On the Initial Condition in Parameter Estimation

LISAN YU* AND JAMES J. O'BRIEN**

Mesoscale Air–Sea Interaction Group, Florida State University, Tallahassee, Florida

15 October 1991 and 14 February 1992

1. Introduction

In an earlier work, Yu and O'Brien (1991; hereafter referred to as YOB) have successfully determined the oceanic vertical eddy viscosity profile in the open ocean and the wind-stress drag coefficient from meteorological and oceanographic observations from LOTUS (Bowers et al. 1986). The variational optimal control method used ensures an optimal estimate of the unknown current field. However, the optimal current field obtained in YOB did not agree well with the observations in the 50–100-m range, though it did have good agreement with the observations in the upper 25 m. The question arose as to how one can minimize these remaining model-data misfits, or if one can do better.

The variational optimal control method (Lewis and Derber 1985; Le Dimet and Talagrand 1986; Talagrand and Courtier 1987; Thacker and Long 1988) uses a numerical model to obtain the solution of the model equations that best fits, in the generalized least-squares sense, the observations within some space–time domain. In YOB, the surface forcing and the turbulent mixing process were treated as the uncertainties of the model dynamics, while the initial condition was assumed to be known a priori by setting it to the observed initial ocean state. The variational optimal control procedure sought the space–time trajectory that best fits the observations throughout the assimilation period. Thus, the best space–time trajectory was established by adjusting only the unknown parameters of surface forcing and turbulent mixing for the model equations, which were used as strong constraints in the minimization process. However, the chosen model initial condition in YOB may be the part that is not entirely consistent with the model's intrinsic dynamics. The model ocean's memory of the initial information is embedded in the model results and contributes to the model deviation from the observations at later times. Therefore, we conjecture that the addition of the optimal estimation of the initial condition to the model of YOB will improve the model results. In this note, we will show that the model-data misfit is greatly reduced by seeking the optimal initial condition for the model equations.

* Also affiliated with Supercomputer Computations Research Institute, Florida State University, Tallahassee, Florida.

** Also Secretary of Navy Professor, Meteorology and Oceanography, Florida State University, Tallahassee, Florida.

Corresponding author address: Dr. Lisan Yu, Mesoscale Air-Sea Interaction Group, Mail Stop B-174, Florida State University, Tallahassee, FL 32306-3041.

2. Model results and discussion

a. Model description

The model dynamics have been described in detail in YOB. The new cost function is constructed as follows:

$$J(w, c_D, A, w_0) = \frac{1}{2}\sum_{\Omega}(w-\hat{w})^{T}\mathbf{K}_m(w-\hat{w}) + \frac{1}{2}K_cTH(c_D-\tilde{c}_D)^2 + \frac{1}{2}K_aT\sum_{\zeta}(A-\tilde{A})^2 + \frac{1}{2}K_0T\sum_{\zeta}(w_0-\tilde{w}_0)^2, \quad (1)$$

where w is a complex vector of current velocity; w_0 is the current field at the initial state; c_D is the drag coefficient; and A is the eddy viscosity. The superscript T denotes transpose, the summation for the first term is over the observational space–time domain Ω, and the summations for the last two terms are over the space domain ζ. The carrot ($\hat{\ }$) denotes observed data, and the tilde ($\tilde{\ }$) denotes estimated value. Here $\mathbf{K}_m$ is a weighting matrix and theoretically should be the inverse of the observation error covariance matrix. By assuming that the errors in the data are uncorrelated and equally weighted, $\mathbf{K}_m$ is reduced to a unit matrix. The coefficients K_c, K_a, and K_0 are the Gauss precision moduli, representing the smoothness in the iterative process, and are calculated as a part of the step size in the minimization algorithm.

