Continuous data assimilation of drifting buoy trajectory into an equatorial Pacific Ocean model

M. Kamachi a, J.J. O’Brien b

a Oceanographic Research Division, Meteorological Research Institute, Japan Meteorological Agency, 1-1 Nagamine, Tsukuba 305, Japan
b Center for Ocean and Atmosphere Prediction Studies, The Florida State University, 020 Love Building, Tallahassee, FL 32306-3041, USA

Received 20 September, 1993; revised and accepted 7 June, 1994

Abstract

A variational adjoint method is developed to assimilate the trajectories of drifting buoys into a simple ocean model. The method is applied to the equatorial Pacific Ocean. In the variational formalism a cost function measuring the distance between the trajectories of the model simulated buoys and of the observed buoys is minimized. Adjoint equations are forced by the model-data misfit that is proportional to the difference of model simulated and observed buoy positions. A hypothetical mean upper layer thickness of the model is used as a control parameter. The optimal spatial structure giving the best fit of the model to the observations is determined, and it is used as the initial condition of the model. Model simulated data are used first in several experiments and later experiments with real observations are discussed. Assimilation periods and regional effects with relations of buoy trajectory and wave propagation on the assimilation processes are examined. Upper layer thickness in the western equatorial region is improved relatively better than in the eastern region for the assimilation period of three months, because the western equatorial region acts as a special "caustic" region in the adjoint system. This east–west inhomogeneity vanishes using a one-year period. The upper layer thickness optimized with real buoy data is compared with independently observed sea level data.

1. Introduction

There is growing evidence that the ocean exhibits low frequency phenomena associated with climate variability. A good knowledge of the upper ocean state is needed in order to make forecasts of the climate variability. Because the tropical ocean is mainly deterministic, knowledge of the forcing is critical to determine ocean state. The forcing is, however, incompletely known and ocean models are not perfect. Also there are some important nonlinear phenomena. Therefore ocean observations must also be used. Because the ocean is sparsely observed, methods must be designed to extract the maximum amount of information from the available observed data.

There are two distinct specifications to describe the ocean circulation (see, e.g., Pickard and Emery, 1990). The first is called the Eulerian description in which physical quantities are defined as functions of position in space and time. This Eulerian description provides a picture of
the spatial distribution of physical quantities (velocity, temperature, salinity, etc.) at each time. The second is called the Lagrangian description in which physical quantities relate to time and certain initial positions of identifiable fluid particles (or drifting matter) such as in dynamics of particles. The physical quantities are functions of time and the fluid particle identification. This method provides a history of dynamical trajectory. Although drifting buoys have a two-dimensional movement in a horizontal plane within an averaged depth (e.g., 15 m) of the ocean, we may regard them as a kind of quasi-Lagrangian particles in the two-dimensional plane. The Eulerian variables are mostly used in numerical studies, but the Lagrangian drifting trajectory data have been useful in describing the ocean circulation. Drifting buoys provide both track and speed of flow, and have also revealed many new details of eddies associated with ocean currents. Buoy data satisfy some necessary conditions for the ocean prediction: near real time and continuous data acquisitions. As for the data assimilation, adoption of the buoy data still seems to be necessary. The data of the Eulerian variables (e.g., temperature and sea level height) have been used in recent developments of data assimilations in oceanography. The inherent variability of the structure of the ocean current system may be understood by data assimilation of the Lagrangian trajectories of drifting buoys. Therefore our aim is a development of a new data assimilation method, which will be explained in Section 3, with the Lagrangian trajectories of the drifting buoy and a simple ocean model.

The variational adjoint adjustment technique of data assimilation based on optimal control methods have been developed recently. This method was pioneered by Sasaki (1958, 1969, 1970a, b) and Marchuk (1974), and have been developed and applied in the oceanography, for example, for initialization with XBT data (Sheinbaum and Anderson, 1990a, b), with altimeter data (Moore, 1991), for parameter estimation with sea level (Smedstad and O'Brien, 1990), with velocity and wind data (Yu and O'Brien, 1991). Recent reviews of this method and applications in the oceanography and meteorology are given by Ghil and Malanotte-Rizzoli (1991), Anderson (1991), Daley (1991), Thacker (1992), and Bennett (1992).

Many problems about the initialization with the adjoint method in the equatorial ocean were discussed in detail in Sheinbaum and Anderson (1990a, b). They used a linear reduced gravity equatorial model, its adjoint and temperature data. Their aim was to obtain an optimal initial condition. The model was unable to match the data in the eastern Pacific for an assimilation period of six months. They discussed east–west contrast in the optimization process and inconsistency in a linear model equation. The inconsistency is due to unbalance among the given phase speed of the gravity wave, pressure gradient determined from the assimilation procedure and the given wind forcing. Those problems were carefully discussed by Sheinbaum and Anderson (1990a, b), however their approach is quite different from ours. We discuss a similar problem where difficulties generally arise in the equatorial data assimilation, with a different model (nonlinear one), a different assimilation scheme (parameter estimation) and moving observation points (drifting buoys). Because the nonlinear model used here does not contain the mean upper layer thickness in the pressure terms explicitly, the unbalanced problem noted by Sheinbaum and Anderson does not arise.

In the variational formalism, described later in detail in Section 3, a cost function measuring the distance between the model and observation states is minimized with a strong constraint, i.e., model equations. The augmented Lagrangian is constructed as a sum of the cost function and the inner product of the model equations and the Lagrange multipliers. In order to minimize the cost function the first variation of the Lagrangian is taken to zero. It yields the original model and adjoint equations that governs the Lagrange multipliers. In order to minimize the cost function the first variation of the Lagrangian is taken to zero. It yields the original model and adjoint equations that governs the Lagrange multipliers. Adjoint equations are forced by the model-data misfits, which are partial derivatives of the cost function with respect to model variables. Used with a gradient information of the Lagrangian with respect to free control parameters and an iterative descent algorithm, the minimum of the cost function and the optimal state
are able to be found. The optimal state satisfies the constraint dynamics in space and time exactly and is close to the observations in a least square sense. Furthermore, we can assimilate any type of observations, if and only if the observed variable can be functionally related to the model variables.

