Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation

By I. HOTEIT1∗, B. CORNUELLE1, A. KÖHL2 and D. STAMMER2

1Scripps Institution of Oceanography, USA
2Institut fuer Meereskunde, Germany

(Received 23 May 2005; revised 15 March 2006)

SUMMARY

A variational data assimilation system has been implemented for the tropical Pacific Ocean using an eddy-permitting regional implementation of the MITgcm. The adjoint assimilation system was developed by the Estimation of the Circulation and the Climate of the Ocean consortium, and has been extended to deal with open boundaries. This system is used to adjust the model to match observations in the tropical Pacific region using control parameters which include initial conditions, open boundaries and time-dependent surface fluxes. This paper focuses on problems related to strong adjoint sensitivities that may impede the model fit to the observations. A decomposition of the velocities at the open boundaries into barotropic and baroclinic modes is introduced to deal with very strong sensitivities of the model sea surface height to the barotropic component of the inflow. Increased viscosity and diffusivity terms are used in the adjoint model to reduce exponentially growing sensitivities in the backward run associated with nonlinearity of the forward model. Simplified experiments in which the model was constrained with Levitus temperature and salinity data, Reynolds sea surface temperature data and TOPEX/POSEIDON altimeter data were performed to demonstrate the controllability of this assimilation system and to study its sensitivity to the starting guesses for forcing and initial conditions.

KEYWORDS: Adjoint method Open boundaries Tropical Pacific

1. INTRODUCTION

Understanding the circulation of the tropical ocean has long been a goal of physical oceanography. As the Southern Oscillation and its relationships to El Niño and the extratropics began to be understood (Bjerknes 1966), more attention was focused on the tropical Pacific ocean. The spectacular global climate impacts associated with the strong 1997–98 El Niño event underscored the need for better understanding of tropical Pacific features. This event was one of the most widely discussed climate events in recent history. The great variability between El Niño events has not yet been completely understood, and leads to other questions, such as the response of the El Niño/Southern Oscillation (ENSO) to large-scale warming.

Accurate description of the spatio-temporal structure of the tropical Pacific ocean is a key step towards understanding these events and better studying the Tropics currents and their interactions. The circulation of the tropical Pacific has been studied extensively in recent years using both observations and numerical models. However, to date most of these studies were limited by the small amount of available data (Lagerloef et al. 1999; Durand and Delcroix 2000; Vialard et al. 2001). On the other hand, numerical ocean models provide a crude approximation of reality because of the many approximations used. The observational dataset is both sparse and inhomogeneous, and one of the challenges of data management is to combine the disparate data types and ocean models to obtain a dynamically coherent and useful four-dimensional picture of the state of the tropical Pacific Ocean (Wunsch 1996).

High-resolution tropical Pacific models require significant computing resources because of the very large basin. Consequently, inexpensive data assimilation techniques were implemented with these models, based on optimal interpolation (e.g. Giese and

∗ Corresponding author: Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA-92122, USA. e-mail: ihoteit@ucsd.edu
© Royal Meteorological Society, 2005.

Carton 1999) or the three-dimensional variational (3D-Var) methods (e.g. Tang et al. 2004). These methods provide a static analysis of the system state using observations over a given time, without enforcing smoothness in the time dimension. At present, most of the advanced assimilation schemes are based on the adjoint method or on the Kalman filter (Wunsch 1996). For instance, Bennett et al. (1998) used the representer method, which solves the dual formulation of the adjoint problem, to assimilate Tropical Atmosphere and Ocean (TAO) data into an intermediate coupled ocean–atmosphere model. Bonekamp et al. (2001) also used the adjoint method to adjust the model wind stress over the tropical Pacific while constraining the model to temperature data profiles using a twin experiments approach. Recently, Weaver et al. (2003) successfully tested an incremental approach of the four-dimensional variational (4D-Var) assimilation problem with a primitive-equation model while assimilating temperature profiles and adjusting initial conditions. Since a full implementation of the Kalman filter is not possible in practice, simplified Kalman filters with different degrees of approximations were also used to assimilate altimetric data and TAO data in the tropical Pacific (e.g. Fukumori 1995; Cane et al. 1996; Hoteit et al. 2002).

Most of the previous assimilation studies in the tropical Pacific used a restricted set of observations or relatively low horizontal or vertical resolutions to reduce computational burdens. We are implementing an eddy-permitting regional adjoint assimilation system for the tropical Pacific ocean using the MITgcm (Marshall et al. 1997). Boundary conditions and forcings come from the Estimation of the Circulation and the Climate of the Ocean (ECCO) global 1° resolution adjoint assimilation (Stammer et al. 2002; Köhl 2005, personal communication) using the same model and methods. The purpose of this paper is to describe the new technical features of the tropical Pacific assimilation system, which is an example of variational data assimilation in a nonlinear model with widely varying adjoint sensitivities. Hereafter we will distinguish between two types of large adjoint sensitivities: the first type of sensitivities remain correct for finite perturbations and the second type are correct only for infinitesimal perturbations. The first type of sensitivities can lead to slow descent in the optimization problem but this can be overcome by preconditioning (Zupanski 1996). The second type is generally associated with rapidly growing perturbations (‘intrinsic variability’) that are not easily predictable or controllable by the system. These sensitivities generally grow exponentially in time and are not useful in the optimization problem because the range of validity of the linear approximation is smaller than the uncertainty in the control parameters. In other words, the nonlinearity of the system invalidates the use of the gradient for descent (Pires et al. 1996; Köhl and Willebrand 2002).

The paper is organized as follows: section 2 describes the assimilation system, including the model and the assimilation method. The main characteristics of the datasets used in the study are reviewed in section 3. The assimilation experiments are described and the assimilation results are evaluated against independent data in section 4. Finally, concluding remarks are given in section 5.

2. The Assimilation System

(a) Ocean model

This study uses a regional implementation of the ECCO MITgcm assimilation system. The model MITgcm is a general-circulation model developed at the Massachusetts Institute of Technology (MIT) (Marshall et al. 1997) to support ocean general-circulation studies over a broad range of scales and physical processes. It is based on the primitive equations on a sphere under the Boussinesq approximation.
The equations are written in height $z$-coordinates and discretized using the centred second-order finite-differences approximation in a staggered ‘Arakawa C-grid’. The numerical code is designed to allow for the construction of the adjoint model using the automatic differentiation tool Transformation of Algorithms in Fortran; see Giering and Kaminski (1998), as described by Heimbach et al. (2005). Stammer et al. (2002) produced the first $2^\circ \times 2^\circ$ global state estimation using ECCO assimilation system.

