SOME NOTES ON DATA ASSIMILATION IN PHYSICAL OCEANOGRAPHY

James J. O'Brien
Mesoscale Air-Sea Interaction Group, The Florida State University, Tallahassee, FL 32306-3041

INTRODUCTION

This paper is a discussion of the author's emphasis on data assimilation in physical oceanography. The work draws on recent work by members of the MASIG Team. Our approach has focused on time-dependent models in which parameters are estimated through data assimilation using the variational adjoint method.

It is useful to adapt a paradigm for classifying all data assimilation methods. I chose to define three groups of assimilation schemes: (A) local polynomial interpolation methods, (B) statistical (including optimal) interpolation methods, and (C) variational numerical analysis methods.

In (A), the idea is to expand the data misfit in terms of some interpolating polynomial in the spatial vicinity of the data location; direct insertion or substitution or "bogusing" are some simple examples; Cressman filters are a commonly used meteorological assimilation technique. No knowledge of the statistical property of the data or the model is used.

In (B), we use statistical information of the data error field or the model variability to determine the adjustment in space and time. In principle, one could estimate the cross-correlation function of the data misfit and adopt some rules to adjust the model solution. The simplest idea is the so-called nudging method where an inverse-time parameter is used to estimate the variability of the data misfit. The most sophisticated example is the Kalman-Bucy filtering method. In all the implementation schemes one should imagine that the physical model is evolving in time, and a moment arrives when a data value is encountered. The data misfit is then added to the model field in time and space. If the covariance matrix structure is primarily spatial, then the simplest time structure for the variability is nudging where a linear time decay processes is added to the prognostic model.

In (C), the assimilation scheme defines a statistically weighted data misfit field, which is minimized in a construct such that the complete physics of the prognostic model is included as dynamical constraints. I will concentrate on examples of this latter method.

THE VARIATIONAL ADJOINT METHOD

The essential ingredients in this data assimilation are a "nice" model, availability of some "useful" data, and a willingness to adjust the model in some manner. Each of these elements must each be appreciated. The model should produce validated solutions that are reasonable and "liked." The data may be estimates of model-dependent state variables or the data may be any function of a dependent state variable as long as an estimate of the function can be calculated from the model output. A simple example would be ocean altimeter cross-over data. The difference in time between two altimeter readings at a point can be estimated from the solution to any ocean model that simulates sea level, and therefore altimeter cross-overs can be assimilated.

For a contrived example, let us consider the following model. Suppose a scalar field, $c(x,t)$, is advected by an unknown advection field, $u(x)$, and other processes are represented by $g(c,\beta)$ where β is a poorly defined parameter. We "like" our model after we guess u,β and the initial conditions, $c'(0,x)$. We acquire some data, $F'(c)$, where $F(c)$ is any function of c which we can estimate from the output of $c(x,t)$. The model is

$$c_t + uc_x = g(c,\beta). \tag{1}$$

There are many avenues to arrive at the variational problem. I choose simply to write down the functional

$$\begin{aligned} H(c,\lambda,u,\beta) = &\int_{x,t}^{T} \lambda(c_t + uc_x - g)\,dx\,dt \\ &+\int_{x,t}^{T} \frac{K_c}{2}(F(c)-F'(c))^2\,dx\,dt \\ &+\int_{x,t}^{T} \frac{K_u}{2}(u-u')^2\,dx\,dt \\ &+\int_{x,t}^{T} \frac{K_\beta}{2}(\beta-\beta')^2\,dx\,dt \end{aligned} \tag{2}$$

where $\lambda(x,t)$ is a Lagrange multiplier, K_c, K_u, K_β are called Gauss precision modulae. The range of space is over all x, say, $x \in [0,L]$ and periodic, e.g., $c(t,x) = c(t,x+L)$. The last three terms are called the cost function, which is to be minimized subject to the contraint that the data, F', and the advection function, u, and the parameter, β, must satisfy the model. The range of time is $[0,T]$; T is a time later than the last observed datum.

