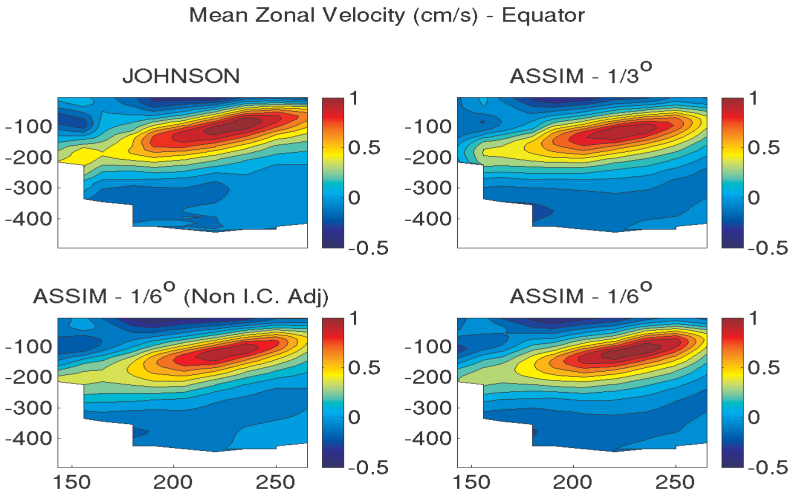

Figure 6. *Mean zonal velocities in a vertical section on the equator for the data, 1/3^o and 1/6^o model runs with adjusted controls.*

In addition to the averaged comparisons, it is again interesting, as in Figure 4, to look at snapshots of SST and SSH. The panels on the left in Figure 7 are the observations, and were in the center of figure 4. The panels in the center are from the 1/3^o model run with the optimized forcing, the same as the right panels of Figure 4. The right panels of figure 7 show the 1/6° model run with the same optimized forcing, and shows that the higher-resolution model run responds well to the optimized forcing, retaining good agreement with the observations even on a weekly basis. The TIW region shows deviations from the observations, due to the lack of controllability of these instabilities.

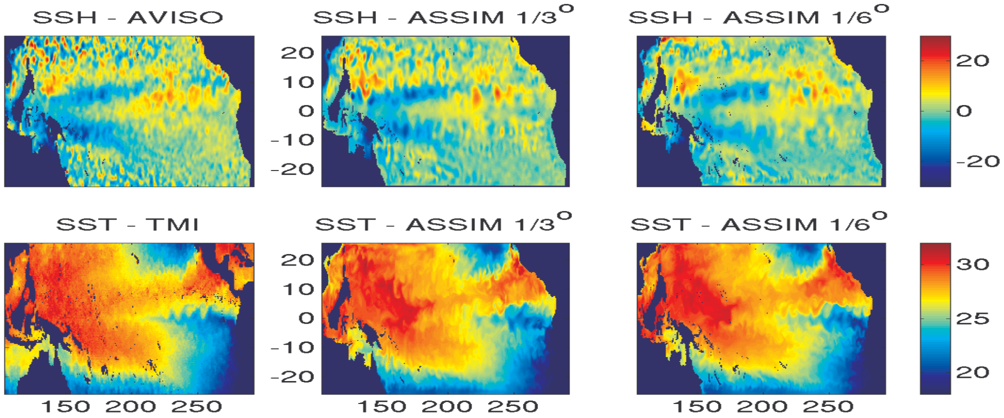

Figure 7. *Weekly fields (end of July) from the data, 1/3° and 1/6° adjusted runs.*

DISCUSSION AND FUTURE WORK

The adjusted model run for 2000 provides a test bed for diagnosing processes such as EUC momentum balance or TIW energy flows. The SST signal in the TIW region is not well-represented by the adjusted solution, since these instabilities were deliberately reduced in the adjoint runs used to adjust the controls. Further experiments to assess the controllability of these instabilities over shorter time-scales are contemplated. The similarity of the 1/3° and 1/6° model runs with the adjusted controls confirms the results of pre-assimilation sensitivity studies. Future work includes:

- Use dynamically balanced background covariance matrix.
- Include mixing parameters as controls.
- Impose smoothness of the control through diffusion.
- Extend the assimilation period.
- Assess the controllability of TIWs.

REFERENCES

Hoteit, I., B. Cornuelle, A. Köhl, and D. Stammer, 2005: Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. QJRMS, **131**, 3659-3682.

Johnson, G., B. Sloyan, W. Kessler and K. McTaggart , 2002: Direct measurements of upper ocean currents and water properties across the tropical Pacific during the 1990s. Prog. Oceanog., **52**, 31-61.

Köhl, A., D. Stammer, and B. Cornuelle, 2006: Interannual to Decadal Changes in the ECCO Global Synthesis. J. Phys. Oceanog**.**, **54**, 406-425.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997: A finite-volume, incompressible Navier-Stokes model for studies of the ocean on parallel computers. J. Geophys. Res., **102**, 5753-5766.

Stammer and coauthors, 2002: The global ocean circulation during 1992 -1997, estimated from ocean observations and a general circulation model. J. Geophys. Res., **107**, 3118-3145.

Zhang, K., and J. Marotzke, 1999: The importance of open boundary estimation for an Indian ocean GCM data synthesis. J. Mar. Res., **57**, 305-334.