The regional domain includes the entire tropical Pacific, extending from $26^\circ$S to $26^\circ$N and from $104^\circ$E to $68^\circ$W. The bathymetry is extracted from the global topography prepared by Smith and Sandwell (1997) and adjusted slightly to leave the Indonesian throughflow (ITF) pathways open and to limit the maximum depth to 6000 m. The model is integrated on a $1/3^\circ \times 1/3^\circ$ Mercator grid, with 39 vertical levels. The vertical levels are spaced at 10 m from the surface to 250 m in depth, with spacing gradually increasing to 300 m below. The model is operated in a hydrostatic mode with an implicit free surface. No-slip conditions are imposed at the lateral boundaries while bottom friction is quadratic with a drag coefficient equal to 0.002 m$^{-1}$. The subgrid-scale physics is a tracer diffusive operator of second order in the vertical, with the eddy coefficients determined by the K-profile parametrization (KPP) model (Large et al. 1994). In the horizontal, diffusive and viscous operators are of second and fourth order, respectively, with coefficients $5 \times 10^2$ m$^2$s$^{-1}$ and $1 \times 10^{11}$ m$^4$s$^{-1}$, respectively. Vertical diffusivity and viscosity are parametrized by Laplacian mixing with values $1 \times 10^{-6}$ m$^2$s$^{-1}$ and $1 \times 10^{-4}$ m$^2$s$^{-1}$, respectively.

Open boundaries (OB) are set at $26^\circ$S and $26^\circ$N, as well as at four straits in the Indonesian throughflow. The OB are implemented as in Zhang and Marotzke (1999). Temperature $T$ and salinity $S$ and the horizontal components of the velocity ($U$, $V$) are specified on the boundary. A smooth transition to the prescribed conditions at the boundaries is achieved by a sponge layer extending $3^\circ$ from the boundary in which the model solution is relaxed to the boundary values. Monthly mean values (centred on the 15th of each month) obtained from the ECCO global state estimate were prescribed at the grid points just outside the OB and the model solution is relaxed to these values within a buffer zone of $3^\circ$ over time-scales varying linearly from 1 day at the boundary and 40 days at the edge of the zone. The normal velocity fields across the open boundaries have been further corrected on a monthly basis to enforce the same transport at $26^\circ$N and $26^\circ$S as in the global ECCO model, and to exactly balance the volume flux into the domain by the transport out in the ITF. The corrections needed at each boundary are added as a barotropic transport uniformly distributed over all grid points on the boundary. In the ITF, 1/2, 1/3 and 1/6 of the correction is added to the transport in the Ombai strait, Timor passage and Lombok strait, respectively.

Fluxes of momentum, heat, and fresh water are prescribed at the ocean–atmosphere interface. Two sets of forcing fields are used in the experiments. The NCEP forcing consists of sea surface fields from the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) re-analysis project (Kalnay et al. 1996). It contains twice-daily wind stress vectors and daily net heat flux, net short-wave radiation and water flux at the sea surface. The second forcing dataset was taken from the ECCO flux fields (Stammer et al. 2004). These are the NCEP forcings adjusted by a variational $1^\circ \times 1^\circ$ global state estimation procedure (Köhl 2005, personal communication).

Prior to the data assimilation experiments, the model was integrated over a nine-year period from 1992 to 2001 with a variety of forcings and resolutions. At $1^\circ \times 1^\circ$ resolution, the solution was consistent with the global ECCO runs used as boundary conditions. At $1/3^\circ$ resolution, the model fidelity to the observations improved significantly.
away from the boundaries. At 1/6° resolution, the solution did not change qualitatively from the 1/3° solution. The model with ECCO forcing showed good agreement with sea surface height (SSH), sea surface temperature (SST), drifter, TAO, and Johnson acoustic Doppler current profiler data in the eastern Pacific (giving reasonable cold-tongue temperature agreement with data and reasonable equatorial undercurrent (EUC), north equatorial counter current (NECC), and south equatorial currents (SEC)). It was also able to reproduce many features of the observed current and temperature variability during the nine-year period, including Tropical Instability Waves (TIWs). The NCEP forcing produced better agreement in the western Pacific, but none of the forcing fields gave good agreement over the entire basin. It was hoped that assimilation could correct the misfits by adjusting the forcing fields along with initial conditions (IC) and open boundary conditions.

(b) Assimilation method

Assuming that the model physics are accurate, the model state can in principle be brought into agreement, within error estimates, with the observations by adjusting an identifiable set of model parameters. This can be posed as the minimization of a cost function measuring the difference between the model solution and data over a specified period of time and constrained by the model equations subject to a set of control variables. The gradient of the cost function is used to determine descent directions toward the minimum in an iterative procedure. An efficient way to compute the gradient of the cost function is to use the adjoint method which provides the gradient by integrating the adjoint of the tangent linear model backwards in time (Le Dimet and Talagrand 1986; Wunsch 1996).

In its general form, the objective (cost) function consists of a weighted sum of quadratic norms of model-data misfit ($J_{\text{obs}}$) and changes to the control variables ($J_{\text{c}}$), between the initial time ($t_0$) and the final time ($t_f$) of the assimilation window,

$$
J(u) = \sum_{t=t_0}^{t_f} \left\{ y(t) - H_t(x(t)) \right\}^T R(t)^{-1} \left\{ y(t) - H_t(x(t)) \right\} + \sum_{t=t_0}^{t_f} \left\{ u(t) - u^b(t) \right\}^T Q(t)^{-1} \left\{ u(t) - u^b(t) \right\},
$$

(1)

where $x(t)$ is the model state vector and $u^b(t)$ a first guess (or ‘background’) of the unknown control vector $u(t)$ at time $t$; $u$ represents uncertainties in the model parameters, the external forcing fields, and the internal model physics which fully control the evolution of the model state. The vector $y(t)$ contains all observations available at time $t$ and is related to the model state according to

$$
y(t) = H_t(x(t)) + \epsilon(t),
$$

(2)

where $H_t$ is the observation operator and $\epsilon(t)$ represents the associated observation errors; $R(t)$ and $Q(t)$ are weight matrices representing the covariance of observation errors and control (‘background’) errors, respectively. In the formulation of (1), it is implicitly assumed that the errors in the control vector are uncorrelated with the observation errors, and that both errors are mutually uncorrelated in time. In the following, the background state is taken as the starting point for minimization.
The weight matrices $R$ and $Q$ are important in determining the solution of the 4D-Var problem. In practice, however, there are insufficient observations to determine these matrices and ad hoc estimates are used instead. Observational errors are frequently assumed to be spatially uncorrelated, so that $R$ is diagonal. The control parameter covariances are generally not assumed to be diagonal in order to enforce smoothness in the estimated adjustments (Weaver et al. 2003). For analysis of the tropical circulation, smoothness is achieved efficiently in this assimilation system by penalizing the Laplacian of the control fields in the cost function. Combining the Laplacian penalty term with a diagonal $Q$ approximates a Gaussian covariance for the control fields (Bennett 2002). In experiments with this system, smoothness penalties increased the complexity of the system without eliminating the problems with large sensitivities, so in this diagnostic study the matrix $Q$ is assumed diagonal for simplicity.