The minimum is determined by the usual approach,

$$\frac{\partial H}{\partial \lambda} \text{yields } c_t + u c_x = g(c,\beta) \tag{3}$$

where u and β are now not known.

$$\frac{\partial H}{\partial u} \text{yields } u(x) = u'(x) - \frac{1}{TK_u} \int_t^T \lambda c_x \, dt \tag{4}$$

where we observe that the correction of u from its guess field depends on the average of the product of the Lagrange multiplier and the spatial gradient of the dependent state variable.

$$\frac{\partial H}{\partial \beta} \text{yields } \beta = \beta' - \frac{1}{TLK_\beta} \int_{x,t}^T \lambda \frac{\partial g}{\partial \beta} dx\, dt \tag{5}$$

and

$$\begin{aligned}\frac{\partial H}{\partial c} \text{yields } \frac{\partial \lambda}{\partial t} + (u\lambda)_x &= -\lambda \frac{\partial g}{\partial c} + K_c[F(c) - F'(c)] \frac{\partial F}{\partial c} \\ &+ \int_x \lambda(0,x)\, c(0,x)\, dx + \int_x c(0,x)\, dx.\end{aligned} \tag{6}$$

The next to last integral vanishes using the lemma that a product of periodic functions is periodic. We have used the natural spatial boundary conditions for λ and chosen $\lambda(T,x) = 0$. Note that the last integral is zero except at $t = 0$.

The solution procedure is as follows:

1. Guess u', β', $c'(0,x)$ and calculate the solution forward over the time $[0,T]$ from Eq. (1).
2. Calculate the data misfit, F-F', and the data misfit transfer function, F_c, and integrate Eq. (6) backwards in time from T to zero.
3. Next adjust the initial conditions and $u(x)$ and β using Eqs. (4, 5, and 6) (for $c(0,x)$).
4. Repeat 1,2, and 3 as often as desired in order to assimilate the data, $F(c)$.

There are many advantages to this algorithm. It will almost always converge; thus all the data are used and it is eloquent. I am told that one can contrive a case where it will not converge. There are some disadvantages. It is very expensive because we have to integrate two models and save the solution from both models, particularly when the physical model is nonlinear. It may take many forward and backward integrations to find the miminum. An emphasis of current research is to identify algorithms which find the minimum in as few integrations as possible. The present view is to implement an efficient conjugate gradient algorithm. A further disadvantage is that this method is difficult to teach to scientists. We are beginning to have several simple examples that will demonstrate the method to scientists.

A SIMPLE REAL OCEAN EXAMPLE

There have only been a few modern superb upper-ocean data expeditions that have measured meteorological and upper ocean currents. One such experiment is LOTUS from which Briscoe, Price and Weller have provided us data to develop data assimilation methods. Suppose we wish to assimilate wind and ocean current data and determine the momentum drag coefficient and the mixing function for momentum, $A(z)$.
The model equation is

$$\frac{\partial w}{\partial t} + ifw = \frac{\partial}{\partial z}(A\frac{\partial w}{\partial z}) \tag{7}$$

where $w = u + iv$. The boundary conditions are at

$$z = 0, \text{ and } \rho_w A\frac{\partial w}{\partial z} = \tau \tag{8}$$

where the wind stress is calculated from

$$\tau = \rho_a c_D |w_a| w_a .$$

At the bottom

$$z = -H, \text{ and } A\frac{\partial w}{\partial z} = 0. \tag{9}$$

The initial condition for this dynamic system is at $t = 0$ and $w = w_0$.We chose to nondimensionalize the system as follows:

$$t' = \frac{t}{T_f},\ w' = \frac{w}{U},\ z' = \frac{z}{D},\ A' = \frac{A}{s_a},\ c_D' = \frac{c_D}{s_c},\ w_a' = \frac{w_a}{U_a}$$

where

$$T_f = f^{-1},\ D = \sqrt{\frac{s_a}{f}},\ \text{and } U = \left(\frac{\rho_a}{\rho_w} s_c\right)\frac{U_a^2}{\sqrt{s_a f}},$$

which yields the model

$$\frac{\partial w}{\partial t} + iw = \frac{\partial}{\partial z}\left(A\frac{\partial w}{\partial z}\right) \tag{10}$$

with

$$A\frac{\partial w}{\partial z} = \begin{cases} c_D |w_a| w_a & \text{for } z = 0 \\ 0 & \text{for } z = -\dfrac{H}{D} \end{cases} \tag{11}$$

and $w = w_0$ for $t = 0$.