The augmented Lagrange function takes the form:

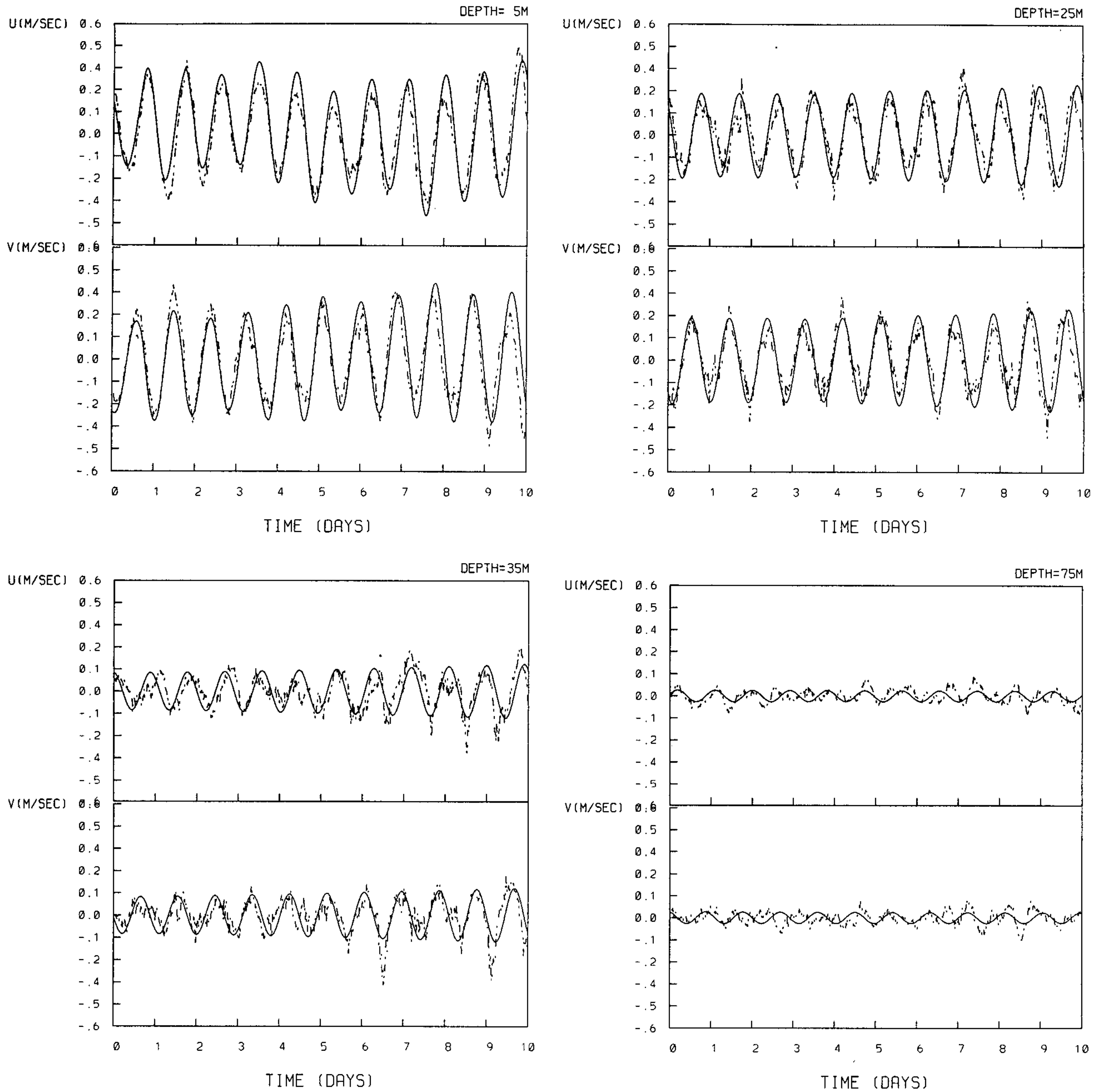

FIG. 1. Comparisons of the modeled (solid) and observed (dashed) current fields at (a) 5 m, (b) 25 m, (c) 35 m, (d) 75 m, and (e) 95 m.

$$L(w, c_D, A, w_0, \lambda) = J(w, c_D, A, w_0) + \sum_{\Omega} \left\{ \lambda, \frac{\partial w}{\partial t} + iw - \frac{\partial}{\partial z}\left(A \frac{\partial w}{\partial z}\right) \right\}, \quad (2)$$

where $\{,\}$ is the inner product of two vectors and λ is an as yet unspecified Lagrange multiplier.