A new variational adjoint method for the drifting buoy trajectories is developed in this study. The method is applied to the equatorial Pacific Ocean. The trajectories have a useful large scale and low frequency information, but extracting the information is difficult. The structure of this paper is as follows: We describe the equatorial Pacific Ocean model and its hindcast in Section 2. In Section 3 description of the new assimilation method is given with a derivation of the adjoint equations. Section 4 describes assimilation experiments and results. Section 5 is devoted to discussion and summary.


Equatorial ocean models have been successful and reproduced many variations of observed sea level changes. This research work was originated by Busalacchi and O'Brien (1980). Nonlinear effects, however, have been reported to play an important role in many papers (e.g., Hurlburt et al., 1976; Cane, 1979; Philander and Pacanowski, 1980). Also a number of theoretical and numerical studies of nonlinear equatorial waves have been reported (e.g., Boyd, 1980; Matsuura and Yamagata, 1982; Greatbatch, 1985). In this study we use the Lagrangian trajectories of drifting buoys that contain nonlinear eddy phenomena. We therefore adopt a nonlinear model with fine resolution.

Let $U = uh$ and $V = vh$ represent the eastward and northward components of the upper-layer transport, respectively, where $(u, v)$ are the depth-independent eastward and northward velocity components in the upper layer, and $h$ the upper layer thickness (ULT). The model used is a simplified form of a reduced gravity, nonlinear transport shallow-water model for the equatorial Pacific Ocean in spherical coordinates in which $\phi$ is longitude and $\theta$ is latitude:

\[
\frac{\partial U}{\partial t} + \frac{1}{a \cos \theta} \frac{\partial}{\partial \phi} \left( \frac{U^2}{h} \right) + \frac{1}{a^2 \cos^2 \theta} \frac{\partial g' h}{\partial \phi} \frac{\partial h}{\partial \phi} - \tau(\phi) - \left( \frac{A}{\rho} \right) U = 0
\]

\[
E_n = \frac{\partial V}{\partial t} + \frac{1}{a \cos \theta} \frac{\partial}{\partial \phi} \left( \frac{UV}{h} \right)
\]

\[
\times + \frac{1}{a \cos \theta} \frac{\partial}{\partial \phi} \left( \frac{V^2}{h} \right) + (2 \Omega \sin \theta) U + \frac{g' h}{a \cos \theta} \frac{\tau(\phi)}{\rho} - \frac{A}{a^2 \cos^2 \theta}
\]

\[
\times \left[ \frac{\partial^2}{\partial \phi^2} + \cos \theta \frac{\partial}{\partial \theta} \left( \cos \theta \frac{\partial}{\partial \theta} \right) \right] V = 0
\]

where $a$ is the radius of the Earth, $\Omega$ is its angular velocity, $g'$ is the reduced gravity and $A$ is a horizontal eddy viscosity. The wind forcing $\tau = (\tau(\phi), \tau(\phi))$ is modeled as a body force acting over the layer.

Eqs. (2.1)–(2.3) are integrated numerically on a sphere over an ocean basin extending between 25°N and 20ºS in latitude and from 124°E to 76°W in longitude. The horizontal resolution is 0.125º in both longitudinal and latitudinal directions. Real coastline geometry is used. We adopt the Arakawa C grid. The equations are integrated in time using a leapfrog scheme. We also used a forward difference scheme every seventy-seventh time step to eliminate the computational mode. The model time step is 30 minutes. The diffusive terms are computed with a DuFort-Frankel scheme. No-slip boundary condition is applied on
at Santa Cruz in Galapagos Island. We applied a twelve months moving average to both data for removing the seasonal cycle. Thin line shows the interannual variability in the model result, and the thick one shows the observed one. The data in the period of 1961 to 1964 are omitted in the model result because the model after the spinup adjusts to the real wind field in the period. This figure depicts that the model upper layer thickness has a good agreement with the observed sea level in the equatorial eastern and western (not shown) Pacific Ocean, and shows major warm and cold events. The model has a good performance, but not excellent. We, therefore, adapt a data assimilation of drifting buoy data for improvement of the model performance.

3. Description of the assimilation method

The assimilation scheme that will be used is a variational method based on the Lagrange multiplier technique. This scheme minimizes a given measure of the distance between the model and the observations. We call the measure the cost function. The model requires additional specification of a control parameter in order to obtain a unique solution. Prior information about the control parameter is useful. The control parameter may consist of some of the parameters in the model, the initial conditions, boundary conditions, or a combination of these. To determine the numerical value of the admissible control parameter we have to extract the maximum amount of information from the observed data using the variational method. We select as the control parameter a hypothetical upper layer thickness $H(\phi, \theta)$, which is a hypothetical coefficient of the pressure gradient terms in the momentum equations [see the paragraph of Eqs. (3.8) and (3.9) for more discussion].

The cost function is constructed as the square of a norm. We use trajectories of drifting buoys as observed data. The cost function $J$ is chosen as the quadratic function

$$J(x; H) = \int_\Sigma \left[ K_x (x - x^0)^2 + K_y (y - y^0)^2 + \frac{K_H}{2} (H - H_c)^2 \right] d\sigma,$$

where $K_x, K_y, K_H$ are the validity coefficients (Gauss' moduli), $(x, y)$ are the positions of the model simulated buoys in longitude and latitude, $(x^0, y^0)$ are the observed buoy positions, $H_c$ is an estimate of the upper layer thickness, and $d\sigma = a^2 \cos \theta \, d\phi \, d\theta \, dr$. The validity coefficients are in principle the inverse of the error covariance matrices. If the observation has a large error, the weight is small. This weighting tendency of the error covariance matrices affects only the data region. If we have no data in a region, we have no value in the cost function. Because ocean observations are unfortunately very sparse, we need to increase the number of data in the model-data misfit. In the last paragraph in this section, we will discuss this spreading of the information with relation to calculation of forcing terms in adjoint equations. By assuming that the errors are uncorrelated and equally weighted, the matrices are reduced to unit matrices multiplied by constants $K_x, K_y, K_H$.