Following the formalism developed in the previous section, we define the cost function, J:

$$J(w,A,c_D)=\frac{1}{2}K_m\int_t\int_z(w-\hat{w})^2\,d\zeta\,dt \\ +\frac{1}{2}K_aT\int_z(A-\hat{A})^2\,d\zeta+\frac{1}{2}K_cTH(c_D-\hat{c}_D)^2. \tag{12}$$

The functional, L, is the sum of the cost function and the constraint

$$L(w,A,C_D,\lambda)=J+\int_t\int_z\left\{\lambda\left(\frac{\partial w}{\partial t}+ifw-\frac{\partial}{\partial z}\left(A\frac{\partial w}{\partial z}\right)\right)\right\}d\zeta\,d\tau. \tag{13}$$

The solution is found as usual by solving

$$\frac{\partial L(w,A,c_D,\lambda)}{\partial\lambda}=0$$
$$\frac{\partial L(w,A,c_D,\lambda)}{\partial w}=0$$
$$\frac{\partial L(w,A,c_D,\lambda)}{\partial A}=0$$
$$\frac{\partial L(w,A,c_D,\lambda)}{\partial c_D}=0.$$

This yields the model plus

$$\frac{\partial\lambda}{\partial t}+i\lambda+\frac{\partial}{\partial z}\left(A\frac{\partial\lambda}{\partial z}\right)=K_m(w-\hat{w}) \tag{14}$$

$$c_D=\hat{c}_D+\frac{1}{K_cTH}\int_t\left(|w_a|u_a\lambda_{uz=0}+|w_a|v_a\lambda_{vz=0}\right)d\tau \tag{15}$$

$$A=\hat{A}+\frac{1}{K_aT}\int_t\left(\frac{\partial u}{\partial z}\frac{\partial\lambda u}{\partial z}+\frac{\partial v}{\partial z}\frac{\partial\lambda v}{\partial z}\right)d\tau. \tag{16}$$

Note that we have assumed that c_D is a constant and A is only a function of depth. We can rescale the parameters, K, by using

$$\frac{\lambda}{K_M}=\lambda',\ \frac{K_c}{K_m}=K_c',\ \text{and}\ \frac{K_a}{K_m}=K_a'.$$

This yields

$$\frac{\partial\lambda}{\partial t}+i\lambda+\frac{\partial}{\partial z}(A\frac{\partial\lambda}{\partial z})=(w-\hat{w}) \tag{17}$$

$$c_D=\hat{c}_D+\frac{1}{K_c TH}\int_t\left(\left|w_a\right|u_a\lambda_{uz=0}+\left|w_a\right|v_a\lambda_{vz=0}\right)d\tau \tag{18}$$

$$A=\hat{A}+\frac{1}{K_a T}\int_t\left(\frac{\partial u}{\partial z}\frac{\partial\lambda u}{\partial z}+\frac{\partial v}{\partial z}\frac{\partial\lambda v}{\partial z}\right)d\tau. \tag{19}$$

In order to solve these equations we need to define a solution space. This is shown in Figure 1.

