Tellus (1998), 50A, 534–544
Printed in UK – all rights reserved

TELLUS
ISSN 0280–6495

On determining initial conditions and parameters in a simple coupled atmosphere-ocean model by adjoint data assimilation

By JINGXI LU and WILLIAM W. HSIEH*, *Oceanography/EOS, University of British Columbia, Vancouver, BC, Canada V6T1Z4*

(Manuscript received 20 August 1997; in final form 13 May 1998)

ABSTRACT

The general problem of retrieving both initial conditions and model parameters for coupled atmosphere-ocean models by the adjoint data assimilation method was formulated. In particular, for a simple coupled equatorial model, where the atmosphere and the ocean were each represented by a linear shallow water model, retrieval of 3 oceanic initial conditions (the sea level height (SLH) and the 2 horizontal current components) together with 6 (damping and coupling) parameters was performed. Identical twin experiments assimilating wind and SLH data were conducted to test (i) the mutual influence of initialization and parameter estimation, (ii) the information transfer efficiency between the atmosphere and the ocean during retrieval, and (iii) the effect of initial guess on retrieving the initial conditions and parameters. By assimilating the wind and SLH data once per day at every grid point, retrieving both the parameters and initial conditions together was found to be more difficult than retrieving either of them separately. Assimilating the same data once per day in the TAO array yielded a much poorer but still acceptable retrieval, as the amount of information might be insufficient. The wind and SLH data were comparable in importance when retrieving the initial conditions, but not so when retrieving parameters, the wind data being more helpful for retrieving the atmospheric damping coefficients, and the SLH data for the oceanic damping coefficients. Errors in the initially guessed values of the parameters or the initial conditions generally affected the retrieval quality. The initial condition retrieval was more sensitive to errors in the guessed initial conditions, especially to phase errors, while the parameter estimation was more sensitive to errors in the guessed parameters.

1. Introduction

Coupled atmosphere-ocean models are important for understanding the climate system as well as for forecasting climate variability such as the El Niño phenomenon (Zebiak and Cane, 1987; Ji et al., 1994). Our two earlier works, Lu and Hsieh (1997), henceforth LH1, and Lu and Hsieh (1998), henceforth LH2, have investigated the potential of the adjoint method, a variational data assimilation method, in determining the model parameters (LH1) and in estimating the initial conditions (LH2) of coupled models. Verifying model parameters and determining initial conditions by data assimilation are very relevant to improving the prediction skills of simple (intermediate) coupled models (Palmer and Anderson, 1994; Barnston et al., 1994; Allen and Davey, 1993; Chen et al., 1995). Using the simple equatorial coupled model of Philander et al. (1984) (henceforth the PYP model), LH1 retrieved 6 coupling and damping parameters, while LH2 retrieved the oceanic initial

* Corresponding author.
e-mail: william@eos.ubc.ca

conditions (for the sea level height SLH and the horizontal currents).

In the context of variational adjoint data assimilation, parameters and initial conditions can be estimated together (Le Dimet and Talagrand, 1986; Talagrand and Courtier, 1987; Thacker and Long, 1988; Tziperman et al., 1989). The model solution that best fits (in the generalized least-squares sense) the observations within some space-time domain, can be obtained by adjusting the model parameters and the initial conditions simultaneously. In real application, it is not realistic to isolate one when studying the other. As a further step towards testing the potential of determining the initial conditions and parameters of coupled models by the adjoint method, we conducted experiments in this paper to estimate both the 6 parameters and the 3 oceanic initial conditions in the PYP model, thus combining the separate approaches of LH1 and LH2. This allows an integrated conclusion to be made, as Yu and O'Brien (1991) have shown that by adjusting both the initial conditions and parameters of their model, the parameters were better estimated and the optimal current fields in deeper ocean were also better retrieved than those obtained by adjusting only the parameters.

Here we investigated: (1) the influence of parameter estimation on initialization and the influence of initialization on parameter retrieval, (2) the efficiency of information transfer between the atmosphere and the ocean and the information sufficiency of the TAO (Tropical Atmosphere and Ocean) array in retrieving both the parameters and the initial conditions, and (3) the effect of initially guessed values (for the parameters or the initial conditions) on the retrieval of both parameters and initial conditions.

In Section 2, a general procedure for estimating model parameters and initial conditions of coupled systems is described. In Section 3, the PYP model and its adjoint model are briefly described. The numerical experiments are described in Section 4, with their results presented in Section 5. Section 6 gives a summary and discussion.

2. A generic procedure for initialization and parameter estimation in coupled models

Let $\boldsymbol{x}$ be the vector of state variables of the atmospheric model, with $\boldsymbol{X}$ the atmospheric model operator. Let $\boldsymbol{y}$ and $\boldsymbol{Y}$ denote the state vector and the operator of the ocean model, respectively. A coupled model can be described by:

$$\begin{aligned} &\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{X}(\boldsymbol{x}, \boldsymbol{p}_x, \boldsymbol{y}, t), \\ &\boldsymbol{x}(0) = \boldsymbol{x}_0, \\ &\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{Y}(\boldsymbol{y}, \boldsymbol{p}_y, \boldsymbol{x}, t), \\ &\boldsymbol{y}(0) = \boldsymbol{y}_0, \end{aligned} \tag{2.1}$$

where the unknown model parameters (e.g., friction parameters, coupling parameters, etc.) and unknown initial conditions for the atmospheric and oceanic models are given by the vectors $\boldsymbol{p}_x$ and $\boldsymbol{x}_0$, and $\boldsymbol{p}_y$ and $\boldsymbol{y}_0$, respectively.