The values of the validity coefficients $K_x, K_y, K_H$ were chosen unity for simplicity. In this study
we are examining the fundamental process of the data assimilation. We have made no attempt to introduce a more realistic weighting from observation errors by means of the inverse of the error covariance matrices, although this should be done in a full implementation of the assimilation technique with a realistic 3-dimensional general circulation model and its adjoint one.

The approach that will be followed here is based on the classical Lagrange multiplier technique. An augmented Lagrangian $L$ can be constructed as a sum of the cost function and inner product of the Lagrange multipliers and the model equations:

$$
L(U, V, h, \lambda_u, \lambda_v, \lambda_h; H) = J(x; H) + \int \left( \lambda_u E_u + \lambda_v E_v + \lambda_h E_h \right) \, d\sigma,
$$

(3.2)

where $\lambda$'s are the Lagrange multipliers for $U$, $V$ and $h$, respectively and $E$'s are the constraints of the forms of Eqs. (2.1)-(2.3). The constrained minimization problem is thus replaced by an unconstrained problem. Using this formalism it is insured that the optimized state will satisfy the constraints exactly.

The optimality condition is given by putting the first derivative of the Lagrangian zero. The first variations of $L$ with respect to $\lambda$'s give the original model equations $E$'s = 0 [Eqs. (2.1)-(2.3)] back. Letting the first variations of $L$ with respect to model variables vanish yields the adjoint equations:

$$
\frac{\partial \lambda_u}{\partial t} - \frac{1}{a \cos \theta} \left[ \frac{2U}{h} \frac{\partial \lambda_u}{\partial \phi} + \frac{V}{h a \theta} \frac{\partial}{\partial \theta} (\lambda_u a \cos \theta) \right] + \frac{V}{h} \frac{\partial \lambda_u}{\partial \phi} + (2\Omega \sin \theta) \lambda_u = \frac{1}{a \cos \theta} \frac{A}{\partial \theta} + \frac{1}{a^2 \cos^2 \theta} \frac{\partial^2}{\partial \phi^2} + \cos \theta \frac{\partial}{\partial \theta} \left( \cos \theta \frac{\partial}{\partial \phi} \right) \lambda_u - \frac{\partial J}{\partial U},
$$

$$
\frac{\partial \lambda_h}{\partial t} - \frac{g' h}{a \cos \theta} \left[ \frac{\partial \lambda_u}{\partial \phi} + \frac{\partial}{\partial \theta} (\lambda_u \cos \theta) \right] - \frac{V}{h} \frac{\partial \lambda_v}{\partial \phi} \frac{\lambda_v}{h} \frac{\partial}{\partial \theta} \left( \lambda_v \cos \theta \right) - \frac{\partial J}{\partial V} = \frac{1}{a \cos \theta} \left\{ \frac{U}{h} \frac{\partial \lambda_u}{\partial \phi} + \frac{V}{h} \frac{\partial \lambda_v}{\partial \phi} \right\} + \frac{V}{h} \left[ \frac{\partial \lambda_u}{\partial \phi} + \frac{\partial}{\partial \theta} \left( \lambda_u a \cos \theta \right) \right],
$$

(3.4)

where the model-data misfit forcing terms are

$$
\frac{\partial J}{\partial U} = K_x (x - x^0) \int \frac{dt}{h},
$$

(3.6)

$$
\frac{\partial J}{\partial V} = K_y (y - y^0) \int \frac{dt}{h}.
$$

(3.7)

We used the relation between the buoy trajectory and the velocity, i.e., kinematic condition, for deriving the misfit terms. The model equations govern the evolution forward in time, while the adjoint equations are going to evolve backward in time.

With our nonlinear simple model we need to conduct a different kind of parameter estimation than the usual one. In the usual parameter estimation, a model has an explicit parameter (e.g., squared gravity wave speed in a linear shallow water model). We at first derive a gradient of Lagrangian with respect to the parameter, and estimate an optimal value of the parameter with the adjoint, gradient and descent schemes. The nonlinear model does not have an explicit parameter except for the eddy viscosity. Although the
technique is the usual parameter estimation technique, we are going to adopt and estimate an optimized hypothetical mean upper layer thickness $H$ that is implicit in the model. The main dynamic balance is the pressure gradient and the wind stress in the equatorial region. Hence we put $H$ in the pressure gradient terms as a model parameter, only when we derive the gradient of the Lagrangian: We replace the original pressure gradient terms with the following ones

$$\frac{g' H}{a \cos \theta} \frac{\partial h}{\partial \phi} \quad \text{and} \quad \frac{g' H}{a} \frac{\partial h}{\partial \theta}$$

and derive the gradient of the Lagrangian

$$\frac{\partial L}{\partial H} = TK_c(H - H_e)$$

$$+ \int g \left[ \frac{\lambda_u}{a \cos \theta} \frac{\partial h}{\partial \phi} + \frac{\lambda_v}{a} \frac{\partial h}{\partial \theta} \right] dt,$$

where $T$ is an assimilation period. The gradient is a function of both longitude and latitude.

With the gradient for obtaining an optimized state, the thickness $H$ may contain the large scale phenomena as the control parameter, since the terms Eq. (3.8) balance with the wind forcing as the main balance in the model. The replacements of the terms Eq. (3.8) are done only for deriving the gradient. We do not insert the optimized $H$ into the pressure gradient terms for the model integration but use it as an initial condition. When we integrate the model forward in time, we use the original nonlinear model equations. The expression of the gradient is derived from the model equations with the replaced pressure gradient terms, Eq. (3.8) which acts as the main balance. Because the time evolution is governed by the original nonlinear model, the value of the gradient contains an effect of the nonlinear evolution as a small perturbation. Because the gradient term is used for searching the optimized state, the procedure searches a state that the main balance governs as the lowest order with the small perturbation of the nonlinear time evolution.

This optimization method in this study is quite different from the usual parameter estimation. It has an effect of continuous adjustment of the hypothetical parameter to the data over an assimilation period, while the variational adjoint method usually optimizes a control parameter over the assimilation period. In order to distinguish our method from usual parameter estimation and initialization, for the time being, we designate our method continuous data assimilation, though it has a rather wide meaning.