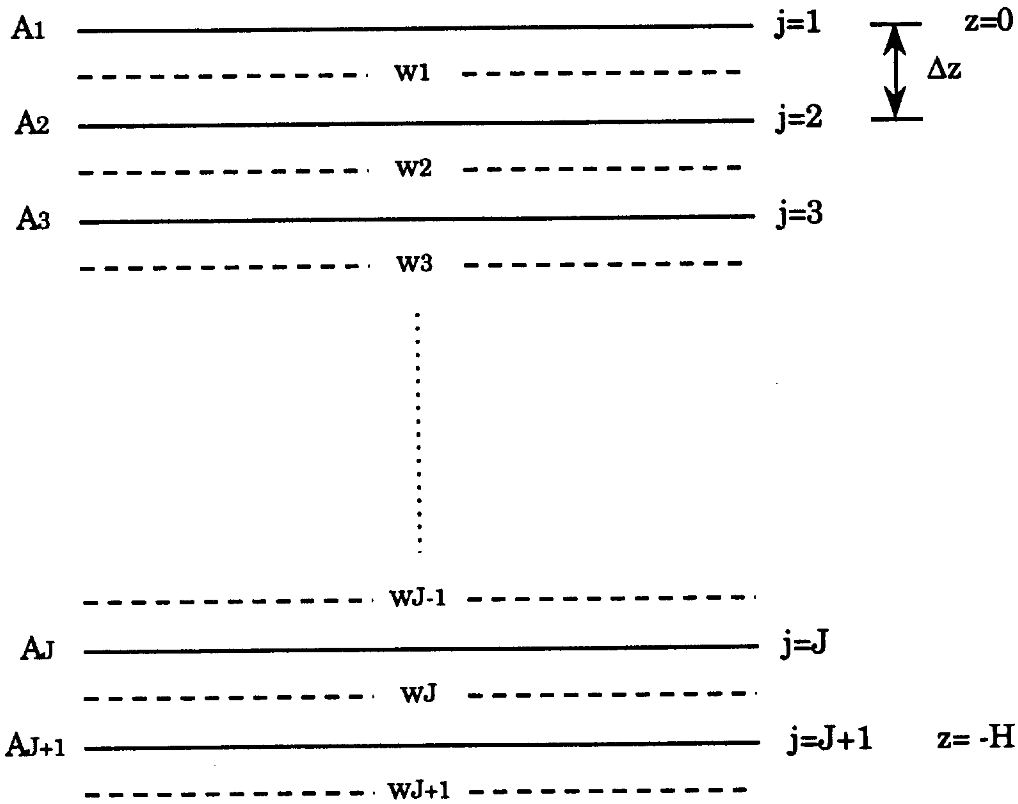

Figure 1. Diagram of the vertical structure of the numerical model.

Our procedure for using the variational method to solve this system can be fully described:

(1) Begin with a best initial estimate for the control parameters A and c_D.
(2) Integrate the model equation (7) forward in time and calculate the value of the cost function.
(3) Compute the data misfits $(w - \hat{w})$.
(4) Integrate the adjoint equation (17) backward in time.
(5) Use equations (18) and (19) to calculate the gradients of the cost functions ∇J corresponding to A and c_D with solutions for λ and w from steps (2) and (4).
(6) With the gradient information, apply the descent algorithm to obtain the new values of A and c_D simultaneously.
(7) Check if the minimization process is done. The convergence criterion is satisfied if $|\nabla J| / |\nabla J_0| < 10^{-2}$, where ∇J_0 is the value at the initial iteration.
(8) Return to step (2) if the optimal solution is not found.

We will demonstrate a solution using currents over 10 days in the summer in the North Atlantic during the LOTUS experiment. Figure 2 shows the observed currents at 5 and 15 meters. Only a low frequency trend has been omitted from the original data. The cost function is shown and the gradient are shown in Figure 3 as a function of interation. Note that the cost function reaches a "practical" minimum in four iterations. The profile of the eddy viscosity coefficient and the drag coefficient are shown in Figure 4. The surface value of 0.003 implies an "Ekman Layer Depth" of about 6-8 m. The comparison of the assimilated data with the data is shown in Figures 5 and 6. It is seen that the model reproduces the current meter data above 65 m quite well and very poorly below. This is a simple example of ocean data assimilation. This research is available in detail in Yu and O'Brien (1991). In Yu and O'Brien (1992), we also change the initial condition with improvement (Table 1). There are additional, completed examples of this work showing how to assimilate sea level. In this report I have not tried to reference all the important works by other research teams.