The 4-D variational problem involves minimizing a quadratic cost function $J = J(\boldsymbol{x}(\boldsymbol{p}_x, t), \boldsymbol{y}(\boldsymbol{p}_y, t), \boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{x}_0, \boldsymbol{y}_0)$ subject to model constraint (2.1), which is defined as:

$$J(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{x}_0, \boldsymbol{y}_0) = \int_0^T D(\boldsymbol{x}, \boldsymbol{y}, t)\,\mathrm{d}t, \tag{2.2}$$

where T is the assimilation time window, and D measures the distance between the model state and the observations at time t, i.e.,

$$\begin{aligned} D(\boldsymbol{x}, \boldsymbol{y}, t) = {} & [\boldsymbol{x} - \boldsymbol{x}_{\mathrm{obs}}]^{\mathrm{tr}} \boldsymbol{W}_x(t) [\boldsymbol{x} - \boldsymbol{x}_{\mathrm{obs}}] \\ & + [\boldsymbol{y} - \boldsymbol{y}_{\mathrm{obs}}]^{\mathrm{tr}} \boldsymbol{W}_y(t) [\boldsymbol{y} - \boldsymbol{y}_{\mathrm{obs}}], \end{aligned} \tag{2.3}$$

where the subscript obs denotes the observations, the superscript tr, the transpose, and $\boldsymbol{W}_x(t)$ and $\boldsymbol{W}_y(t)$, the weighting matrices.

The gradients of J with respect to, $\boldsymbol{p}_x$, $\boldsymbol{p}_y$, $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ needed for finding the optimal parameters and initial conditions can be found by introducing two Lagrange multiplier vectors $\boldsymbol{x}^*(t)$ and $\boldsymbol{y}^*(t)$ to form a Lagrange function L:

$$\begin{aligned} & L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{x}_0, \boldsymbol{y}_0) \\ & \quad = J + \int_0^T \left\{ \boldsymbol{x}^*(t)^{\mathrm{tr}} \left[\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} - \boldsymbol{X} \right] \right. \\ & \qquad \left. + \boldsymbol{y}^*(t)^{\mathrm{tr}} \left[\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} - \boldsymbol{Y} \right] \right\} dt. \end{aligned} \tag{2.4}$$

The differentiation of L with respect to $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ will simply recover the coupled model. After integration by parts, L becomes

$$L=[\boldsymbol{x}^*(t)^{\mathrm{tr}}\boldsymbol{x}+\boldsymbol{y}^*(t)^{\mathrm{tr}}\boldsymbol{y}]_{t=0}^{t=T} + \int_0^T \left[D - \left[\frac{\mathrm{d}\boldsymbol{x}^*(t)^{\mathrm{tr}}}{\mathrm{d}t}\boldsymbol{x} + \boldsymbol{x}^*(t)^{\mathrm{tr}}\boldsymbol{X} \right] - \left[\frac{\mathrm{d}\boldsymbol{y}^*(t)^{\mathrm{tr}}}{\mathrm{d}t}\boldsymbol{y} + \boldsymbol{y}^*(t)^{\mathrm{tr}}\boldsymbol{Y} \right] \right\} \mathrm{d}t\,. \tag{2.5}$$

The first order variation of L is:

$$\begin{aligned}\delta L = {} & [\boldsymbol{x}^*(t)^{\mathrm{tr}}\delta\boldsymbol{x}+\boldsymbol{y}^*(t)^{\mathrm{tr}}\delta\boldsymbol{y}]_{t=0}^{t=T} \\ & + \int_0^T \left\{ \nabla_x D^{\mathrm{tr}}\delta\boldsymbol{x} + \nabla_y D^{\mathrm{tr}}\delta\boldsymbol{y} - \left[\left(\frac{\mathrm{d}\boldsymbol{x}^*}{\mathrm{d}t}\right)^{\mathrm{tr}}\delta\boldsymbol{x} \right.\right. \\ & + \left(\frac{\partial\boldsymbol{X}}{\partial\boldsymbol{x}}\right)^{\mathrm{tr}}\boldsymbol{x}^*(t)\delta\boldsymbol{x} + \left(\frac{\partial\boldsymbol{X}}{\partial\boldsymbol{y}}\right)^{\mathrm{tr}}\boldsymbol{x}^*(t)\delta\boldsymbol{y} \\ & \left. + \left(\frac{\partial\boldsymbol{X}}{\partial\boldsymbol{p}_x}\right)^{\mathrm{tr}}\boldsymbol{x}^*(t)\delta\boldsymbol{p}_x \right] - \left[\left(\frac{\mathrm{d}\boldsymbol{y}^*}{\mathrm{d}t}\right)^{\mathrm{tr}}\delta\boldsymbol{y} \right. \\ & + \left(\frac{\partial\boldsymbol{Y}}{\partial\boldsymbol{y}}\right)^{\mathrm{tr}}\boldsymbol{y}^*(t)\delta\boldsymbol{y} + \left(\frac{\partial\boldsymbol{Y}}{\partial\boldsymbol{x}}\right)^{\mathrm{tr}}\boldsymbol{y}^*(t)\delta\boldsymbol{x} \\ & \left.\left. + \left(\frac{\partial\boldsymbol{Y}}{\partial\boldsymbol{p}_y}\right)^{\mathrm{tr}}\boldsymbol{y}^*(t)\delta\boldsymbol{p}_y \right]\right\} \mathrm{d}t\,. \end{aligned} \tag{2.6}$$

The coupled adjoint model can then be obtained by letting $\delta L/\delta\boldsymbol{x} = \delta L/\delta\boldsymbol{y} = 0$:

$$\begin{aligned} & -\frac{\mathrm{d}\boldsymbol{x}^*}{\mathrm{d}t} = \left(\frac{\partial\boldsymbol{X}}{\partial\boldsymbol{x}}\right)^{\mathrm{tr}}\boldsymbol{x}^*(t) + \left(\frac{\partial\boldsymbol{Y}}{\partial\boldsymbol{x}}\right)^{\mathrm{tr}}\boldsymbol{y}^*(t) - \nabla_x D\,, \\ & \boldsymbol{x}^*(T) = 0\,, \\ & -\frac{\mathrm{d}\boldsymbol{y}^*}{\mathrm{d}t} = \left(\frac{\partial\boldsymbol{Y}}{\partial\boldsymbol{y}}\right)^{\mathrm{tr}}\boldsymbol{y}^*(t) + \left(\frac{\partial\boldsymbol{X}}{\partial\boldsymbol{y}}\right)^{\mathrm{tr}}\boldsymbol{x}^*(t) - \nabla_y D\,, \\ & \boldsymbol{y}^*(T) = 0\,. \end{aligned} \tag{2.7}$$