The evolution of the model state is determined by the initial condition, the surface forcing, and the turbulent process; that is, once the optimal values of these control variables are obtained, the corresponding current field will best fit the observations through the data-assimilation period. The problem is then to find the best values of the control parameters, that is, those in which the cost function reaches the minimum and the Lagrange function reaches an extremum. At this point, derivatives of L with respect to all the variables w, c_D, A, w_0, λ are equal to zero. This requirement yields the same adjoint equation and correction equations for c_D

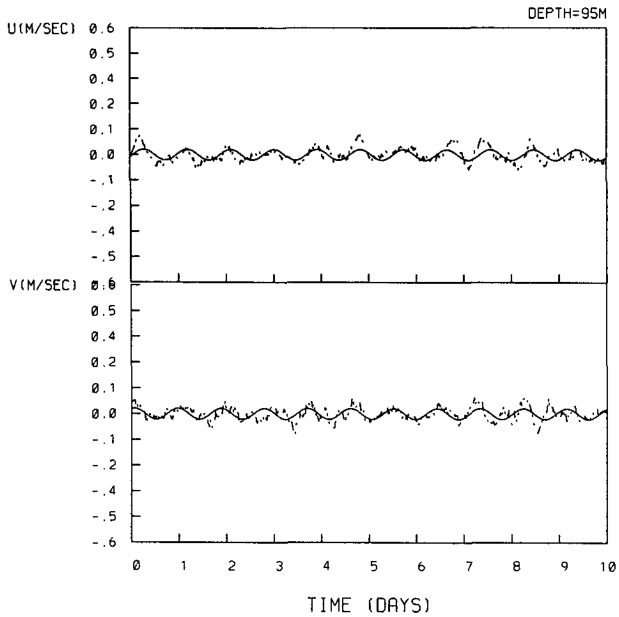

FIG. 1. *(continued)*

and A as those in YOB and, in addition, the correction equation for the initial condition, which is

$$w_0 = \tilde{w}_0 + \frac{1}{K_0 T}\left\{\frac{\partial\lambda}{\partial t} + i\lambda - \frac{\partial}{\partial z}\left(A\frac{\partial\lambda}{\partial z}\right)\right\}\Bigg|_{t=0}. \quad (3)$$

b. Results

The model uses the same datasets as those in YOB. At the end of the assimilation period, the current fields, improved by the optimal values of the initial condition, reproduce the observations more closely. This is demonstrated in Figs. 1a–e. The model-data fits at 5 and 25 m were already quite good in YOB, so no obvious improvement has taken place here except some phase correction for a few oscillation cycles. At 35 m the amplitude changes agree with the observed current field, and the phase shift shown in YOB has improved greatly. We also see much improvement of the modeled current states at 75 and 95 m. The current fields oscillate at the inertial frequency and fall in the major oscillation cycles of the observation, though the inertial signal of the observed current data is, to some degree, contaminated by the noise. These improvements can be seen clearly in Fig. 2, which is the plot of the evolution of the norm of the gradient, the cost function, and the data-misfit in the minimization process. The procedure reaches a minimum after 12 iterations, which is far less than the number of control variables. Because the data are noisy, the model state could not fit the data exactly; that is, the cost function cannot be reduced to zero. Extending the iterative process cannot

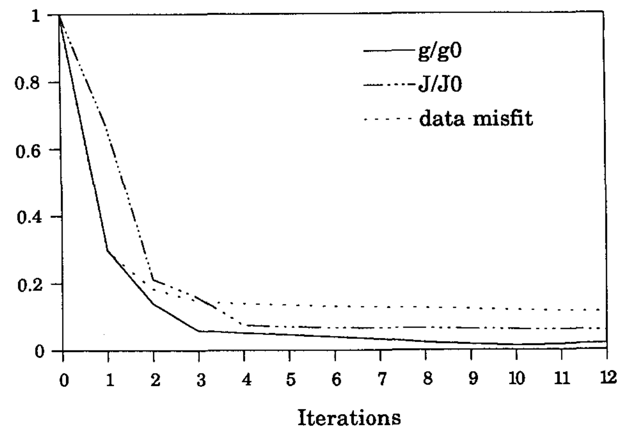

FIG. 2. The evolution of the norm of the gradient (solid), the cost function (dot), and the data misfit (dash) in the minimization process. The procedure reaches a minimum after 12 iterations.