Table 1.
Change of Correlation Coefficient with Depth

Depth Z	New r*	Old r
5	0.92	0.87
25	0.88	0.81
35	0.71	0.67
75	0.34	0.28
95	0.53	0.44
•Max A	1.4×10^{-3}	2.9×10^{-3}
c_D	1.2×10^{-3}	1.3×10^{-3}

*** Initial condition adjusted.**

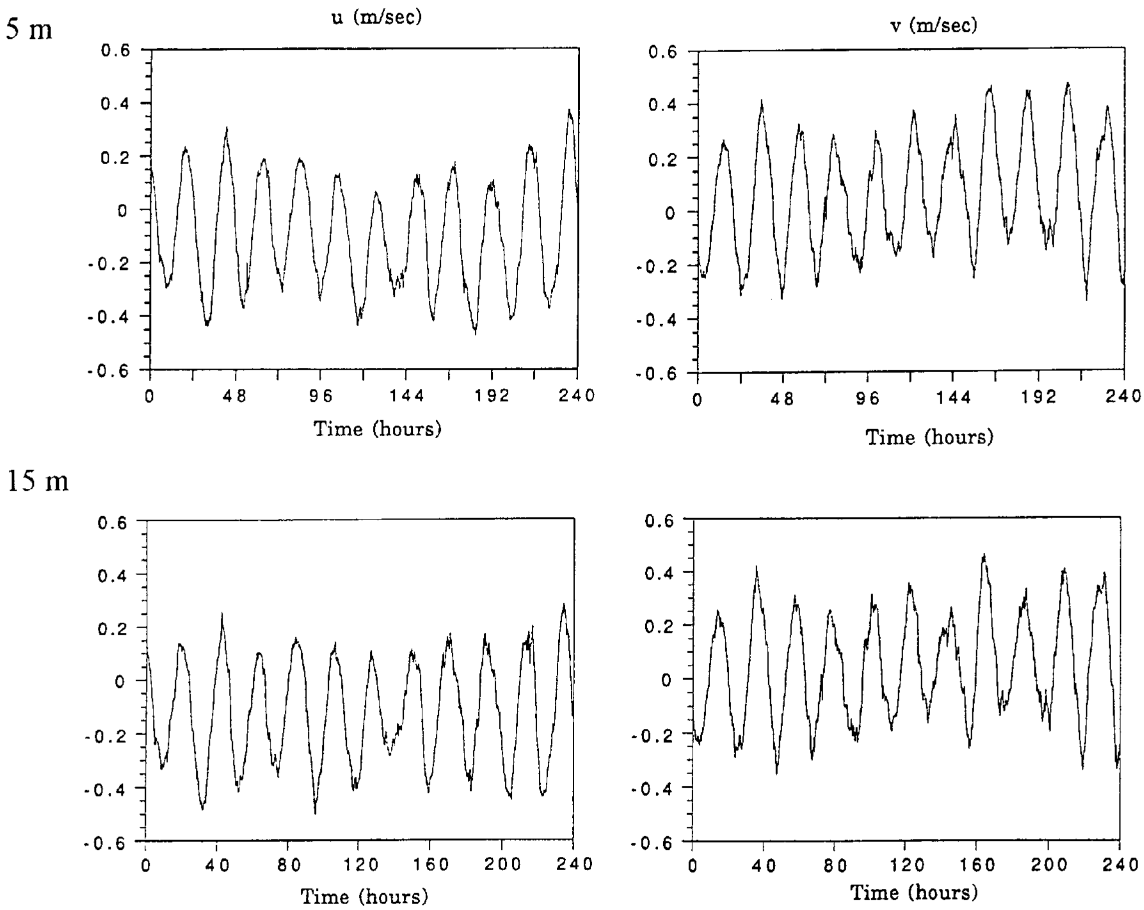

Figure 2. Current observations at 5 m (top) and 15 m (bottom).