According to (2.7), the formulae for computing the gradients of J with respect to $\boldsymbol{p}_x$, $\boldsymbol{p}_y$, $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ can be obtained by differentiating (2.6) with respect to these unknowns:

$$\begin{aligned} & \frac{\partial J}{\delta\boldsymbol{p}_x} = -\int_0^T \left(\frac{\partial\boldsymbol{X}}{\partial\boldsymbol{p}_x}\right)^{\mathrm{tr}}\boldsymbol{x}^*(t)\,\mathrm{d}t\,, \\ & \frac{\partial J}{\delta\boldsymbol{p}_y} = -\int_0^T \left(\frac{\partial\boldsymbol{Y}}{\partial\boldsymbol{p}_y}\right)^{\mathrm{tr}}\boldsymbol{y}^*(t)\,\mathrm{d}t\,, \\ & \frac{\partial J}{\delta\boldsymbol{x}_0} = -\boldsymbol{x}^*(0)\,, \\ & \frac{\partial J}{\delta\boldsymbol{y}_0} = -\boldsymbol{y}^*(0)\,. \end{aligned} \tag{2.8}$$

In summary, the general computational procedure for estimating coupled model unknown parameters and initial conditions using the adjoint method is:

(i) specify an initial guess of the unknown $\boldsymbol{p}_x$, $\boldsymbol{p}_y$, $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$;
(ii) integrate the coupled model (2.1) from $t=0$ to $t=T$;
(iii) compute J within $[0, T]$ using forward model trajectory and data;
(iv) compute the gradient from (2.8) by a backward integration of the adjoint model (2.7) from $t=T$ to $t=0$;
(v) use a descent algorithm to find a new estimate of $\boldsymbol{p}_x$, $\boldsymbol{p}_y$, $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$;
(vi) if the convergence criteria are satisfied, terminate the procedure, otherwise redo the procedure from (ii) using the new guess of $\boldsymbol{p}_x$, $\boldsymbol{p}_y$, $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$.

3. A simple equatorial coupled model

The atmosphere (denoted by subscript a) and the ocean (subscript o) of the PYP model are each governed by a set of shallow water equations in spherical coordinates:

$$\begin{aligned} & \frac{\partial u_\mathrm{a}}{\partial t} - f v_\mathrm{a} + \frac{g_\mathrm{a}}{r\cos\phi}\frac{\partial h_\mathrm{a}}{\partial\lambda} = -A u_\mathrm{a}\,, \\ & \frac{\partial v_\mathrm{a}}{\partial t} + f u_\mathrm{a} + \frac{g_\mathrm{a}}{r}\frac{\partial h_\mathrm{a}}{\partial\phi} = -A v_\mathrm{a}\,, \\ & \frac{\partial h_\mathrm{a}}{\partial t} + \frac{d_\mathrm{a}}{r\cos\phi}\left(\frac{\partial u_\mathrm{a}}{\partial\lambda} + \frac{\partial(v_\mathrm{a}\cos\phi)}{\partial\phi}\right) = -B h_\mathrm{a} - \alpha h_0\,, \end{aligned} \tag{3.1}$$

$$\begin{aligned} & \frac{\partial u_\mathrm{o}}{\partial t} - f v_\mathrm{o} + \frac{g_\mathrm{o}}{r\cos\phi}\frac{\partial h_\mathrm{o}}{\partial\lambda} = -a u_\mathrm{o} + \gamma u_\mathrm{a}\,, \\ & \frac{\partial v_\mathrm{o}}{\partial t} + f u_\mathrm{o} + \frac{g_\mathrm{o}}{r}\frac{\partial h_\mathrm{o}}{\partial\phi} = -a v_\mathrm{o} + \gamma v_\mathrm{a}\,, \\ & \frac{\partial h_\mathrm{o}}{\partial t} + \frac{d_\mathrm{o}}{r\cos\phi}\left(\frac{\partial u_\mathrm{o}}{\partial\lambda} + \frac{\partial(v_\mathrm{o}\cos\phi)}{\partial\phi}\right) = -b h_\mathrm{o} \end{aligned} \tag{3.2}$$

where u_a and v_a are the zonal and meridional wind velocity components, h_a the depth perturbation against the equivalent depth d_a, u_o and v_o are the zonal and meridional currents, and h_o the depth perturbation against the equivalent depth d_o. Motion is damped by Rayleigh friction and

Newtonian cooling with respective parameters A and B for the atmosphere and a and b for the ocean. The dynamic coupling coefficient is γ and the thermodynamic coupling parameter is α. The values chosen for α and γ are as in Yamagata (1985). The Coriolis parameter is f, gravity is g, the reduced gravities for the atmosphere and the ocean are g_a and g_o, respectively, and the coordinates (r, λ, ϕ) are the radial distance, the longitude and the latitude, respectively. The atmospheric long gravity wave speed $c_a = 66\ \mathrm{ms}^{-1}$ is chosen such that the equatorial radius of deformation is about 2000 km and the corresponding equivalent depth is around 400 m. For the ocean, a long gravity wave speed of $c_o = 1.4\ \mathrm{m\,s}^{-1}$ gives an equatorial radius of deformation of 250 km and an equivalent depth of about 20 cm. This value of the equivalent depth is appropriate if the sharp, shallow tropical thermocline is considered to be an interface between the warm surface water and the cold deep water.

Eqs. (3.1–3.2) were solved numerically following PYP. The ocean model was run in a 200° longitudinal extent between 30° S and 30° N. Zonally, the atmospheric domain was larger by 20° and was cyclic. The solid-wall boundary condition was set for the ocean and for the atmosphere along the northern and southern boundaries. The grid spacing was 1° latitudinally and longitudinally in both the atmosphere and the ocean. The coupling between the atmosphere and the ocean took place once a day. For the rest of the day, the driving forces for the atmosphere and ocean were held fixed.

