

THE FLORIDA STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

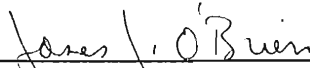
USING OCEAN CURRENT DATA TO
DETERMINE THE WIND STRESS

By
LISAN YU

A Thesis Submitted to the
Department of Oceanography in partial
fulfillment of the requirements for
the degree of Masters of Science

Degree Awarded:
Spring Semester, 1990

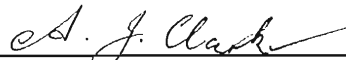
The members of the Committee approve the thesis of Lisan Yu defended on March 26, 1990.



James J. O'Brien
Professor Directing Thesis



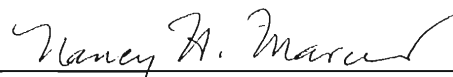
Steven L. Blumsack
Committee Member



Allan J. Clarke
Committee Member

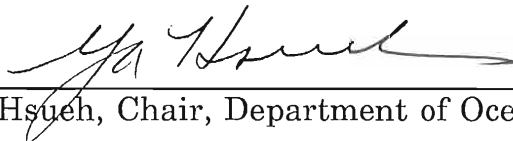


William K. Dewar
Committee Member

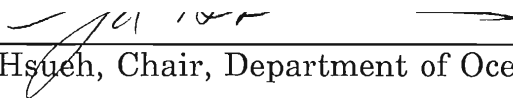


Nancy H. Marcus
Committee Member

Approved:



Ya Hsueh, Chair, Department of Oceanography



Ya Hsueh, Chair, Department of Oceanography

Abstract

A variational-adjoint analysis technique is used to assimilate both meteorological and oceanographical observations with the aid of a modified oceanic Ekman layer model. By fitting the model results to the data, the unknown boundary condition (wind stress) and the unknown vertical eddy viscosity distribution can be deduced simultaneously from the data, and an optimal estimate of the current field can be obtained as well.

The variational-adjoint analysis is conceptually simple. It provides a practical procedure for handling the data in a systematic, quantitative way. The wind stress and eddy viscosity are varied until a cost function, which measures the misfit between the model results and the observations, is minimized while the model equations are treated as strong constraints. The best fit is determined by a system which consists of a dynamical model and a corresponding adjoint model driven by the model-data misfit. The Lagrange multipliers calculated from the adjoint model transform the information about the data misfit into the gradient of the cost function, which then determines the amount that the wind stress and eddy viscosity should be changed. Thus, the misfit can be reduced and the cost function can be modified toward its optimal value.

can be reduced and the cost function can be modified toward its optimal value.

Four identical twin experiments are performed in which the "observations" are created by the dynamical model. The fast convergence and the excellent results promise that the variational-adjoint method is a feasible approach for data assimilation.

The field measurements from LOTUS-3 are also used in our study. The results show that, though the model is simple, the variational assimilation technique is capable of extracting useful information from the available observations. The wind stress and eddy viscosity estimated are all in their own reasonable ranges.

Acknowledgement

It is with the deepest gratitude that I thank Dr. James J. O'Brien as my major professor. His guidance, encouragement and support throughout the course of this research are sincerely appreciated. I wish to express my deep appreciation to Dr. I. M. Navon for his many valuable suggestions to this research. I am also grateful for the time taken by Drs. Steven L. Blumsack, Allan J. Clarke, William K. Dewar and Nancy H. Marcus while serving on my committee. Thanks are extended to Drs. M. G. Briscoe, J. F. Price and R. A. Weller from WHOI for kindly providing the LOTUS-3 data.

I would like to thank my colleagues at the Mesoscale Air-Sea Interaction Group for their comments and support. Special thanks go to Mr. Jiayan Yang for numerous discussions and constructive suggestions, to Mr. Brian Kelly for reviewing the manuscript and correcting an embarrassing number of errors, and to Mr. Allan Davis for technical help.

Finally, it is with the deepest affection that I thank my parents for introducing me to this wonderful world of physical oceanography and for their love and unconditional support.

Table of Contents

Abstract	iii
Acknowledgement	v
List of figures	viii
1. Introduction	1
2. Problem specification	6
2.1 Model dynamics	6
2.2 Variational analysis	9
2.3 Numerical method	15
3. The variational analysis technique applied to a simulated current field	24
3.1 Experiment 1: Wind is uniformly eastward, estimate only the eddy viscosity profile	26
3.2 Experiment 2: Wind is given as a sinusoid pattern and the eddy viscosity profile only is determined	30
3.3 Experiment 3: Using the identical twin data in experiment 2 to estimate the wind stress parameter and the eddy viscosity profile simultaneously	33
3.4 Experiment 4: The simulated current field is generated by a real wind, the wind forcing and the eddy viscosity by a real wind, the wind forcing and the eddy viscosity	

profile are simultaneously recovered	37
4. The variational analysis technique applied to an observed field	43
4.1 Data	43
4.2 Results	48
5. Summary and conclusion	61
References	64
Appendix Derivation of the finite difference equations	68