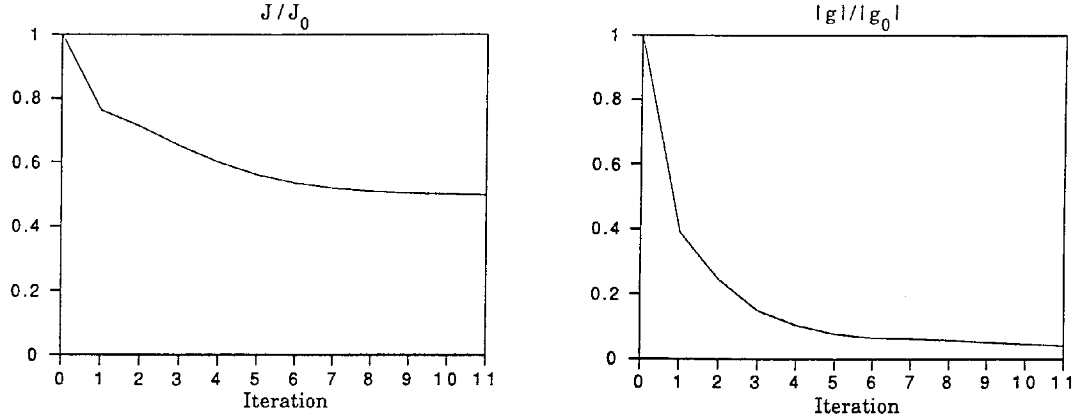

Figure 3. The variation of (left) the cost function and (right) the gradient with the number of iterations.

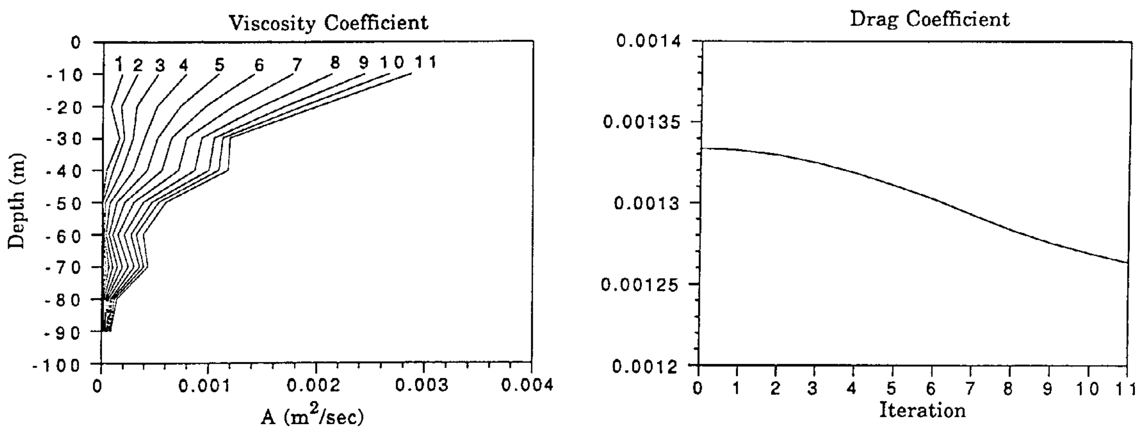

Figure 4. (left) The variation of the eddy viscosity coefficient during the iterative process, and (right) the variation of the drag coefficient with the number of iterations.