Since most real operational simple models have an equilibrium atmosphere, the initial conditions for the atmospheric model were specified to be zero and only the oceanic initial conditions will be adjusted by the adjoint data assimilation. Wind and SLH data were assimilated as in LH1 with the cost function

$$J = \int_0^T \int_S \frac{1}{2} \{(u_a - u_a^{obs})^2 + (v_a - v_a^{obs})^2 + (h_o - h_o^{obs})^2\}\, ds\, dt\,, \tag{3.3}$$

where the integrations are over the spatial domain S and the time period T, with the "observed" values indicated by the superscript "obs". Since we scaled the 6 model variables to be of the order 1 by using the typical values in our model, which are 400 m and 20 cm for the atmospheric and oceanic equivalent depths, and $1\ \mathrm{m\,s}^{-1}$ for the wind and the ocean current, the three data fields were weighed equally, i.e., the weighting matrices were chosen as 1/2 times the identity matrix.

Following the procedure in Section 2, the continuous adjoint model is derived from (3.1–3.2) with the cost function specified in (3.3):

$$\begin{aligned}
&\frac{\partial u_a^*}{\partial t} - f v_a^* + \frac{d_a}{r\cos\phi}\frac{\partial h_a^*}{\partial \lambda} = A u_a^* - \gamma u_o^* + (u_a - u_a^{obs})\,,\\
&\frac{\partial v_a^*}{\partial t} + f u_a^* + \frac{d_a}{r}\frac{\partial h_a^*}{\partial \phi} = A v_a^* - \gamma v_o^* + (v_a - v_a^{obs})\,,\\
&\frac{\partial h_a^*}{\partial t} + \frac{g_a}{r\cos\phi}\left(\frac{\partial u_a^*}{\partial \lambda} + \frac{\partial v_a^*\cos\phi}{\partial \phi}\right) = B h_a^*\,,
\end{aligned} \tag{3.4}$$

$$\begin{aligned}
&\frac{\partial u_o^*}{\partial t} - f v_o^* + \frac{d_o}{r\cos\phi}\frac{\partial h_o^*}{\partial \lambda} = a u_o^*\,,\\
&\frac{\partial v_o^*}{\partial t} + f u_o^* + \frac{d_o}{r}\frac{\partial h_o^*}{\partial \phi} = a v_o^*\,,\\
&\frac{\partial h_o^*}{\partial t} + \frac{g_o}{r\cos\phi}\left(\frac{\partial u_o^*}{\partial \lambda} + \frac{\partial v_o^*\cos\phi}{\partial \phi}\right) = b h_o^* + \alpha h_a^* + (h_o - h_o^{obs})\,.
\end{aligned} \tag{3.5}$$

Variables with an asterisk denote the adjoint variables (i.e., Lagrange multipliers) of their counterparts in (3.1)–(3.2). In the adjoint model, the forcing terms on the RHS of the atmospheric momentum equations and the oceanic mass equation measure the misfit between the model solution and the data, and the damping processes take opposite signs because of backward integration. Notice that there exists an inverse coupling in the coupled adjoint model in which the atmospheric adjoint model is driven "dynamically" by the "adjoint current" and the oceanic adjoint model is driven "thermodynamically" by the atmospheric "adjoint depth perturbation". These terms allow information exchanges in the coupled adjoint model, so that the model trajectories of the two single models can be tuned together.

4. Setup of numerical experiments

A 40-day run simulating the unstable local-growth of the atmosphere-ocean coupled system

Table 1. *Parameters used in the coupled atmosphere-ocean model*

Parameter	Values	
A	2.3×10^{-6}	s^{-1} [(5 days)$^{-1}$]
B	7.7×10^{-7}	s^{-1} [(15 days)$^{-1}$]
$c_a = (gd_a)^{1/2}$	66	m s^{-1}
g_a	40	cm s^{-2}
a	1.2×10^{-7}	s^{-1} [(100 days)$^{-1}$]
b	1.2×10^{-7}	s^{-1} [(100 days)$^{-1}$]
$c_o = (gd_o)^{1/2}$	1.4	m s^{-1}
g_o	2.0	cm s^{-2}
α	1.0×10^{-2}	s^{-1}
γ	5.0×10^{-7}	s^{-1}

in the tropical Pacific was taken as a control (see PYP for details on this control run). The parameter set (Table 1) used was that of Yamagata (1985). The three initial oceanic conditions (h_o, u_o, and v_o) in the control run were the oceanic states after having relaxed the ocean model (in the absence of atmospheric forcing or coupling) for 10 days from an initial sea level height anomaly of Gaussian distribution centered around the equator in the middle of the ocean, i.e.,

$$h = 0.5 \exp[-(x^2 + y^2)/(500\ km)^2]\ \text{cm}, \tag{4.1}$$

where the perturbation amplitude of 0.5 cm relative to an equivalent depth of 20 cm would correspond to a 2.5 m displacement of the thermocline from a mean thermocline depth of 100 m. The formulae calculating the gradients of the cost function with respective to the six parameters and the three oceanic initial conditions are:

$$\begin{aligned}
\delta J/\delta A &= \textstyle\iint_{T,S} (u_a^* u_a + v_a^* v_a)\, ds\, dt\,, \\
\delta J/\delta B &= \textstyle\iint_{T,S} (h_a^* h_a)\, ds\, dt\,, \\
\delta J/\delta \alpha &= \textstyle\iint_{T,S} (h_a^* h_o)\, ds\, dt\,, \\
\delta J/\delta a &= \textstyle\iint_{T,S} (u_o^* u_o + v_o^* v_o)\, ds\, dt\,, \\
\delta J/\delta b &= \textstyle\iint_{T,S} (h_o^* h_o)\, ds\, dt\,, \qquad (4.2) \\
\delta J/\delta \gamma &= \textstyle\iint_{T,S} (u_o^* u_a + v_o^* v_a)\, ds\, dt\,, \\
\delta J/\delta u_o &= -u_o^*(0)\,, \\
\delta J/\delta v_o &= -v_o^*(0)\,, \\
\delta J/\delta h_o &= -h_o^*(0)\,.
\end{aligned}$$

In all experiments conducted, the wind and SLH data were available once per day, the temporal density of the TAO array network. When studying the sufficiency of observations from the TAO array in retrieving the parameters and/or the initial conditions, the wind and SLH data were sampled from the model domain between 8° S and 8° N, at 2° latitude by 15° longitude resolution (about 350 data per day). Unless specifically mentioned, guessed parameters used values 10% smaller than the control, while guessed initial conditions were generated by first systematically scaling up the magnitude of h in (4.1) by 10% and then relaxing the ocean model the same way as when obtaining the control initial conditions.