The control variables for an efficient 4D-Var assimilation system should both control most significant ocean state variability and have minimum dimension to reduce computational burden and to speed the decrease of the cost function. For this reason, many investigators have restricted the control vector $u$ to a subset of the possible parameters or have reduced the dimension of the control variables through parametrizations. In early work, ocean models were controlled through the modification of the initial conditions only (e.g. Luong et al. 1998; Weaver et al. 2003). The influence of the adjusted initial conditions on the ocean state decays in time, and this limits the length of the assimilation window (Luong et al. 1998). Bonekamp et al. (2001) investigated a 4D-Var scheme that adjusted the wind-stress forcing which is a primary driving force in the tropical Pacific. These controls are effective over long periods in the regions where the adjusted forcing can reach over time. However, surface forcing cannot completely correct for all uncertainties in the initial conditions (Vossepoel et al. 2004). A more general control vector which is made of the initial conditions and all atmospheric fluxes was considered by Stammer et al. (2002) and Köhl (2005, personal communication).

The latter control vector was adopted for the system used here and augmented with temperature, salinity, and velocities at the open boundaries for better control of the regional model. The term representing the constraints on the control variables in (1) has the form

$$J_c = (x_0 - x_0^b)^T Q_0^{-1} (x_0 - x_0^b) + \sum_{t=t_0}^{t_f} (f(t) - f^b(t))^T Q_f^{-1} (f(t) - f^b(t))$$

$$+ \sum_{t=t_0}^{t_f} (x_{ob}(t) - x_{ob}^b(t))^T Q_{ob}^{-1} (x_{ob}(t) - x_{ob}^b(t)).$$

where $x_0$, $f(t)$, and $x_{ob}(t)$ represents, respectively, the initial conditions, the forcing fields and the open boundaries conditions, and $Q_0$, $Q_f$, and $Q_{ob}$ the associated error covariances.

Uncertainties in the internal model physics were not represented in the control vector, so that regions out of reach of the surface forcing and boundary controls are determined by initial conditions alone. This may lead to uncontrollable misfits, as discussed above, such as model drifts at depths where the controls do not penetrate during short-time integrations. Recently Stammer (2005) and Ferreira et al. (2005) introduced errors in the internal physics of the ECCO MITgcm by adding control of the model mixing parameters, which represent some of the least-understood physical approximations in the model. This is far from the complete control used by Bennett et al. (1998), but provides an efficient compromise. In nine years of integration of the tropical Pacific model without assimilation, no large errors or drifts were observed in the
deep ocean away from the coasts. In the assimilation runs made here the surface controls were able to reduce the cost function and no large uncontrollable misfits were seen over one year of integration. This does not mean that model errors are not significant, and the forcing adjustments may include components that are compensating for model errors. We will include mixing parameter controls in the future, but each added set of control parameters must be evaluated by the reduction in the cost function that it provides, and the data and duration of the present experiments are not yet large enough to force the adoption of model error terms in the control vector.

(c) Control of the normal velocities at the open boundaries

One of the benefits (or problems) of assimilating with a regional model is the use of the open boundaries as adjustable controls. In the MITgcm, the boundary conditions require the complete specification of the ocean temperature, salinity and velocities on the boundary. Accordingly, four separate penalty terms were added to the cost function, one for each variable (Zhang and Marotzke 1999). This allows dynamically unbalanced open boundaries which can cause deterioration in the model solution. Following Gebbie (2004), an additional cost-function term was added that penalizes the deviation from the thermal-wind balance on the boundary (a ‘soft constraint’). The velocities were adjusted to enforce a zero net mass flux.

The estimation of velocities at the open boundary is often poorly conditioned when simultaneously adjusted with other control variables. The model SSH is very sensitive to changes in the total barotropic component of the normal velocities at the open boundary. This means strong sensitivity of the total SSH term in the cost function to the normal velocities at the open boundary. Large contrasts in sensitivities can lead to strongly anisotropic topology of the cost-function manifold, which can severely slow gradient descent optimization (Zupanski 1996). This topology should be regularized by preconditioning, ideally using the inverse of the Hessian matrix. The ECCO system uses diagonal preconditioning and its descent subroutine builds an approximated Hessian (Gilbert and Lemaître 1989). The standard preconditioning by the background-error covariance matrix was not sufficient to speed the descent in the cases tried here. Descent might be improved by using decreased weighting for all the boundary normal velocity control terms in the cost function. However, down-weighting the velocity acts on both barotropic and baroclinic components of the normal velocities, but only the barotropic velocities have high sensitivities. A better way to deal with this problem is to decompose the normal velocity at the open boundary into baroclinic and barotropic components and then apply reduced weights only for the barotropic component (Gebbie 2004). Defining the baroclinic component as the difference between the absolute and the barotropic velocities creates an additional constraint: the vertical integral of the baroclinic component must be zero. To enforce this hard constraint, Gebbie (2004) removed the first layer from the baroclinic component and set it to enforce zero net flow. However, this can result in unrealistic vertical shear between the first and lower layers.

In order to avoid unrealistic baroclinic structure, we adopted a normal mode decomposition to the velocities normal to the boundaries. Following Pedlosky (1987), quasigeostrophic normal modes are the solution of the vertical Sturm–Liouville problem

$$\frac{\partial^2 h_k}{\partial z^2} + \frac{N^2}{C_k^2} h_k = 0,$$

with $z$ the vertical coordinate and $h_k$ the $k$th vertical mode; $N(z)$ is the buoyancy frequency given by $N^2 = -g\rho_z/\rho_0$ ($g$ is gravitational acceleration, $\rho$ is the density...
profile, and $\rho_0$ is a constant reference density), and $C_k$ is the phase speed of long Rossby waves with mode $k$ vertical structure. This can be used to enforce smoothness in the vertical, but we only use it here as an orthogonal decomposition of the control parameters to separate the barotropic and the baroclinic components, and it is not used by the forward model.