The procedure of the assimilation can be formulated in the following way: (1) Choose a first guess for the model parameter. (2) Integrate the model forward in time for the period of the assimilation. (3) Integrate the adjoint equations backward in time forced by the model-data misfits. (4) Calculate the gradient of the cost function with respect to the model parameter. Use the gradient in a descent algorithm [Shanno and Phua's (1980) limited memory quasi-Newton conjugate gradient algorithm in this study] to find a new value of the model parameter that makes the cost function move towards its minimum. (5) Check if the optimal solution has been found. This can be done by checking the norm of the gradient or the value of the cost function to see if it is less than a prescribed tolerance. (6) If this is not the case, the new value is used as the initial condition and the above procedure can be repeated. We continue the iterative process until a satisfactory solution has been determined.

There are some points which one should notice about the adjoint equations. (1) The adjoint equations are forced by the model-data misfits. The wind stress does not appear in these equations. (2) The adjoint equations correspond to an evolution backward in time. Since the system of equations governing the Lagrange multipliers are similar to the model equations, the information propagates in the form of equatorial Kelvin and equatorial/mid-latitude Rossby waves. The equatorial Kelvin wave propagates westward in the adjoint equations; the Rossby waves propagate eastward. (3) The adjoint equation is always a "linear" equation even though the model equation is a nonlinear equation (see Kamachi, 1994 for further discussions). (4) When the model is a system of simultaneous nonlinear differential equations, the adjoint equations have some extra interaction terms. These extra terms are the in-
teractions between the model variables and the Lagrange multipliers. The additional terms cause the nonlinear propagation of waves in the linear adjoint system (see Kamachi 1994 for example).

We derived the adjoint Eqs. (3.3)-(3.5) in its analytic form in order to examine the characteristics of the equations. Practically we use a finite difference form of the adjoint equations that are derived directly from the finite difference form of the model equations and the Lagrangian. The finite difference form of the adjoint equations is thus consistent with the model equations.

We also derive the Lagrangian, model equations, adjoint equations, and the gradient of the cost function in the non-dimensional forms, because the scaling of those equations is related to the shape of the cost function and its second derivative (Hessian matrix). The scaling is important for a well-conditioned problem and the speed of the convergence of the optimization process. We adopted the scaling of variables (see Gill et al., 1981). Three Gauss' moduli are not independent with each other and two of them are independent. In the numerical procedure all equations are normalized so that only the ratios \( K_x/K_H \) and \( K_y/K_H \) appear. As typical scales we used the equatorial radius of deformation and the time scale, which are the typical length and time scales in the equatorial ocean and related to the Kelvin wave speed and the meridional variation of the Coriolis parameter (Gill, 1982).

Smedstad and O'Brien (1990) stated that the estimated parameter has been assumed to be a function of longitude only in their data assimilation study. If the phase speed would be a function of both longitude and latitude, the limited number of observations would not yield a feasible problem. In addition they restricted the attention to the large scale variation, thus reduced the dimensionality of the parameter. Buoy observations presented here may give enough information to make it possible to determine the spatial structure of the mean upper layer thickness that is a function of both longitude and latitude. Furthermore a nonlinear fine mesh model used can express mesoscale eddy phenomena. Also buoy data show eddy phenomena (Hansen and Maul, 1991). We, therefore, assume that the optimal parameter is a function of both longitude and latitude.

The model-data misfit terms Eqs. (3.6) and (3.7) give values at one point with each buoy at a time. Even though it appears that 40 to 100 buoy "trajectories" used in this study cover the whole ocean, the number of the applied model-data misfit may not be enough in every time step of the adjoint equations. Hence we are going to increase the number of the grid points at which the model-data misfit terms can apply. Because the model's dynamic state has a typical length scale, i.e., the radius of deformation, the state has similar behavior in the region. The value of the misfit may hence have the similar value within the radius of deformation and it may have a distribution in space. We adopt the Gaussian type as the distribution of the misfit. This spreading of the information of the misfit has the same effect as that, in the cost function, the difference of the model and observed quantities has a uniform distribution in the region, and the validity coefficients have the Gaussian distribution. The distribution means that the observation errors have an exponentially increasing function from the original observation point. The practical procedure of the calculation of the misfit is as follows. Every day we get a new position of an observed buoy. We calculate a model buoy trajectory started from the observed position for 24 h. Then the model-data misfit terms are calculated using Eqs. (3.6) and (3.7) at the end position of the model buoy for 24 h. We calculate the spatial autocorrelation of velocity data on \( 32 \times 32 \) grids to get a decorrelation scale, which is defined as the e-folding scale of the autocorrelation. At last we obtain a spatial Gaussian distribution of the misfit with the decorrelation scale. The above procedure is repeated every day in the model calculation.

4. Assimilation experiments and results

The primary objective of this research work is to be able to determine the implicit time-independent upper layer thickness \( H \) in the model, and to get the better description of the ocean
state. We at first examine the performance of the data assimilation algorithm. We examine experiments in which model simulated data are used as the “observations.” The observation run has a seasonal variation. The equatorial Pacific Ocean model was spun up from rest and uniform depth, \( U = V = 0 \) and \( h = 300 \text{ m} \), using the Florida State University wind stress monthly mean climatology, which is obtained from ship observations in the period 1966–1985. The integration for the spinup is done for twenty years. In the twenty-first year the forty “observed” buoy trajectories are calculated with an iterative method in which the kinematic conditions between the model velocity and buoy position are used. The data are collected each day in the observation run (IT0OBS and IT1OBS; the experimental conditions are described in Table 1). In the spinup period, the model equations are forced with different wind fields for the different integrations (i.e., simulation and assimilation). Wind fields used have the same annual mean value as the FSU wind climatology with different seasonal variations (5%, 10%, 20% stronger). We also used 30, 50 and 100% stronger seasonal wind fields. However those winds are unrealistic. The meridional wind at the Galapagos Island is northward, because the Inter-Tropical Convergence Zone (ITCZ) is in the Northern Hemisphere throughout the year in the eastern Pacific Ocean. The 30% stronger wind has the southward meridional component in the winter season at the Galapagos Island. It means that the ITCZ is in the Southern Hemisphere or the intertropical divergence zone appears in the Northern Hemisphere. This is an unrealistic situation even if we adopt wind fields with observational error in the assimilation experiments. Therefore we adopted 5, 10 and 20% stronger wind fields. In the experiments, position data of the model simulated buoy are stored every day, and the model variables are stored every five days. Elements of the variable parameters for the sensitivity in the assimilation experiments are shown in Table 1.