As in LH1, all experiments were performed using a limited-memory quasi-Newton method in the NAG FORTRAN Library routine, E04DGF. The root mean square error (RMSE) and the correlation coefficient between retrieved initial conditions and the true initial conditions were used to score the quality of the estimations for the initial conditions, while the relative estimation error (i.e. (estimated value − true value)/(true value)) was used to appraise the quality of the estimated parameters. All the experiments presented in this paper have converged, with the minimization routine requiring typically 10^2 iterations to converge.

5. Results from numerical experiments

5.1. *The mutual influence of parameter estimation and initialization*

This issue was tested first by conducting experiments with the wind and the SLH data provided everywhere in the model domain but only once per day, retrieving (i) both the 6 parameters and the 3 initial conditions, (ii) the parameters only or (iii) the initial conditions only (Fig. 1). Although the data amount was far more than the unknowns, the retrieval quality when retrieving both parameters and initial conditions was generally poorer than those retrieving either of them, with the RMSE for retrieving the initial conditions and the relative estimation error for retrieving the parameters in (i) being typically one order of magnitude larger than those in (ii) and (iii). This agrees with many studies which found that more model unknowns require more data, as information insufficiency could cause the retrieval quality to deteriorate (Thacker and Long, 1988; LH1).

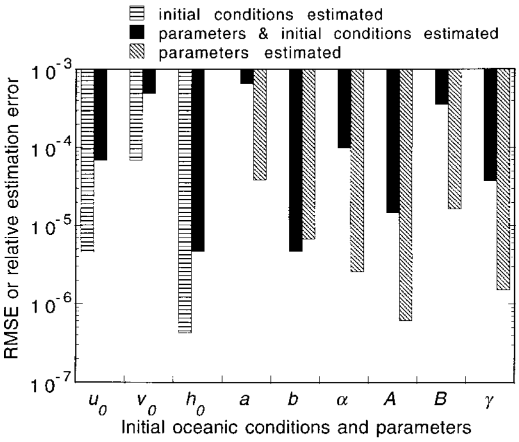

Fig. 1. The RMSE in retrieving the 3 oceanic initial conditions and the relative estimation error in retrieving the 6 parameters by assimilating the wind and SLH data $1\times$ per day at every spatial grid point. In one run, both the oceanic initial conditions and the parameters were retrieved together, while in other runs, only the oceanic conditions or only the parameters were retrieved. As this bar chart is plotted with a logarithmic vertical axis, a *longer* bar represents a *smaller* error. The RMSE has no units because the model variables have been non-dimensionalized.

However, one must be careful when drawing this conclusion for problems involving parameter estimation, e.g., parameter b was better estimated when retrieved together with the initial conditions (Fig. 1). Thus information insufficiency may not be the only cause for poor estimation, the existence of secondary minima in the cost function can result in poor parameter estimation even when relatively more information is available. The errors among different experiments in Fig. 1 allows us to compare the capability of retrieving the true solution in each experiment. The differences in the retrieval errors, judging from their small magnitudes, are not likely to have large effects on the model, especially for the retrieved parameters. Fig. 2 shows that the cost function and the gradient norm for the experiment estimating both the parameters and initial conditions can be minimized to the order of 10^{-5}–10^{-7}.

To investigate this further, we ran experiments identical to those in Fig. 1 except that only the wind and SLH data in the TAO array were assimilated (Fig. 3). The estimation errors were typically 1–3 orders of magnitude larger than those in Fig. 1. Though the data amount was still

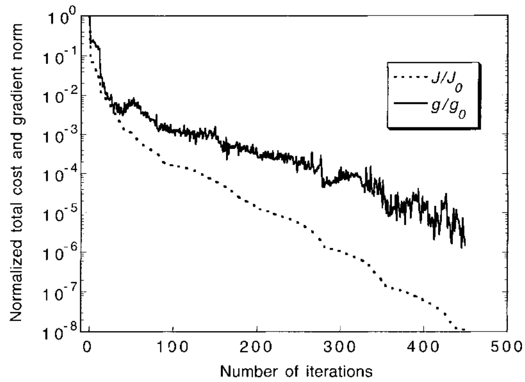

Fig. 2. Change of the total cost function J and the gradient norm $g = |\boldsymbol{g}|$ with iterations during the minimization. Both the parameters and initial conditions were estimated and the winds and the SLH data were provided once per day at every grid point. Both the cost function (dashed curve) and the gradient norm (solid curve) have been normalized with respect to their initial values (J_0 and g_0) before plotting.

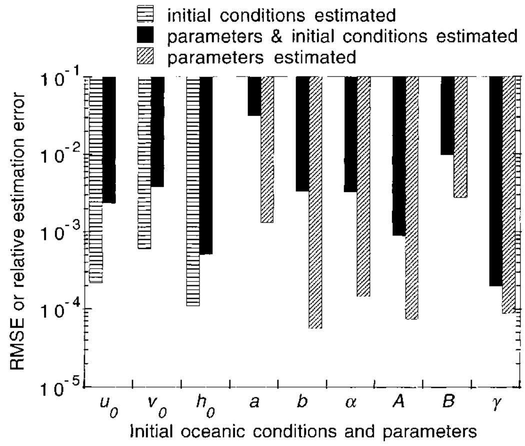

Fig. 3. Same as Fig. 1 except that only the wind and SLH data in the TAO array were assimilated $1\times$ per day.

more than the unknowns in these experiments, the quality of retrieving both the parameters and the initial conditions was poorer than those retrieving either one, reflecting the fact that now information insufficiency becomes dominant over the effect of secondary minima. Nevertheless, the secondary minima effect can still be detected in Fig. 3, e.g., relative to the results in Fig. 1, the estimation of the h_0 and B in Fig. 3 deteriorated more in the experiments retrieving only the parameters or only the initial conditions than in the experiment estimating both parameters and initial conditions.