This decomposition depends on local water depth and stratification and the number of modes is set equal to the number of layers. Taking fewer modes, or reducing the expected variance of high modes would enforce smoothness in the boundary control. In order to apply the decomposition in a model with varying topography, a different barotropic-baroclinic decomposition operator is determined for each location using the local depth and stratification. To keep the decomposition simple we assume a constant buoyancy frequency at all locations, so that the decomposition basis functions are sinusoids and depend only on local water depth. The velocity vector at each point along the open boundaries is decomposed into barotropic and baroclinic modes using the decomposition operator that corresponds to the local depth.

\[ (d) \quad \text{Treating undesirable adjoint sensitivities} \]

The use of the adjoint method for data assimilation with nonlinear models can be problematic. When the model is sufficiently nonlinear the cost function becomes non-convex, implying the existence of multiple local minima (Pires et al. 1996; Li 1991). This may prevent significant reduction in the cost function when using a gradient descent optimization algorithm, since these algorithms are only designed to converge toward local minima. In the case of multiple minima, the choice of the initial guess for the control variables becomes crucial.

A solution to this problem has been suggested by Pires et al. (1996) for the control of initial conditions. They propose starting with assimilation over a short period in which the linear approximation holds and then gradually increasing the length of the assimilation window while starting each successive (longer) optimization from the adjusted initial conditions determined in the previous (shorter) run. The aim is to maintain the initial conditions in the basin of the absolute minimum where the cost function is expected to be convex. Although this method was shown to improve the performance of the adjoint method, its usefulness over long time spans can be questioned since the number of local minima can grow exponentially with the length of the assimilation window. In this case, the cost function becomes too irregular to reduce with a gradient descent algorithm (Swanson et al. 1998). This approach also depends on the existence of sufficient information early in the assimilation period to give a good estimate of the control parameters and assumes a perfect model with the initial condition as the only source of error.

The presence of multiple local minima due to nonlinearity is associated with large gradients of the cost function. The adjoint solution for the tropical Pacific showed small, irregular regions of very high sensitivity which first appeared after 2–3 months in the western basin of the tropical Pacific and then spread to a large area centred on the equator. These small-scale but strong gradients indicate the presence of many small-amplitude but tightly packed extrema (Köhl and Willebrand 2002). Similar patterns in the adjoint solution were also reported in several recent papers (Janiskova et al. 1999; Lea et al. 2002; Zhu et al. 2002). Sensitivities grow exponentially large because the linearized model lacks the nonlinear interactions that otherwise slow or stop the exponential growth once the perturbations reach finite amplitude (Zhu et al. 2002). These strong sensitivities have been associated with positive Lyapunov exponents which characterize the limits of predictability of the model (Köhl and Willebrand 2002).
Likewise, the sensitivities obtained from the adjoint model become unrealistic when
the integration is long enough that the corresponding tangent linearization is not valid
for finite perturbations (Lea et al. 2002; Zhu et al. 2002). Although the adjoint is correct
for sufficiently small parameter adjustments, in a practical assimilation problem the size
of the adjustments will be similar to the size of the model state uncertainties (Errico and
Reader 1999), which are not generally infinitesimal.

Once the model is integrated beyond 2 or 3 months, the range for which the
linearization remains valid becomes much smaller than the range of possible values
for the control parameters. Therefore, the large gradients do not provide any useful
information for the optimization of the cost function or for sensitivity studies over long
periods (the actual length depends on the system under study). If these instabilities are
not damped, the assimilation time range must be limited so that a linearized model
remains valid and the adjoint sensitivities can be useful over finite-size optimization
steps. This is too short to optimize all the atmospheric forcing fields, which can require
years of integration time before the effects of uncertainties in the forcing fields are
visible above the noise in the observations. This happens both because the forcing effects
at the surface accumulate and because the effects penetrate to greater depths where the
signal-to-noise ratio is larger. Surface forcing can take years to affect the ocean below
the surface mixed layer, particularly outside the equatorial waveguide. Because model
dynamical errors are not included in the control vector, these deep regions are controlled
only by the initial condition, and accurate measurements there should be able to isolate
model errors by showing uncontrollable misfits contributing to the cost function. As
mentioned above, large misfits were not obvious in the integrations.

Exponentially increasing sensitivities might be treated by preconditioning as in section
2(c), by a separation into stable and unstable modes of the adjoint. However, since
these modes depend on the state and change with each iteration no fixed preconditioning
can be found. Imposing full covariance matrices for the controls might help attenuate
high sensitivities through the application of appropriate preconditioning (Weaver et al.
2003), but for short periods only since unstable sensitivities grow exponentially in time
without any limit. Over long periods, smoothing of the adjoint will only mix the strong
sensitivities together (Vögeler and Schröter 1995). In summary, the problem with strong
sensitivities is not that they cause roughness in the solution, but that they restrict changes
in the control parameters. This is because large adjoint values mean that small changes in
the control parameters make huge changes in the estimated data, and so changes in the
control parameters are strongly restricted. In experiments using the adjoint with normal
dissipation, the cost function barely decreased (<1% per iteration). If these sensitiv-
ities were real, meaning that they obtained for finite-amplitude perturbations as well
as infinitesimals, this would mean that the error bars on the estimates of the control
parameters would be very small. In the tropical Pacific example, these sensitivities do
not apply to finite perturbations, and so they must be eliminated in order to allow the
descent method to proceed with reasonable step sizes.

To extend the limits of the adjoint method in the presence of nonlinearity, Lea
et al. (2002) recently proposed averaging the sensitivities of several adjoint runs in
order to filter the effect of secondary minima. Such an ‘ensemble adjoint approach’
could become computationally prohibitive for long assimilation periods since this might
require a large increase in the number of adjoint runs to filter out very large sensitivities.
Another approach replaces the original unstable adjoint model with the adjoint of a
tangent linear model which has been modified to be stable. This involves a simplification
of the original tangent linear model and is usually obtained by omitting strongly unstable
modes. The gradients calculated by the modified adjoint do not ‘see’ the secondary
minima and approximate the full gradients to the envelope of the cost function. This is similar to performing the optimization in a smooth subspace, which means loss of accuracy due to the omission of strongly nonlinear variations, generally related to small-scale phenomena. This is necessary, however, in order to optimize the cost function over long assimilation periods. For instance, Köhl and Willebrand (2002) successfully optimized the cost function of their highly nonlinear variational assimilation problem by constructing an adjoint model of the mean state that was integrated on a coarser grid employing larger mixing than the forward model. It is also common to simplify the adjoint associated with the turbulence closure models by disabling the physical terms that were the main sources of instability (Zhu et al. 2002; Weaver et al. 2003). More related to the current study, Köhl (2005, personal communication) has found that the inherent nonlinearity of KPP (the mixing coefficient depends on the state) leads to the development of linearly unstable modes after integrating beyond a few weeks. A comparison of the adjoint with and without KPP during the first few weeks showed that the adjoint of KPP makes only minor contributions to the gradient before the large sensitivities appear. Removing KPP from the adjoint eliminates the instability with apparently minimal impact on the descent. This simplification was adopted here. Since KPP is enabled in the forward runs, the reduction of the cost function is a check on the validity of this approach.