Fig. 3 shows examples of drifting buoy trajectories of (a) observation (IT1OBS) and (b) simulation (IT1sim20) experiments for one-year period. We used 40 buoys in the experiments for the seasonal variation. Those trajectories generally show the main equatorial currents: NEC, NECC and SEC. The trajectories in the simulation are different in detail from those in observation. In

<table>
<thead>
<tr>
<th>EXP.</th>
<th>First guess</th>
<th>Data region</th>
<th>Period</th>
<th>Wind forcing</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT0OBS</td>
<td>no assim</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0sim20</td>
<td>no assim</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0all05</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0all052</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0all053</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0all05FG</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0all10</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0all20</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0NECW20</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0NECE20</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0NECCW20</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0NECE20</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0SECW20</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
<tr>
<td>IT0SEC20</td>
<td>all</td>
<td>all</td>
<td>1 yr</td>
<td></td>
</tr>
</tbody>
</table>
the western part those trajectories show wavy lines, which is identical to the mixed Rossby-gravity waves, inertia gravity waves and eddy phenomena. Especially large amplitude Rossby-gravity waves may cause problems since they can give rise to unrealistic meanders in the equatorial currents (see Hansen and Maul, 1991). The Rossby-gravity waves are generated in September at about 150° E on the equator, and last until April in the next year. The wave is a monochromatic wave which wavelength is 850 km and period is 28 days. The wavelength and the period satisfy the dispersion relation of the mixed Rossby-gravity wave. The westward phase speed is 0.352 ms⁻¹, and the eastward group velocity is 0.263 ms⁻¹. Using differences of the trajectories above mentioned we proceed to the assimilation experiments.

The upper layer thicknesses are contoured in Fig. 4. Fig. 4a shows the “true” upper layer thickness that is a result of the experiment ITOOBS, in which the observed buoy trajectories are calculated, and Fig. 4b shows the simulated upper thickness in the experiment ITOsim20. The true upper layer thickness would not be known to observers. We have only the observed buoy trajectory data and the simulated upper layer thickness. Using these two types of information, we are going to recover the true upper layer thickness.

In the first series of the assimilation experiments we adopted three months as the assimilation experiments.
tion period since the period may contain and represent a seasonal variation of the equatorial currents and upper layer thickness. Fig. 5 shows the normalized values of the cost function and its gradient versus successive iterations for experiment IT0all20. The cost function and its gradient are normalized by the values of the 0-th iteration, i.e., simulation, respectively. The final value of the gradient is close to zero, though the value of the cost function is still larger than zero. This tendency suggests that the shape of the cost function may be nearly flat. The tendency also suggests that the initial guess of the control parameter is close to a minimum state. Is this minimum state the global minimum state? Although the sensitivity of the first guess to the optimization is beyond this study, we carried out several simple experiments in which the first guess is different from the experiment IT0all20. The first guesses in these experiments are related to different wind fields, in which seasonal increments are 5, 10 and 20% (IT0all05, IT0all10, IT0all20), and to different spinup time: no spinup (IT0all05FG3), 5, 15 and 20 years (IT0all052, IT0all053, IT0all05). We get the same optimized ocean state with those wind fields and 15/20 years' spinup. It suggests that the optimized state may be on the global minimum, though it is not a direct proof of the optimization for the global minimum. We could not, however, get an optimized state from the first guess of 5 year spinup and no-spinup state. The value of cost function did not decrease and...

![TRUE ULT IN IDENTICAL TWIN EXP.](image)

![UPPER LAYER THICKNESS IN IDENTICAL TWIN EXP.](image)

Fig. 4. Comparison of (a) the "true" upper layer thickness and (b) the simulated upper layer thickness.
the norm of the gradient oscillates with the different order of the magnitude with iteration number. There may be multiple minima around those first guess points, which may be far away from the global minimum point. We examined a time-evolution of the total kinetic energy for checking the statistically equilibrium state. The figure shows that the total kinetic energy after 5 year integration is about 60% of the statistical equilibrium and the currents are rather weak not only in the mid-latitude region but also in the equatorial region. Then the wind forces the ocean state toward a different statistical equilibrium even if we use a data assimilation. When the ocean state is not fully spun up, the assimilation process may be deteriorated with wind forcing and strong non-linearity that causes the multiple local minima. The used upper layer thickness after spin-up of 20 years may be a good first guess. It is due that the prior information contains useful information. A complete fit to the observations could be done with an experiment of the complete observation and complete model dynamics without error. Every observation contains four types of information: (1) controllable phenomena that have the same dynamics as the model, (2) phenomena with other dynamics, (3) uncontrollable phenomena even if the dynamics is the same as the model one (e.g., model and observation have inertial gravity waves, but model may not be able to control the wave with the different external conditions such as wind forcing and topographic geometry) and (4) observation or model errors.

In Fig. 6, time series of the upper layer thickness are plotted. Fig. 6a shows the time series of the upper layer thicknesses of simulation, assimilation and true state at Canton Island in the western Pacific (Niño4 area), and Fig. 6b shows the time series at Galapagos Island in the eastern Pacific for experiment IT0all20. The root-mean-square value of residuals of simulation from true upper layer thicknesses is 4.2 m at Canton Island. The value of the assimilation from true ULT is 1.8 m. On the other hand the value of the simulation from true ULT is 3.7 m and the assimilation counterpart is 2 m at Galapagos Island. These values show that the assimilation procedure improved. The figures generally show that the time variation of the assimilated ULT is similar to that of simulation, and the mean level is near the true value. In this experiment with three months of the assimilation period, the upper layer thickness in the western region is relatively going to be recovered in four iterations, but not so much in the eastern region.