Though retrieving both the parameters and initial conditions could in general result in the degradation of their estimation quality, some characteristics of the results tended to follow those from experiments where only the parameters or the initial conditions were retrieved. Among the parameters, A, α and γ were generally the most accurately retrieved (as found in LH1), and among the initial conditions, h_o was easier to retrieve than the current components (as found in LH2).

The major factor determining the accuracy of parameter estimation and initialization is likely to be information sufficiency, which seems dependent on the nature of the link between model unknowns and input data. Unknowns directly associated with input data through the constraining model equations were more accurately estimated than those without direct relation with input data. That parameters A and b were relatively better estimated than B and a reflects the fact that parameter A was bounded by the atmospheric momentum equations where winds were available directly, and b was constrained by the oceanic thermodynamic equation where SLH data were available. That h_o was better estimated than the currents is because SLH data were directly available. In principle, the information available in the whole system could be used for estimating all the model unknowns. This problem involving information transfer between the atmosphere and the ocean is discussed in the Subsection 5.2.

5.2. *The information transfer between the atmosphere and the ocean*

The effectiveness of information transfer between the atmosphere and the ocean was studied by conducting experiments assimilating the wind data or the SLH data in the TAO array. Both the parameters and initial conditions were estimated together in Fig. 4. By assimilating only the wind data, the estimation of the oceanic damping parameters (a, b) and the thermodynamic and dynamic coupling parameters (α, γ) deteriorated, while the atmospheric damping parameters (A, B) were better estimated. In contrast to the wind case, the atmospheric damping parameters (A, B) and the thermodynamic and dynamic coupling parameters (α, γ) deteriorated when assimilating only the SLH

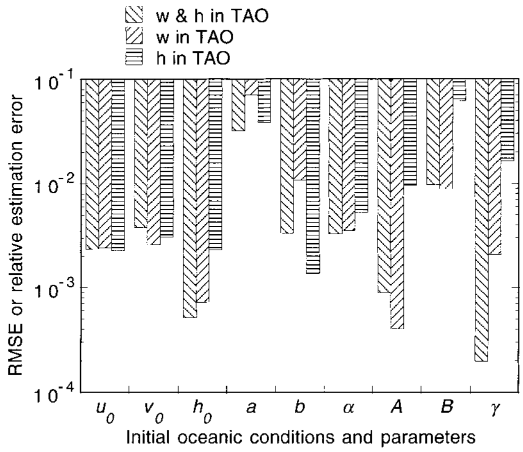

Fig. 4. RMSE in retrieving the 3 oceanic initial conditions and the relative estimation error in retrieving the 6 parameters by assimilating the wind or the SLH data (or both) in the TAO array $1\times$ per day. The initial conditions and the parameters were retrieved together.

data. That the estimation accuracy of the atmospheric damping parameters depends more on the wind data and the estimation accuracy of the oceanic damping parameters depends more on the SLH data indicates that information transfer has less impact on damping parameter estimations. However, the wind and the SLH data were both important for retrieving the coupling parameters because the parameters α and γ were constrained by the whole system and extracted information from both the wind and SLH data.

Compared with parameter estimation, the retrieval of initial conditions was generally less sensitive to the assimilated data type (wind or SLH). With only wind assimilated, the estimated oceanic initial conditions were comparable to those using both the wind and SLH data (Fig. 4). This indicates that information transfer from atmosphere to ocean is important in estimating initial conditions. By assimilating only the SLH data, the retrieval quality of the initial condition h_o was poorer than using both the wind and SLH data, or just the wind data. This is consistent with LH2 that the SLH data in the TAO array were less critical than the wind data.

5.3. *The effect of initial guess on the retrieval of parameters and initial conditions*

When parameters and initial conditions are retrieved together, the influence of the initially guessed values (for the parameters or the initial conditions) on the retrieval quality is of practical interest. The sensitivity of the retrieval quality to 3 types of initial guesses in which different errors were introduced into the initially guessed values was studied. In case 1, the initial guesses for the initial conditions were systematically scaled up by 10% (case 1a) or 50% (case 1b) from those used for Fig. 1, (while the initial guesses of the 6 parameters were unchanged). In case 2, the initial guesses for the initial conditions were generated by placing the origin (x_0, y_0) of the initial perturbation h in (4.1) 4 or 8 grid points (case 2a and case 2b respectively) to the left of where the initial perturbation h for the control run was generated. Again, the initial guesses for the parameters were unchanged. In case 3, the 6 initially guessed parameters were systematically scaled up by 50% (case 3a) or 100% (case 3b) from those for Fig. 1, while keeping the initially guessed initial conditions unchanged. Case 1 brought larger "magnitude" errors to the initially guessed initial conditions than those used for Fig. 1, while the initially guessed initial conditions in case 2 had "phase" errors as the warm anomaly had been shifted zonally.

Fig. 5 depicts the influence of initial guesses on the retrieval quality (the RMSE) of the three oceanic initial conditions with both the wind and the SLH data provided everywhere and once per day. While errors in the initial guesses generally led to poorer retrieval quality, the drop in retrieval quality was different for the various types of initial guess errors. For instance, increasing the phase error (i.e., going from csse 2a to case 2b) seemed far more detrimental than increasing the magnitude error (i.e., going from case 1a to case 1b). The effect of the errors in the guessed parameters on the retrieval quality of the initial conditions were relatively small (cases 3a, b), presumably with well guessed initial conditions, the errors in the guessed parameters did not cause much damage to the initial condition retrieval.