In the tropical Pacific, there are three regions with strong sensitivities, which we suggest are related to the growth rates of intrinsic variability. These regions all exhibit large relative vorticity, indicating current shear, and show relatively low gradients of potential vorticity (PV). The Charney–Stern condition requires a sign change of the PV gradient and regions with small PV gradient are expected to be disposed to instability (Pedlosky 1987). In all three regions, the growth of the Perturbation Kinetic Energy occurs through both barotropic and baroclinic instability, as in observations (Qiao and Weisberg 1998) where the growing perturbations extract energy from the low-frequency flow. The relative contributions of baroclinic and barotropic energy varies with region. In the eastern and central Pacific, the energy source is primarily barotropic instability, feeding off horizontal current shear between the SEC and the NECC in regions of low potential-vorticity gradient (low stability), and baroclinic instability associated for example with the TIWs on both sides of the equator (Baturin and Niiler 1997; Contreras 2002). In the western Pacific, baroclinic instability dominates the growth as seen by Tozuka et al. (2002), apparently including, in this model at least, interactions with bottom topography producing significant vertical velocity fields.

We first examined the nonlinearities by carrying out several forward runs to compute first and second derivatives of the model state with respect to various control variables using finite differences in control parameter space. The second derivative is an indicator for the nonlinearity of the dynamics of the perturbations and thus for the breakdown of the assumption that sensitivities can be represented by linearized models. Moreover, since the model remains stable they also indicate regions of large nonlinear damping of the linear growth. To compute the differences we performed three forward runs in the 1998 assimilation period in which the zonal wind stress $\tau$ was perturbed as follows:

(i) R1: A forward run with reference wind stress.

(ii) R2: The same run as R1 but with a spatially and temporally constant perturbation $(size = \beta = 10^{-4} \text{ N m}^{-2})$ on the wind stress.

(iii) R3: The same run but with perturbation size $= -\beta$. 

The approximate first and second derivatives of the SSH at every point with respect to this constant (in space and time) zonal wind-stress perturbation were estimated by, respectively, subtracting $R_3$ from $R_2$, and adding the $R_2$ and $R_3$ solutions and subtracting twice the $R_1$ solution. The left column of Fig. 1 shows the time evolution of the second derivatives normalized by $\beta$ and the r.m.s. of the first derivative spatially smoothed over a radius of six grid points. This is a comparison of the sizes of the first two terms in the Taylor expansion

$$\frac{\Delta \text{SSH}}{\Delta \tau} = \frac{\partial \text{SSH}}{\partial \tau} + \frac{\beta}{2} \frac{\partial^2 \text{SSH}}{\partial \tau^2} + \ldots,$$

so magnitudes larger than 1 indicate significant nonlinearity. The model SSH sensitivity appears to be linear during the first month. Nonlinearities appear in the western Pacific and in the central Pacific near the equator after 40 days. They spread over the domain and become stronger as the integration time increases. Small nonlinearities in the eastern Pacific are probably due to weak TIWs during the first half of 1998, which is the end of an El Niño period (Contreras 2002).
The time evolution of the gradient (from the adjoint model) of the total cost function with respect to the zonal wind stress is shown in the right column of Fig. 1. This suggests that the adjoint model is stable over roughly the same time range as the dynamics of the perturbations remain linear. Because the time of the adjoint is running backwards from the end of 1998 and that of the forward model from the beginning of 1998, the linearization and the associated unstable modes are not identical. In addition, the growth of the unstable modes is forced by the model-data differences in the adjoint, but by perturbing the wind field in the forward model. Although these two different ways of triggering the most unstable modes usually result in slightly different timings, the qualitative agreement between the left and right columns of Fig. 1 suggests that exponentially growing sensitivities in the adjoint run are an indicator for strong nonlinearity of the forward model. The nonlinearities seen in the western Pacific during the forward run correspond to limiting of the linear growth. This indicates that the linear infinitesimal sensitivities do not correctly describe the sensitivity of the model to finite perturbations. As suggested by Köhl and Willebrand (2002), the existence of secondary minima that prevent the convergence of the adjoint method can also be diagnosed from an exponential increase of the norm of the adjoint variables. Figure 2 shows the natural logarithm of the Euclidean-norms of the adjoint variables and shows a nearly exponential growth.

Next, we examined the norm of the adjoint variables when using larger viscosity and diffusivity terms. Comparing to the run with the regular viscosity and diffusivity (Fig. 2), the norm of the adjoint solution appears to be stabilized when the viscosity and diffusivity were set to $3 \times 10^{12} \text{ m}^4\text{s}^{-1}$ and $1.5 \times 10^4 \text{ m}^2\text{s}^{-1}$ (30 times the original values), respectively. To check that the disappearance of the large adjoint sensitivities comes from increased linearity, we performed the same forward runs as described before but using the higher viscosity and diffusivity terms. These runs confirmed (not shown) that the viscous model remains almost linear to these perturbations as hypothesized from the results of the backward run. In summary, at 30 times viscosity and diffusivity, the model remains nearly linear during the first 12 months and the adjoint is stable.
It is still necessary to check that increasing the viscosity and diffusivity extends the allowable integration time of the adjoint-based assimilation system without rendering the adjoint useless for an assimilation which uses forward model runs with the original viscosity and diffusivity terms. Figure 3 shows the gradients of the cost function with respect to the heat flux as obtained from the adjoint model for increasing values of the viscosity and diffusivity parameters. The use of larger values of these parameters appears to significantly reduce the high sensitivities while preserving the large-scale patterns. We hypothesize that the extra damping discards some ‘uncontrollable’ small-scale intrinsic variability, but makes it feasible to carry out the assimilation over longer periods. Increasing model dissipation to reduce nonlinearities is similar to the method of Köhl and Willebrand (2002) but it was simpler to implement in this case. However, because the adjoint model is run at full resolution, it is more demanding in computing resources. The remaining (and most important) test is to confirm that the gradients calculated using increased dissipation produce useful descent in the iterations, which will be discussed in section 4. Using viscosity for damping out uncontrollable variability further allows for a multi-scale approach, if small-scales are of interest. Such a procedure is similar in spirit to that of Pires et al. (1996), but where the assimilation starts with high viscosity and diffusivity and runs over the maximum time range. Once the uncertainty in the control parameters has decreased, the adjusted values can be used as the starting guesses for successive runs with decreased values of the viscosity and diffusivity parameters.