improve the model result because the model is inconsistent with (cannot reproduce) the residual noise in the data. The plot of the correlation coefficient versus the iterations is shown in Fig. 3. Compared with Fig. 7 in YOB, it clearly displays the much better model-data fit in this study, with the correlation coefficient of 0.92 (against 0.87 in YOB) at 5 m, 0.88 (0.81) at 25 m, 0.71 (0.67) at 35 m, 0.34 (0.28) at 75 m, and 0.53 (0.44) at 95 m at the end of the data assimilation cycle.

c. Discussion and conclusion

The ocean has some "memory" of previously existing dynamic processes. By including the initial conditions in the set of adjustable model parameters, it is obvious that the modeled upper-ocean current field has better reproduced the observations. The derived coefficients, that is, the drag coefficient and the eddy viscosity coefficient, are slightly smaller than before.

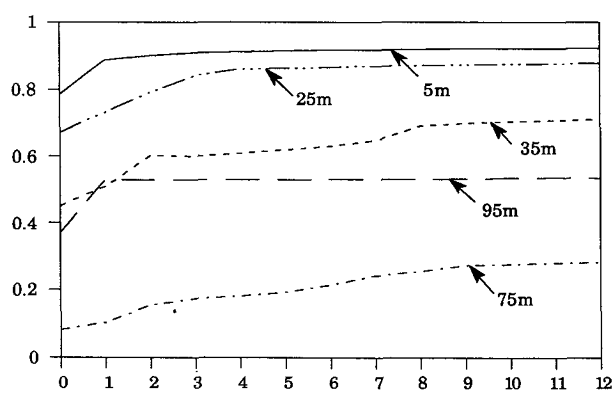

FIG. 3. The variation of the correlation coefficient at different depths with the number of iterations.

The eddy viscosity profile has a maximum value of 1.44×10^{-3} (compare with 2.9×10^{-3} in YOB) at the surface, and the drag coefficient is 1.20×10^{-3} (compare with 1.26×10^{-3} in YOB). This is because in the previous case, by specifying observations as the initial values, the observed current field (inconsistent with the model) inputs too much momentum at the beginning of the assimilation period. This inconsistency in the initial data leads to contributions to the cost function at later times, which can cause unreasonable modifications of the viscosity and drag coefficients in achieving a cost minimum. Therefore, a larger eddy viscosity coefficient and a stronger wind field are needed to achieve the balance required by the model intrinsic dynamics in YOB.

This study also indicates that the same adjoint equation can be used both for assimilating data and for identifying the model coefficients. It is not more costly to control both the initial condition and the model parameters than to control only the model coefficients. In fact, by adding initial condition to the set of the adjustable parameters, the data at all times are given equal weight in fixing the optimal model solution. Since the data at all times of the assimilation cycle influence the recovered initial conditions at zero time, data noise will tend to average out and a less noisy initial model state should be recovered. We conclude that the processes of parameter estimation and initial condition adjustment should not be separated in order to obtain the best model-data fitting.

Acknowledgments. The Mesoscale Air–Sea Interaction Group is supported by ONR Physical Oceanography, NSF (Oceanography and Climate Dynamics), NASA (Oceanic Processes), and NOAA Global Program. We appreciate all this support. This is a contribution to GFDI and SCRI at FSU.

REFERENCES

Bowers, C. M., J. F. Price, R. A. Weller, and M. G. Briscoe, 1986: Data tabulations and analysis of diurnal sea surface temperature variability observed at LOTUS. Woods Hole Oceanographic Institution Tech. Rep., WHOI-86-5, 51 pp.

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

Lewis, J. M., and J. C. Derber, 1985: The use of the adjoint equations to solve a variational adjustment problem with the advective constraints. *Tellus,* **37A,** 309–322.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equations. Part I: Theory. *Quart. J. Roy. Meteor. Soc.,* **113,** 1311–1328.

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

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