The correlation between the retrieved initial condition and the true initial condition was shown in Fig. 6 for the various types of initial guess errors, with h_o showing high correlations for all experiments. For the currents, errors in the guessed parameters (cases 3a, b) had negligible effects on their retrieval, but phase errors in the guessed initial conditions (cases 2a, b) caused far more damage than magnitude errors (cased 1a, b).

The retrieval quality of the 6 parameters were also affected by errors in the initial guess (Fig. 7). The magnitude errors in the guessed initial conditions (case 1a) caused the degradation of the parameter estimation, with the relative estimation error increasing by 1–2 orders of magnitude. Larger magnitude errors in case 1b did not cause poorer estimation than case 1a, except for the parameter B. Large phase errors in the guessed initial conditions (cases 2a, b) also showed detrimental effects on parameter estimation, with the

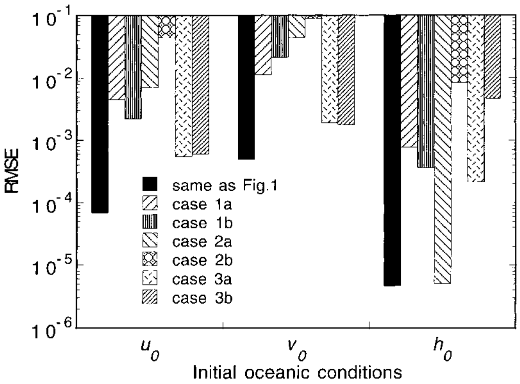

Fig. 5. RMSE in retrieving 3 oceanic initial conditions, for different cases of initial guesses, by assimilating the wind and SLH data 1× per day at every spatial grid point.

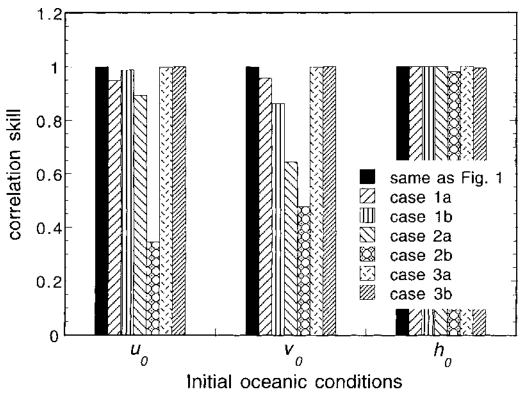

Fig. 6. Correlation between the retrieved initial conditions and the true initial conditions, where the 3 oceanic initial conditions were retrieved from different cases of initial guesses by assimilating the wind and SLH data 1× per day at every spatial grid point.

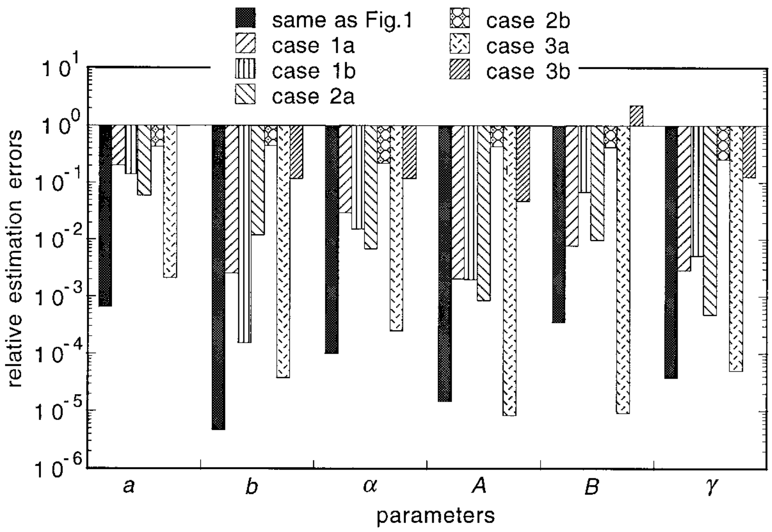

Fig. 7. Relative estimation error in retrieving the 6 parameters, from different cases of initial guesses, by assimilating the wind and SLH data 1 × per day at every spatial grid point.

damage increasing substantially with increased phase error (cf. case 2b with case 2a). Errors in the guessed parameters also lead to degradation in the retrieval quality for all the parameters, with very strong degradation as the errors were increased from case 3a to case 3b.

6. Summary and discussion

In this paper, a general formalism for simultaneously retrieving both model parameters and initial conditions in coupled atmosphere-ocean models by adjoint data assimilation was developed. This method was applied to the Philander et al. (1984) simple equatorial coupled model (with the atmosphere and the ocean each represented by a one-layer linear shallow-water model), where 3 initial oceanic fields (SLH h_o and the current components u_o and v_o) and 6 parameters (4 damping parameters and 2 coupling parameters) were estimated. A series of 40-day identical twin experiments was designed to study (i) the sensitivity of initial condition retrieval to simultaneous parameter estimation, and vice versa, (ii) the information transfer efficiency between the atmosphere and the ocean in determining the parameters and initial conditions, and (iii) the impact of initial guesses on the retrieval of the parameters and initial conditions.

When retrieving simultaneously the parameters and initial conditions of simple models by assimilating both the wind and SLH data once per day at every grid point, the retrieval quality of the parameters and initial conditions was generally poorer than those from the retrieval of only the parameters or only the initial conditions. However, the retrieval quality for an individual parameter, where both parameters and initial conditions were retrieved, could be better than when only the parameters or the initial conditions were retrieved. This resulted from the existence of secondary minima, which could cause poor retrieval even for problems with relatively more information than unknowns. It was found that parameter estimation, though dealing with far fewer unknowns than initialization (where for our model, there were 7420 unknowns in the three initial oceanic conditions), needed as much information as initialization did, and that the parameters A, α and γ were the most readily retrieved among the parameters. Among the initial conditions retrieved, h_o was easier to retrieve than the current components. It seemed that data amount versus the number of unknowns was not a good measure for estimating the retrieval outcome, especially for problems involving parameter estimation. Therefore, having a priori information for both parameters and initial conditions, or at least for the parameters, should be helpful in obtaining a better retrieval (see LH1).