3. Description of the datasets

The observational datasets used in this study are described in this section.

(a) Altimetric data

SSH measurements were provided by the TOPEX/POSEIDON (TP) mission. To eliminate errors associated with uncertainties in the geoid, the mean and time-varying
components of the SSH data were considered separately. The mapped 1998 TP mean SSH minus the Earth Gravitational Model 1996 (EGM96) geoid (Lemoine et al. 1997) has been used to constrain the mean model SSH during the assimilation. For the time-varying component, along-track daily TP data obtained from NASA’s PO-DAAC at the Jet Propulsion Laboratory and processed as described by Stammer and Wunsch (1994) were fitted at the observation points.

(b) Analysed data

Although ocean climatologies are not usually considered ‘observations’, they provide a pre-smoothed analysis of observations and they can be used to constrain the model solution in test cases. Here, the Levitus climatology of monthly mean T and S (Levitus and Boyer 1994), and the Reynolds monthly SST (Reynolds and Smith 1994) were used. The Levitus climatology is based on historical hydrographic data that are merged and spatially averaged. The Reynolds optimum interpolation SST analyses are produced on a one-degree grid using buoy and ship data as well as satellite SST data. All these data were interpolated onto the model grid first horizontally and then vertically using linear interpolation procedures.

(c) Cost function

In this test of the assimilation system, the time-mean dynamic topography TP—EGM96 (SSH<sub>TP—EGM96</sub>), daily TP SSH anomalies (SSH<sub>TP</sub>'(t)), monthly Reynolds SST (SST<sub>Rey</sub>(t)) and monthly Levitus T (T<sub>Lev</sub>(t)) and S (S<sub>Lev</sub>(t)) analyses were assimilated. The model mean SSH (SSH) and daily mean SSH anomalies (SSH'(t)) were constrained to altimetric SSH data while monthly means of the model SST (SST(t)), T (T(t)) and S (S(t)) were constrained to match the climatological T and SST. The explicit form of the model-data misfit term in the cost function is then

\[
J_{\text{obs}} = (\text{SSH} - \overline{\text{SSH}}_{\text{TP—EGM96}})^T R_{\text{TP—EGM96}}^{-1} (\text{SSH} - \overline{\text{SSH}}_{\text{TP—EGM96}}) + \sum_{t=\text{days}} (\text{SSH}'(t) - \overline{\text{SSH}}_{\text{TP}}(t))^T R_{\text{TP}}^{-1} (\text{SSH}'(t) - \overline{\text{SSH}}_{\text{TP}}(t))
\]

\[
+ \sum_{t=\text{months}} (\text{SST}(t) - \overline{\text{SST}}_{\text{Rey}}(t))^T R_{\text{Rey}}^{-1} (\text{SST}(t) - \overline{\text{SST}}_{\text{Rey}}(t))
\]

\[
+ \sum_{t=\text{months}} (T(t) - \overline{T}_{\text{Lev}}(t))^T R_{\text{Lev}}^{-1} (T(t) - \overline{T}_{\text{Lev}}(t))
\]

\[
+ \sum_{t=\text{months}} (S(t) - \overline{S}_{\text{Lev}}(t))^T R_{\text{Lev}}^{-1} (S(t) - \overline{S}_{\text{Lev}}(t)).
\]

(6)

4. Assimilation experiments

(a) Set-up

To evaluate the performance of the system and to study its sensitivity to different set-ups, we carried out several experiments over a one-year assimilation period starting from 1 January 1998. The control variables were the initial conditions, the atmospheric forcing fields adjusted every two days, and boundary conditions adjusted every week. Data and model errors were prescribed only on the diagonal of the error covariance matrices and are the same as used by Köhl (2005, personal communication) in the global ECCO state estimation. The error bars for T and S (both for data and initial conditions) were taken from the uncertainties given with the Levitus climatology. A uniform r.m.s.
error of 3.5 cm was used for the SSH observations. Prior uncertainties for the wind stress were taken from the standard deviations of the differences between NCEP and QuickSCAT scatterometer wind fields. One-third of the local standard deviation (over time) of the NCEP forcing was used as the prior error for the net heat and freshwater fluxes. For the boundary conditions, Levitus errors were used for $S$ and $T$, and the standard deviation (over time) of the ECCO velocities was used at the boundaries for $U$ and $V$ uncertainties. Error bars for the baroclinic modes of the normal velocities at the boundary were estimated from the standard deviation of the same modes of the ECCO velocities. The weight for the barotropic mode was empirically set to a small value as described above.

The absence of cross-correlations (non-diagonal terms) in the data error covariance matrices means that the data are treated as independent. Closely spaced data can then be strongly weighted in the cost function, which could result in over fitting by the optimization. This is particularly true for the analysed Levitus and Reynolds data, which were interpolated to model grid points. This is inappropriate for the estimation of the true state of the tropical Pacific, but we only focus here on the ability of the assimilation system to fit the test datasets within the margin of prescribed uncertainties.

The descent directions towards the minimum were iteratively determined using the Quasi-Newton M1QN3 algorithm which has been developed by Gilbert and Lemaréchal (1989). After thirty iterations the rate of cost-function decrease was relatively small and the differences between the model and the observations were reduced to about the levels of expected uncertainty for most of the data. We started the optimization using either Levitus climatology or ECCO analyses as the initial conditions, NCEP or ECCO analyses for the atmospheric forcing fields, and ECCO analyses for the open boundaries. Three assimilation runs were performed as summarized in Table 1 and compared to the model run with the starting guess for the control parameters (which will be called the ‘reference run’).