Fig. 7 depicts a residual between ULT of the 0th iteration (simulation) and that of the last iteration of the assimilation in the latitude-longitude plane. The difference between simulation and the true field is larger (e.g., 10 m) in the eastern Pacific than in the western Pacific (e.g., 5 m), though a figure is not shown. Fig. 7 shows information on how the upper layer thickness is improved. Western equatorial, NECC and SEC regions are generally improved well but not in the NEC region between 15° to 20° N (see Figs. 1 and 3 for the equatorial current system). The figure also shows that the eastern region is affected zonally because the buoy trajectories are almost zonal. The central to western regions have rather meridional pattern, because the mixed Rossby-gravity wave affects the buoy trajectories. The off-equatorial upper layer thickness is not improved. The improvement is confined around the buoy trajectory in the equatorial region.
Why were not the upper layer thicknesses in the NEC and eastern equatorial SEC regions improved? We examine the relations between the wave propagation, which is generated from the model-data misfit terms in the adjoint equations, and buoy trajectories, which signify advected source points of the wave propagation.

At first, in Fig. 8, we show an example of longitude-time diagram of the Lagrange multiplier at 3°S in the adjoint process. It should be noticed that the evolution of the adjoint process is backward in time. The western (150°E, March and early April) and eastern (110°W, late March) regions are the main generating area in this experiment in a similar longitude-time figure (not shown) of the model-data misfit terms. Fig. 8 shows propagation of equatorial Kelvin and Rossby waves from these generation areas. In the western region, 155°E to 180°E, at the earlier time, February to January, the amplitude of the Rossby wave is increased because of the reflection of the equatorial Kelvin wave. These indicate that the western region has a dense wave characteristics.

Next we show schematic diagrams that show relations of the ocean current, the buoy trajectory and wave propagation in the adjoint system in Fig. 9. Fig. 9a shows relations of the directions of major currents and wave propagation in a longitude-latitude domain. In the NEC region mid-latitude Rossby wave propagates westward, and the direction is the same as that of the buoy trajectory. In the NECC region equatorial Rossby wave propagates toward the opposite direction of the drifting buoy, i.e., westward. Situation is similar to that in the SEC region, though the wave is the equatorial Kelvin wave. Fig. 9b, c and d show relations of buoy trajectories and wave propagation in the adjoint system in longitude-time domain for NEC, NECC and SEC regions, respectively. In these figures the broken lines show buoy trajectories. The solid lines show the propagation of waves that are generated at the end time of the assimilation period as an example in the adjoint system. Typical values of the period in which the waves propagate in the basin scale are 10 years, 9 months, and 3 months for the mid-latitude Rossby, the equatorial Rossby, and the

Fig. 6. Comparison of the upper layer thickness of simulation, assimilation and true state at (a) Canton Island (3°S, 172°W) and (b) Galapagos Island (0.75°S, 90.3°W) for experiment IT0ail20.
equatorial Kelvin waves, respectively. Fig. 9b shows that buoys (i.e., the wave generation positions) have drifted westward in the real time and the mid-latitude Rossby waves propagate eastward in the inverse time in the NEC region. Those two trajectories are almost parallel and overlapped with each other. It means that the adjoint equations and the gradient of the cost function use the almost the same model-data misfit information in the whole assimilation period. The assimilation procedure, therefore, is not effective in the NEC region. Fig. 9c depicts that buoys are advected eastward in the real time and the equatorial Rossby waves propagate eastward in the inverse time in the NECC region. Generated waves are scattered over the longi-

![Fig. 7. Difference of the upper layer thickness of the simulation and assimilation experiments.](image)

**Fig. 8.** Longitude–time diagram of the Lagrange multipliers at 3°S in the adjoint process. Horizontal axis is the longitude and the vertical one is the time (month).

![Fig. 9. Schematic diagram of the buoy trajectory and wave propagation in the adjoint system. (a) Relationship of major currents and wave propagation in longitude–latitude domain. NEC: North Equatorial Current. NECC: North Equatorial Counter-Current. SEC: South Equatorial Current. MR: Mid-latitude Rossby wave. ER: Equatorial Rossby wave. EK: Equatorial Kelvin wave. (b), (c) and (d) show relations of buoy trajectories and wave propagation in the adjoint system in longitude–time domain. (b) NEC region. (c) NECC region. (d) SEC region.](image)

titude-time domain. The adjoint equations use the model-data misfit information in the basin scale. The SEC region is different from the NEC and NECC regions (see Fig. 9d). Buoys are drifted eastward in the real time and the equatorial Kelvin wave propagates westward in the inverse time. Some Kelvin waves reflect at the western boundary, then the equatorial Rossby waves propagate more slowly than the Kelvin waves do. Therefore the density of the characteristics of the Rossby waves is rather high in the western region. The western region acts as a special caustic region for the adjoint system with the assimilation period of three months. The assimilation procedure can improve the upper layer thickness relatively better in the western region than in the eastern region. This effect came out in the difference of the Figs. 6a, b and 7. This east-west inhomogeneous effect would be dissolved in the assimilation experiments with a longer period.

We also performed some experiments with a longer assimilation period. One year is adopted as the assimilation period. A result of evolution of optimization process is shown in Fig. 10. In the figure, the values of the normalized cost function and its gradient are plotted versus the iteration number for experiment IT1all20. Both values of the cost function and its gradient decrease gradually with oscillation as the iteration number increases. Both of the oscillatory decreasing curves suggest that the shape of the cost function is bumpy, which is an example of the nonlinear optimization phenomena.