By using the wind and SLH data only in the TAO array, the information was insufficient to retrieve the parameters and initial conditions, as the RMSE for the retrieved initial conditions and the relative estimation retrieval for the retrieved parameters were 1–3 orders of magnitude larger than those obtained by having data available everywhere.

When retrieving parameters and initial conditions simultaneously, wind and SLH data were found to be of comparable value in determining the initial conditions, but not in parameter estimation, i.e., the wind data were more favorable for determining atmospheric damping parameters, and the SLH data for the oceanic damping parameters. The wind and the SLH data were both important for retrieving the coupling parameters, especially for γ, the dynamic coupling parameter.

Poor initial guesses for the initial conditions and parameters in general lead to degradation of the retrieval quality, with phase errors in the initial conditions more detrimental than magnitude errors, especially for the current field retrieval. The effects of the guessed parameters on the retrieval quality of the initial conditions were weaker than those of the guessed initial conditions. The errors in the guessed parameters and the phase errors in the guessed initial conditions produced detrimental effects on parameter retrieval.

Our conclusion "when retrieving simultaneously the parameters and initial conditions of simple models by assimilating both the wind and SLH data once per day at every grid point, the retrieval quality of the parameters and initial conditions was generally poorer than those from the retrieval of only the parameters or only the initial conditions" is true only because in our runs retrieving only the initial conditions, we assumed knowledge of the true parameters, and vice versa for the runs retrieving only the parameters. In reality, one does not know the true parameters when one is retrieving the initial conditions. In some experiments (not shown), we found that when the model parameter values were modestly off from the control, the retrieval for only the initial conditions failed to converge to meaningful values. Hence, in practice it may not be possible to retrieve initial conditions without simultaneously retrieving the parameters. This highlights the theoretical and practical importance of this work (where parameters were retrieved together with initial conditions) over our earlier work LH2, where only initial conditions were retrieved for models with true parameters. For the other situation, when retrieving parameters only, the lack of accurate initial conditions tends to be less of a problem. Finally, we would point out that as our simplistic model has many unrealistic features, e.g., the heat loss from the ocean leaves the ocean unaffected, and an increase in the thermocline depth necessarily causes heating of the atmosphere, it can only be run on an intraseasonal scale. For operational El Niño prediction models, a period of several months is probably needed to assimilate the tropical atmospheric and oceanic data into a more sophisticated model. Therefore, some of the conclusions presented here may not necessarily hold for longer periods of data assimilation with more realistic models.

7. Acknowledgements

We are grateful to Ron C. Pacanowski for kindly providing us the GFDL C-grid shallow-water model code, and to Ralf Giering for his AMC compiler in helping us to develop the adjoint code. This work was supported by research grants to W. Hsieh from the Natural Sciences and Engineering Research Council of Canada and Environment Canada.

REFERENCES

Allen, M. R. and Davey, M. K. 1993. Empirical parameterization of tropical ocean atmosphere coupling: "The inverse Gill Problem". *J. Climate* **3**, 509–530.

Barnston, A. G., Van den Dool, H. ., Zebiak, S. E., Barnett, T. P., Ji, M., Rodenhuis, D. R., Cane, M. A., Leetmaa, A., Graham, N. ., Ropelewski, C. R., Kousky, V. E., O'lenic, E. A. and Livezey, R. E. 1994. Long-lead seasonal forecasts; where do we stand? *Bull. Amer. Meteor. Soc.* **75**, 2097–2114.

Chen, D., Zebiak, S. E., Busalacchi, A. J. and Cane, M. A. 1995. An improved procedure for El Niño forecasting: implications for predictability. *Science* **269**, 1699–1702.

Ji, M., Kumar, A., and Leetmaa, A. 1994. An experimental coupled forecast system at the National Meteorological Center: Some early results. *Tellus* **46A**, 398–418.

Le Dimet, F., and Talagrand, O. 1986. Variational algorithms for analysis and assimilation of meteorology observations: theoretical aspects. *Tellus* **38A**, 97–110.

Lu, J. and Hsieh, W. W. 1997. Adjoint data assimilation in coupled atmosphere-ocean models: Determining model parameters in a simple equatorial model. *Q. J. R. Meteorol. Soc.* **123**, 2115–2139.

Lu, J. and Hsieh, W. W. 1998. Adjoint data assimilation in coupled atmosphere-ocean models: determining initial conditions in a simple equatorial model. *J. Met. Soc. Japan,* in press.

Palmer, T. N. and Anderson, D. L. T. 1994. The prospectives for the seasonal forecasting. A review paper. *Q. J. R. Meteorol. Soc.* **120**, 755–793.

Philander, S. G. H., Yamagata, T. and Pacanowski, R. C. 1984. Unstable air-sea interactions in the tropics. *J. Atmos. Sci.* **41**, 604–613.

Talagrand, O. and Courtier, P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Q. J. R. Meteorol. Soc.* **113**, 1311–1328.

Thacker, W. C. 1988. Fitting models to inadequate data by enforcing spatial and temporal smoothness. *J. Geophys. Res.* **93**, 10 655–10 665.

Thacker, W. C. and Long, R. B. 1988. Fitting Dynamic to data. *J. Geophys. Res.* **93**, 1227–1240.

Tziperman, E., and Thacker, W. C. 1989. An optimal control/adjoint equations approach to studying the oceanic general circulation. *J. Phys. Oceanogr.* **19**, 1471–1485.

Yamagata, T. 1985. Stability of a simple air-sea coupled model in the tropics. Coupled Ocean-Atmosphere Models, Nihoul, J. C. J. (ed). *Elsevier Oceanogr. Ser.* **40**, 637–657.

Yu, L. and O'Brien, J. J. 1991. Variational estimation of the wind stress drag coefficient and the oceanic eddy viscosity profile. *J. Phys. Oceanogr.* **21**, 709–719.

Zebiak, S. E. and Cane, M. A. 1987. A model El Niño-Southern Oscillation. *Mon. Wea. Rev.* **115**, 2262–2278.