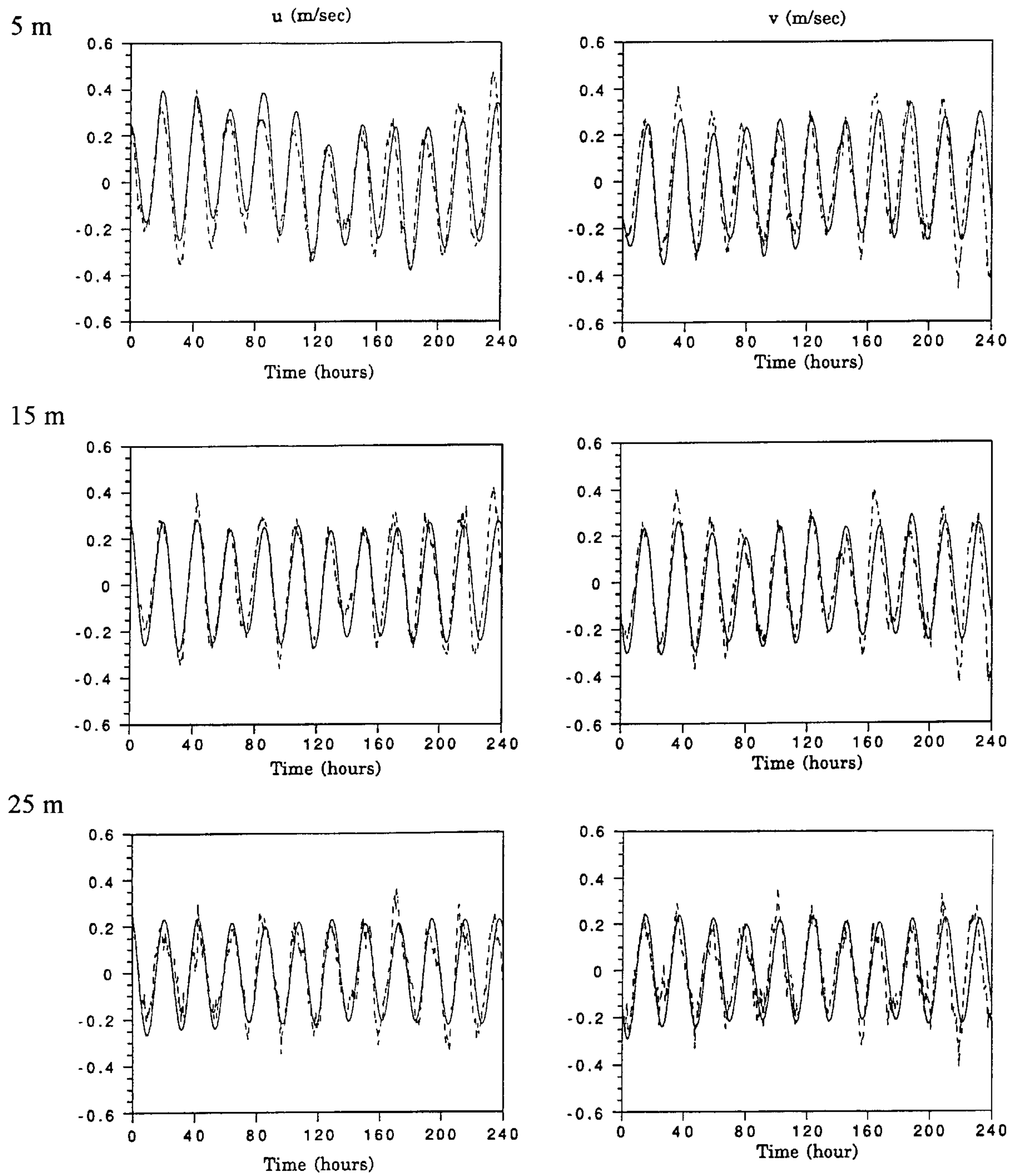

Figure 5. Comparison of modelled (solid lines) and observed (dashed lines) current speeds u (left) and v (right) for 5, 15, and 25 m.