We show time series of the upper layer thickness. Fig. 11a shows the time series of the upper layer thicknesses of simulation, assimilation and the truth at Canton Island and Fig. 11b shows the time series at Galapagos Island for experiment IT1all20. Generally the upper layer thicknesses are improved in the eastern as well as western Pacific. We calculated the root-mean-square values of the simulation/assimilation from true upper layer thickness. The value of simulation from the true ULT is 6.8 m and the assimilation counterpart is 1.6 m at Canton Island. The value of the simulation from true ULT is 3.7 m and the assimilation counterpart is 1.1 m at Galapagos Island. The assimilated value is nearer to the true field than the simulated one. The figures also depict that the time variation of the assimilated ULT is not similar to that of the simulation, which is different from the result with the period of three months. The correspondences of ULT at the islands and the slow decrease of the cost function in Fig. 10 indicate that the ULT in the equatorial region is improved but not in the off-equatorial region, in which the information of the model-data misfit cannot be advected toward the off-equatorial region. The assimilated hindcast shows the good agreement with the true field over the whole period except for the early time, e.g., January to March. This is the different feature of the continuous adjoint assimilation in this study from the initialization with adjoint method, because the variational initialization method can improve the ocean state in the earlier period but not in the later stage (see, e.g., Sheinbaum and Anderson, 1990a and b).

Next we examine experiments with real data. We adopt real observation data of buoy trajectories in 1983. Quality controls used are checking and elimination of undrogued buoys and removing energetic inertial oscillation using some low-pass filters (e.g., Butterworth type, or Savitzky-Golay filter). Details of the quality control are
explained in Hansen and Maul (1991). The experimental condition of the assimilation is shown in Table 1. We adopted one year as the assimilation period. For twenty years, the model was spun up from rest and uniform depth, $U = V = 0$ and $h = 300$ m, using the Florida State University wind stress monthly mean climatology. After the spinup the model was integrated with the FSU monthly wind stress from 1961 to 1990. The integration was started from 16 January 1961.

Here we are going to compare the assimilated ULT with independent observations of sea level in 1983 as an example. At first, in Fig. 12, we show observed buoy trajectories in 1983. These trajectories are used as observation in the assimilation experiment Rd83. Some trajectories show Costa Rica eddies and Legeckis waves, and many trajectories show inertial gravity wave motion. The data cover the eastern Pacific region only. Even if the data do not cover the western Pacific, we found, in a figure (not shown) similar to Fig. 7 for this experiment, that ULT in the western region is improved with wave propagation in the adjoint system as discussed with Fig. 9.

Fig. 13 shows the evolution of optimization process. The normalized cost function is plotted versus the iteration number for experiment Rd83 in the figure. The line of the normalized gradient is omitted. The cost function gradually decreases as the iteration number increases. The normalized cost function has almost the same values at the 3-4 and 5-6 iterations that means no improvement. The curve also suggests that the shape of the cost function may be flat.

We compared the optimal upper layer thickness with the independently observed sea level record. Fig. 14 depicts time series of the ULT of the simulation, of the assimilation experiments and the observed sea level at Santa Cruz in Galapagos Island. The amplitudes of the simulated and assimilated ULT’s are smaller than the observed values, if the ULT’s are converted to the equivalent depth, which conversion factor is 0.002. Though the assimilation process tries to adjust to the observed time variation, the improvement depends on month. In January and April ULT is rather improved than other months. The variation from December 1983 and January 1984 is fairly high. The value of January 1984 is used of the 1984 experiment which is independent of the 1983 experiment, then the value is affected from the period in 1984.

![ULT Clim (Canton)](ULT_CLIM_CANTON.png)

![ULT Clim (Galapagos)](ULT_CLIM_GALAPAGOS.png)

Fig. 11. Comparison of the upper layer thickness of simulation, assimilation and true state at (a) Canton Island and (b) Galapagos Island for experiment ITall20.
5. Discussion and summary

Drifting buoy data have useful information about low frequency ocean variability, though extracting the useful information is difficult. It is well known that the drifting buoy trajectory has a chaotic nature in mid-latitude ocean where mean currents are unstable and midocean eddies form. On the other hand it is still unknown whether the trajectories are chaotic or regular in the equatorial region, because the equatorial region is mainly laminar flow, moreover unknown whether the chaotic nature is destructive to the data assimilation. We estimated the typical time scale of the chaotic trajectory, if it would be chaotic. The time scale is rather larger than our assimilation period.

Fig. 12. An example of real buoy trajectories in 1983. Released points are designated by star sign. The trajectories show Costa Rica eddies, Legeckis waves and inertial gravity wave motion.

Fig. 13. Evolution of optimization process of the normalized cost function for experiment Rd83.

Fig. 14. Comparison of the optimal upper layer thickness and the independently observed sea level record at Santa Cruz in Galapagos Island. Thick line shows the observed sea level time series and the thin lines show assimilated upper layer thickness.
(3 months, 1 year) and the period of the calculation of the data misfit (1 day). It may mean that the assimilation procedure works well (see Appendix).

In many experiments we showed that the assimilation algorithm is able to determine the spatial structure of the hypothetical mean upper layer thickness, i.e., the model is capable of fitting the observed data in the equatorial Pacific Ocean. We discussed regions that are not influenced by the buoy data, moreover examined relations of wave propagation and the generation area determined by the model-data misfit along the buoy trajectory. We examined the different assimilation effects in the NEC, NECC and SEC regions. The adjoint system cannot get much information about the model-data misfit in the NEC region. Then the UL T is not improved well in the region. In the NECC region the UL T is improved well. The SEC region has a different effect on the wave propagation in the adjoint system from NEC and NECC regions. Because a western boundary acts as a caustic region, the SEC region has an east–west inhomogeneous effect on the adjoint system with short assimilation period (e.g., three months). If we adopt a longer assimilation period (e.g., one year), the east–west inhomogeneity is dissolved and we get a similar improvement in both of western and eastern regions.

Continuous data assimilation in this study has rather a different effect on the adjoint system than initialization. For the continuous data assimilation in this study, the time series of the optimized upper layer thickness has a similar variation to the true field in the middle of the period, but not initially. For initialization, the assimilated upper layer thickness fits to the true field in the beginning of the assimilation period but not in the middle to the end period (see, e.g., Sheinbaum and Anderson, 1990a and b). In the initialization experiments the data near the initial time has a larger effect on the optimal initial condition in the adjoint system. This tendency is due that the gradient of the cost function is effective for the initial period, and the information about the model-data misfit from the end period dissipates, though the dissipation effect depends on the value of the eddy viscosity. The dissipation time is, for example, about 600 days with a length scale 200 km and the horizontal eddy viscosity 750 m² s⁻¹ in this paper. In the continuous data assimilation the optimization is effective for the time mean or low frequency variation over the period. This is the different point in the two types of the methods.