**TABLE 1. SUMMARY OF THE ASSIMILATION RUNS**

<table>
<thead>
<tr>
<th>Runs</th>
<th>First guess</th>
<th>Initial conditions</th>
<th>Forcing</th>
<th>Open boundaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCEP-LEVic</td>
<td>Levitus</td>
<td>NCEP</td>
<td>ECCO</td>
<td></td>
</tr>
<tr>
<td>ECCO-LEVic</td>
<td>Levitus</td>
<td>ECCO</td>
<td>ECCO</td>
<td></td>
</tr>
<tr>
<td>ECCO-ECCOic</td>
<td>ECCO</td>
<td>ECCO</td>
<td>ECCO</td>
<td></td>
</tr>
</tbody>
</table>

(b) Assimilation results

(i) Cost function. First we discuss the decrease of the total cost function as well as of particular parts of the data misfits for each of the three runs NCEP-LEVic, ECCO-LEVic and ECCO-ECCOic (Table 1). The various cost functions vs. iteration count are shown in Fig. 4. The initial total cost function is smallest when the assimilation starts from the ECCO forcing, showing that these forcing fields, which were optimized on a $1^\circ \times 1^\circ$ grid, have skill for SSH and SST in the higher-resolution model. Using Levitus as a first guess for the initial conditions also provides a better initial cost function since the $T$ and $S$ terms account for the dominant values in the total cost function. The decrease in the total cost function is greatest during the first iterations, particularly for the NCEP-LEVic run. However, after 5 iterations the overall performance of the NCEP-LEVic run, as measured by the cost function, is as good as the run starting from the ECCO forcing. Furthermore, the cost function seems to reach about the same level for both runs. The rate of cost function decrease for the ECCO-ECCOic run is rather slow and the optimization is unable to adjust certain features present in the ECCO initial
Figure 4. (a) The total cost vs. iteration number, and the contribution from the components of the individual misfits of the observed variables taken from the assimilation runs starting from NCEP forcing and Levitus initial conditions (IC) (black curve), the assimilation runs starting from ECCO forcing and Levitus IC (dashed curve), and the assimilation runs starting from ECCO forcing and ECCO IC (grey curve). (b)–(e) The cost function evolution for (b) temperature, (c) salinity, (d) SST, and (e) SSH. Note that in several cases the lower limits of the cost function axes are not zero.
conditions after 30 iterations. This was found to be due to large differences between the ECCOic and LEVic in the western Pacific north of the equator and in the cold-tongue area. In both these regions, the specified background errors were too small to allow the optimization to make large enough adjustments to the ECCO initial fields in order to fit Levitus data. In the western Pacific in particular, a few large anomalies were seen in the climatology near the coast. These have since been corrected, but degrade the overall $T$ and $S$ fit somewhat in these examples. For LEVic, after 30 iterations, the total cost function is reduced by more than 50% starting from ECCO forcing, and 75% starting from NCEP forcing. Overall, similar conclusions can be made from the individual cost terms, although the rate of decrease of the salinity term is slower than the other data terms. The weighted misfits for salinity in the reference run are one half the size of the weighted temperature misfits, so they are reduced more slowly. In addition, the vertical motions of the tropical pycnocline have a smaller weighted effect on salinity than on temperature, due to the relatively low vertical salinity gradient at the pycnocline. Thermohaline anomalies on isopycnals must be advected and so spread more slowly than density and current anomalies, which can propagate in the equatorial waveguide. The improvements made to the data cost function terms after the 20th iteration are rather small and they tend to be balanced by the increase of the cost function terms for the control variables.

We focus now on the results from the ECCO-LEVic assimilation (starting from ECCO forcing and Levitus IC). The temporal and spatial distributions of cost-function terms for selected data types from the ECCO-LEVic run are plotted in Fig. 5. The curves have been normalized by the number of observations in the sum over time and/or space, so a value of 1 would roughly indicate that the solution fits the data within the specified uncertainties. A quantitative discussion of Chi-squared consistency (e.g. Bennett et al. 1998) was omitted since it requires a linear system and perfectly known, independent Gaussian distributions of the control and observation uncertainties, which is not met in our case. The monthly averaged reference cost function for Levitus temperature is small in January, since the assimilation starts from the ‘data’ field, and then increases over time as the model drifts from Levitus. The assimilation reduces the drift and brings the model close to the observations over the entire assimilation period. The daily SSH cost term has a ‘U’ shape and again the assimilation is also able to control and improve its skill over the entire assimilation period, although the normalized misfit variance is greater than 1 for the SSH. The spatial contributions to the cost function summed over the entire interval are shown for the salinity term, and the model/data misfit is reduced over the entire domain. Normalized misfit variance remains high near Indonesia, although it is less than 1. This is due to complicated topographic interactions in the region, the slow adjustment of salinity in the west, and from anomalies in the climatology fields near the coast as previously mentioned. The results for the other data cost terms (not shown) are generally similar.

(ii) Estimated state. To assess the fit to the assimilated observations in a more physical way, we compared a zonal cross-section of the mean temperature field at the equator from the reference and the assimilation runs to the Levitus field (Fig. 6(a)). The thermocline from the reference run is shallower than in the data in the western part of the Pacific and deeper in the eastern part. The thermocline tilt (and the EUC that moves along it) are thought to be strongly influenced by the wind stress, since inviscid wind-driven thermocline models reproduce many features of the observations (Pedlosky 1987; McCreary and Lu 1994). In oversimplified terms, it is possible that the NCEP wind stress used in the reference case is too weak and produces a weaker EUC and a smaller
zonal thermocline tilt than is observed. The assimilation successfully corrects the tilt of the thermocline and significantly improves the agreement of the temperature field with climatology, even in the deep layers. This is to be expected from the reduction in the temperature errors seen in the cost function, since these are not independent measurements. Assimilation does not, however, improve the strength of the EUC, which remains too weak (not shown) probably because the mean SSH gradient was weakly constrained due to large uncertainties in the geoid data and the Levitus climatology was too smooth to provide much information on subsurface velocities.

The variability of the SSH from the model (with and without assimilation) was compared to the gridded Archiving, Validation and Interpretation of Satellite Oceanographic (AVISO) SSH variability. Figures 6(d)–(f) give maps of the SSH standard deviation from these solutions. The spatial structures of the SSH variability after assimilation is in better agreement with the analysed AVISO variability than the variability of the reference run. The unrealistically strong variability over the NECC area in the reference run is completely removed and shifted to the eastern Pacific after assimilation. The ECCO forcing has strong wind-stress curl in the NECC region, forcing a stronger NECC than is observed and contributing to the variability. The assimilation reduces this curl, reducing the NECC current shear, which is a source of energy for the variability. The variability of the assimilated field, however, is still rather weak in the western Pacific near the
open boundaries. Note that the assimilation introduces small-scale features to the SSH variability compared to the reference solution. As mentioned above, the normalized SSH misfit exceeded the prior error bars, since it had less weight in the solution than $T$ and $S$.