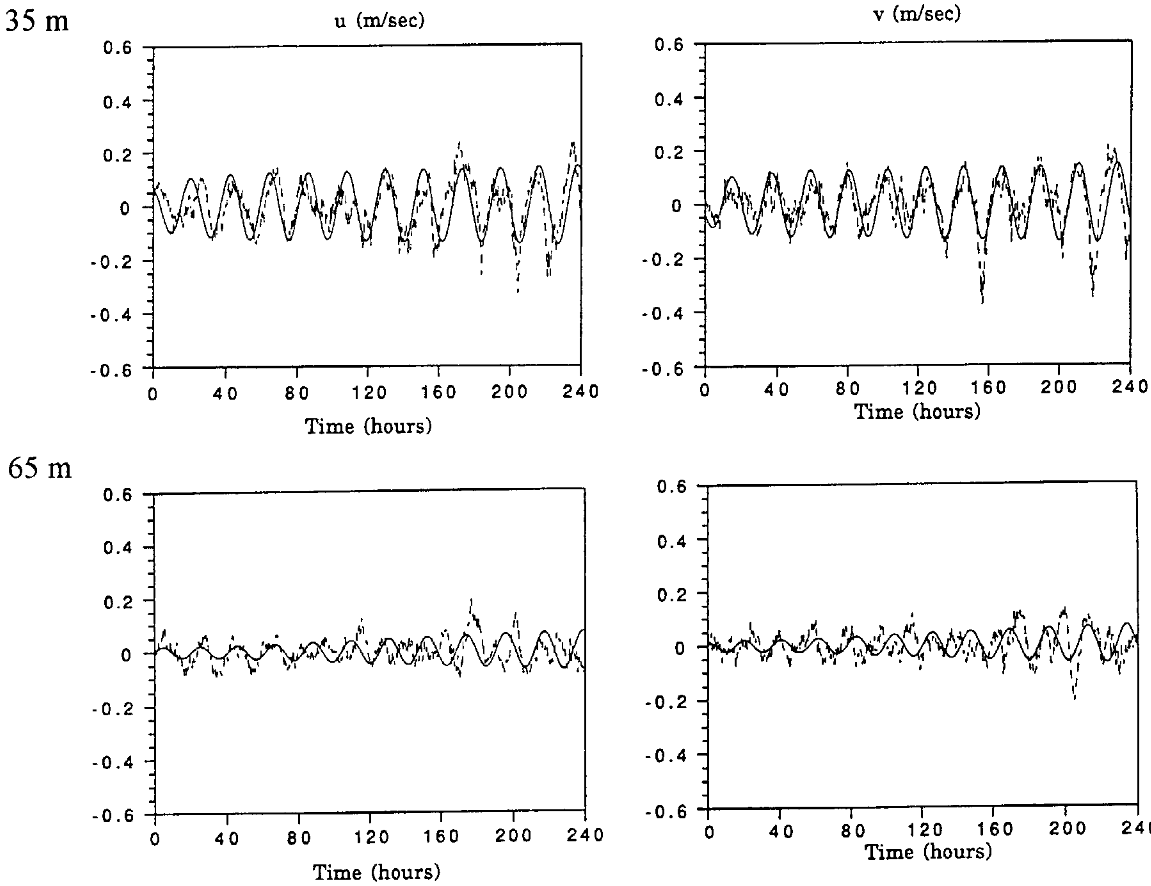

Figure 6. Comparison of modelled (solid lines) and observed (dashed lines) current speeds for u (left) and v (right) at, 35 and 65 m.

Acknowledgments

This work has been supported by NASA Oceanic Processes. Additional support from ONR is important. The data were supplied by Bob Weller, WHOI. In addition, Mel Briscoe and Jim Price were very helpful. The calculations were done by Lisan Yu. Rita Kuyper did a great job on the typing and everything else.

REFERENCES

Yu, L. and J. J. O'Brien, 1991: Variational Estimation of the Wind Stress Drag Coefficient and the Oceanic Eddy Viscosity Profile, *J. Phys. Oceanog.*, 21, No. 5, 709-719.

Yu, L. and J. J. O'Brien, 1992: On the Initial Condition in Parameter Estimation, *J. Phys. Oceanog.*, 22, 1361-1364.