The optimal assimilation period may be about one year because the equatorial Kelvin and Rossby waves can propagate in the basin scale in a year and then the information of the model-data misfit can propagate also. Since the model-data misfit terms have a confined form, the terms contain a continuous spectrum for the generation of the waves. The waves generated from the misfit terms have a broad band of the spectrum and those waves disperse in the period more than one year. Since the dispersion plays as a similar role to the dissipation, the method may not be so effective in the period more than one year.

The main balance in the equatorial region is between the zonal wind stress and the pressure gradient. A linearized model cannot fit to the data in the eastern Pacific Ocean well in initialization experiments, because there is an inconsistency between the model parameter, the wind forcing and the pressure gradient (Sheinbaum and Anderson 1990a, b). The wind stress and the parameter are given and they determine the zonal pressure gradient in the model main balance. On the other hand the observed data also determine the zonal pressure gradient in assimilation experiments with the linearized model. The pressure gradient in the model is therefore inconsistent with the data. The initialization is difficult with a linear model. The problem of the inconsistency can be solved automatically in the parameter estimation with a linear model, in initialization with a nonlinear model, and in the continuous data assimilation with a nonlinear model in this study.

It should be noted that the ability of a numerical scheme to generate a realistic nowcast depends on the numerical and physical accuracy of the model and performance of the assimilation method. For the assimilation presented in this study, the model dynamics match the dynamics of
the "true" solution in the experiments with model simulated data but not so much in the real assimilation experiments. Although reduced gravity models are able to simulate a variety of oceanic phenomena, it is necessary to develop assimilation schemes for multi-vertical mode models for more realistic ocean response (see Busalacchi and Cane, 1985 for example). Future assimilation experiments will utilize higher horizontal resolution and multi-vertical mode models with a sophisticated assimilation and optimization schemes. The simple experiments presented in this study are, however, encouraging for the possible use of satellite tracking Lagrangian trajectories of drifting buoys in ocean monitoring and prediction.

Acknowledgments

We wish to thank D. Hansen (AOML/NOAA) for graciously providing the drifting buoy data. Special thanks to A. Davis for programming assistance of data analyses, to M. Johnson (University of Alaska) for developing the model code and to S. Meyers for discussion of the chaotic problem of the drifting buoy trajectories. The computer work was done on Cray ymp/432 at the Florida State University. This work was supported by a Grant from the Earth Science Division of NASA. M. Kamachi did most of this work while visiting the Mesoscale Air-Sea Interaction Group (MASIG) at the Florida State University. The Office of Naval Research provides the basic support for MASIG through the Secretary of Navy Chair to J. O'Brien. NOAA, Office of Global Program provides the support to prepare the wind fields.

Appendix

It has been well known that drifting buoy trajectories show chaotic features in mid-latitude, e.g., Gulf stream region, because the currents are unstable and midocean eddies appear. It has not been shown whether buoy trajectories are chaotic or regular in the equatorial region. The only study concerning the chaotic nature of equatorial buoy trajectories is Meyers et al. (1994). They used model simulated buoy trajectories in the Indian Ocean, and estimated scaling dimension and did a spectral analysis for fractal trajectories.

We examine the chaotic problem in the deterministic equatorial system. We estimate the Lyapnov exponent as a measure of chaotic phenomena. Let \( d(0) \) denote the distance between two adjacent particles at initial time and \( d(t) \) denote the evolution at later times. The initial adjacent particles become separated linearly under a regular motion and exponentially under a chaotic motion. The Lyapnov exponent measures linear or exponential separation as defined mathematically

\[
\sigma = \lim_{t \to \infty} \lim_{\Delta d(0) \to 0} \frac{1}{t} \ln \frac{d(t)}{d(0)} \tag{A.1}
\]

For a regular trajectory \( d(t) \) grows linearly and \( \sigma = 0 \). For a chaotic trajectory \( d(t) \) grows exponentially \( \{d(0) \exp(\sigma t)\} \) and \( \sigma \) has a positive value. The Lyapnov exponent is interpreted as estimating the rate of loss of information with a relation of time \( T_m \) over which the state of chaotic system can be predicted: \( T_m = (1/\sigma) \ln[1/d(0)] \). Precise predictions are possible for a period that is smaller than \( T_m \) (Schuster, 1989).

We adopt the simplest approach in which we calculated the evolution of distance between initially close trajectories. We estimated the Lyapnov exponent in three cases, two are with model simulated buoys, one with real buoys. Forty pairs are selected with an initial distance for each pair is one model grid spacing \((1/4 \text{ degree})\) in latitude. In this case, the non-dimensional \( d(0) \) scaled with the equatorial radius of deformation is 0.0851. After one year integration we calculate the Lyapnov exponent. The average value is 0.00371 \((\text{day}^{-1})\) with a time scale \( T_m \) of 664 days. A second experiment uses the initial distance of one model grid spacing in longitude. The Lyapnov exponent is 0.00141 \((\text{day}^{-1})\) with a time scale \( T_m \) of 1749 days. Next we used real buoy trajectories for 1983. We selected 6 pairs with initial separation distances much smaller than the equatorial radius of deformation. After one year we obtain the Lyapnov exponent of 0.00695 \((\text{day}^{-1})\)
with a time scale $T_m$ of 250 days. These data suggest that for a period less than one year we can estimate the trajectory to be approximately regular. In our experiments, we used assimilation periods of 3 months and 1 year. Thus we conclude that we may treat the buoy trajectories as regular. Furthermore, when we calculate the model-data misfit terms everyday, we start the model simulated buoy exactly at the same point of the observed buoy position. We renewed the initial positions of the model simulated buoys every day. This is much smaller than the typical time scale derived from the Lyapnov exponent. Therefore the chaotic effects should not affect the calculation of the model-data misfit.

References