As an independent evaluation of the state estimate from the assimilation, we compared the optimized state to observations which were not used in the assimilation. Comparison with high-resolution expendable bathythermograph (XBT) data (Roemmich et al. 2001) (not shown) shows that the reference run is in better agreement with the XBT data than the assimilation run, particularly in the thermocline. We hypothesize that these differences arise from constraining the model to fit the Levitus climatology, which has large differences from observations during the 1997–98 El Niño event. This suggests that we are able to obtain a simultaneous fit to Levitus climatology and 1998 SST and SSH data which should not be consistent with the climatology. Comparison of the mean zonal currents with the TAO measurements (McPhaden et al. 1998) along the equator (not shown) also shows the reference run in better agreement with the observations in the western Pacific where the assimilation deepens the thermocline while weakening the EUC. In the wind-driven theory (Pedlosky 1987; McCreary and Lu 1994) the absolute SSH gradient along the equator is related to the strength of the EUC. Because we assume a large uncertainty for the geoid in this assimilation, absolute SSH gradients are poorly determined. These results suggest the importance of both geoid and subsurface
Figure 7. The time-mean ECCO adjustments (N m$^{-2}$) to (a) the NCEP net zonal wind stress, (b) the time-mean difference with respect to NCEP of the zonal wind stress from the assimilation that started from NCEP forcing and Levitus IC, and (c) the time-mean difference with respect to NCEP of the zonal wind stress from the assimilation that started from ECCO forcing and Levitus IC. (d)–(f) Same as (a)–(c) but for the time-mean net heat flux (W m$^{-2}$) (note different scale in (e) to show detail). See text for further explanation.

data in an assimilation, since if the surface data alone were sufficient to determine the model state, the agreement with the $T$ and $S$ climatology should not have been possible. This is only a hypothesis, though, since the normalized SSH misfit exceeded the prior uncertainty bounds, and the slowed descent rate of the iteration may be a symptom of the incompatibility.

(iii) Adjusted forcing and open boundaries. Here we compare the mean adjustments to the NCEP analysis for the zonal wind stress and the net heat flux in Fig. 7, and the mean adjustments to the ECCO analysis for the temperature and the normal velocity at the southern open boundary in Fig. 8. These result from the separate assimilation runs NCEP-LEVic and ECCO-LEVic. The mean ECCO adjustments to NCEP forcing are plotted in Figs. 7(a) and (b) to show the differences between the two.

Whether starting the optimization from NCEP or ECCO forcing, the assimilation reaches visually similar solutions for the mean zonal wind stress after 30 iterations. Starting from the NCEP forcing, the model requires strong adjustments on both sides of the equator probably related to the overturning circulation and thermocline adjustment. The assimilation also removed the strong ECCO adjustments over the NECC and in the western Pacific over the SEC. The agreement between the optimized fluxes is poor for the net heat flux (and net freshwater flux; not shown). More precisely, the adjustments
to the heat flux were generally less than 10 W m\(^{-2}\) whether the optimization starts from NCEP forcing or ECCO forcing. Relative to NCEP, similar patterns were obtained from both runs, with less heat over the cold tongue and more heat over the subtropical gyres. However, the heat flux adjustments (relative to NCEP) starting from NCEP were only a fifth of those starting from ECCO. This is because the mean difference between ECCO and NCEP heat flux was of the order 50 W m\(^{-2}\).

This result can be rationalized by arguing that the wind stress drives most of the equatorial circulation, and the penetration of wind forcing changes can be large near the equator. The sensitivities with respect to the wind were therefore the largest and the optimization adjusted this control variable first. After roughly 25 iterations, the cost-function terms for the other control variables started to increase, once the main form of the wind adjustments was set. More iterations are probably needed for the stabilization of the heat and freshwater fluxes. Another reason for the slow adjustment of the heat flux is the relatively short assimilation window since the impact of this flux on the tropical circulations takes several years to emerge.

Finally, the decomposition of the normal velocities into barotropic and baroclinic modes efficiently solved the problem of huge sensitivities with respect to the barotropic component allowing the adjustments of the velocities with the rest of the control variables. Moreover, the adjustments to the ECCO state analysis at the southern boundaries are shown to approach similar solutions for both NCEP-LEVic and ECCO-LEVic runs in Fig. 8. The open boundary adjustments are not very large for any of the variables and this is perhaps due both to the short length of the assimilation window and to the relatively limited number of iterations.
5. Discussion

We have described an eddy-permitting four-dimensional adjoint data assimilation system for the tropical Pacific Ocean that is nested in the global ECCO assimilation results. The implementation of this system was not simple due to two main difficulties: the control of the normal velocities at the open boundaries and the nonlinearity of the model. A vertical mode decomposition was introduced for the open boundary velocities in order to down-weight only the barotropic component in the optimization. Higher viscosity and diffusivity terms were used in the adjoint model to damp small-scale instabilities associated with local minima allowing the optimization of the cost function over year-long assimilation periods.

Several experiments were performed over a one-year period in 1998 to evaluate the performance of the assimilation system while constraining the model to altimetric SSH, Reynolds SST and monthly Levitus $S$ and $T$. The system was shown to be able to improve the model fit to these data while providing estimates weakly sensitive to the optimization first guess, providing similar final solutions except where the constraints on the IC adjustments were too strong. The system closely fits the Levitus climatology, which degrades the agreement with XBT observations in comparison to the reference run. This possibly resulted from the use of relatively small uncorrelated errors for the Levitus data in the optimization and because no other subsurface data were included in the data assimilation. Reduced weights for the climatological data and additional smoothness constraints for the forcing should produce better estimates. In this case, the small errors highlight the controllability of the $T$ and $S$ structure and the compatibility of Levitus $T$ and $S$ with observed SSH and SST, within relatively coarse error bars. We hypothesize that the inclusion of subsurface data is helpful to produce an unambiguous model state, at least when the assimilation period is relatively short. The large geoid uncertainties may contribute to the errors in the EUC.

The purpose of these experiments was to test the convergence of the fit and to explore the interplay of surface and subsurface data in specifying the ocean state. Other datasets (obtained from the TAO array, XBT and ARGO networks, and Drifters) will be included in the system in the future with improved error covariance matrices to produce analysed states and forcing fields for the tropical Pacific Ocean.

Acknowledgements

Thoughtful comments and suggestions by the associate editor Anthony Weaver and two anonymous reviewers are gratefully acknowledged. We would also like to thank Patrick Heimbach and Geoffrey Gebbie for assistance and valuable discussions. Reanalysis surface forcing fields from NCEP/NCAR are obtained through a computational grant at NCAR. Computational support from the National Partnership for Computational Infrastructure and NCAR is acknowledged. This work was supported by National Oceanic Atmospheric Administration through the Consortium for the Ocean Role in Climate and in part by the Consortium for ECCO funded by the National Oceanographic Partnership Program.

References

Bennett, A. F. 2002 Inverse modeling of the ocean and atmosphere. Cambridge University Press


<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Title</th>
<th>Reference</th>
</tr>
</thead>
</table>