List of Figures

	Page
Figure 1. Diagram for the vertical structure of the numerical model.	16
Figure 2. The eddy viscosity profile used to create the "observations".	25
Figure 3. Results from experiment 1. (a) The variation of the cost function J/J_0 with the number of iterations (scaled by its initial value), (b) The variation of the norm of the gradient $ g / g_0 $ with the number of iterations (scaled by its initial value), (c) The variation of the data misfit $(w - \hat{w})^2$ (scaled by its initial value) with the number of iterations, and (d) The variation of the eddy viscosity profile during the iterative process (the number denotes the iteration).	28-29

Figure 4. Results from experiment 2. (a) The variation of the of the cost function J/J_0 with the number of iterations (scaled by its initial value), (b) The variation of the norm of the gradient $|g|/|g_0|$ with the number of iterations (scaled by its initial value), (c) The variation of the data misfit $(w - \hat{w})^2$ (scaled by its initial value) with the number of iterations, and (d) The variation of the eddy viscosity profile during the iterative process (the number denotes the iteration).

31-32

Figure 5. Results from experiment 3. (a) The variation of the of the cost function J/J_0 with the number of iterations (scaled by its initial value), (b) The variation of the norm of the gradient $|g|/|g_0|$ with the number of iterations (scaled by its initial value), (c) The variation of the data misfit $(w - \hat{w})^2$ (scaled by its initial value) with the number of iterations, (d) The variation of the eddy viscosity profile during the of the eddy viscosity profile during the

iterative process (the number denotes the iteration), and (e) The variation of the drag coefficient with the number of iterations. 34-36

Figure 6. Wind observations. 39

Figure 7. Results from experiment 4. (a) The variation of the of the cost function J/J_0 with the number of iterations (scaled by its initial value), (b) The variation of the norm of the gradient $|g|/|g_0|$ with the number of iterations (scaled by its initial value), (c) The variation of the data misfit $(w - \hat{w})^2$ (scaled by its initial value) with the number of iterations, (d) The variation of the eddy viscosity profile during the iterative process (the number denotes the iteration), and (e) The variation of the drag coefficient with the number of iterations. 40-42

Figure 8. Current observations at (a) 5m, (b) 25m, (c) 50m, and (d) 75m. 44-47

Figure 9. Real data assimilation. (a) The variation of the of the cost function J/J_0 with the number of iterations (scaled by its initial value), (b) The variation of the norm of the gradient $|g|/|g_0|$ with the number of iterations (scaled by its initial value), (c) The variation of the data misfit $(w - \hat{w})^2$ (scaled by its initial value) with the number of iterations, (d) The variation of the eddy viscosity profile during the iterative process (the number denotes the iteration), and (e) The variation of the drag coefficient with the number of iterations. 49-51

Figure 10. The variation of the correlation coefficient at different depths with the number of iterations. 54

Figure 11. Comparisons of the modelled (solid) and observed (dashed) current velocities at (a) 5m, (b) 15m, (c) 25m, (d) 35m, (e) 65m, and (f) 75m. 55-60

1. Introduction

An understanding of the direct inference of wind stress and turbulent mixing using meteorological and oceanographical datasets can help us to elucidate the dynamics of the upper ocean currents. Most existing numerical models, though they have become very sophisticated in recent years, normally prescribe a specified wind stress and turbulent parameterization, in order to simulate the ocean. Usually there is no direct information about the observed currents so these numerical models are not capable of using the observations to estimate the wind stress and eddy viscosity distribution.

The purpose of this paper is to demonstrate how the variational-adjoint technique can use a modified Ekman equation to simultaneously deduce the wind forcing and the eddy viscosity distribution from observed data. The variational-adjoint analysis technique, which derives from the optimal control theory (Lions,1971), provides a practical procedure for handling the observed data in a systematic, quantitative way. The basic idea is to define a cost function which is a sum of the squared discrepancies between model results and data and treat the dynamical model as a strong constraint (Sasaki,1970). The resulting functional is minimized by finding its gradient with respect to an unknown wind stress parameter and unknown eddy viscosity profile. Descent algorithms such as conjugate gradient or quasi-Newton are used to find the unknown wind stress parameter and unknown eddy viscosity profile. Descent algorithms such as conjugate gradient or quasi-Newton are

implemented to iterate towards the optimal state. The wind stress parameter and eddy viscosity are computed by fitting the model results to the data, and an optimal estimate of the current fields is obtained as well.

The problem of finding the minimum of the cost function with the dynamical model as a constraint is referred to as a constrained problem. A systematic approach for solving problems of this type is to define an Augmented Lagrange function by using undetermined Lagrange multipliers to enforce the model constraints. The introduction of Lagrange multipliers greatly simplifies the derivation of the gradient of the cost function. The minimum of the cost function coincides with a saddle point of the Augmented Lagrange function. The condition that a saddle point exists is equivalent to the condition that the gradients of the Augmented Lagrange function with respect to the model variables vanish. This yields the adjoint equations from which the Lagrange multipliers are calculated. The Lagrange multipliers transform the model-data misfit into the gradient of the cost function, which then provides the information to adjust wind forcing and eddy viscosity. Hence, the misfit can be reduced and model results can be modified.

Meteorologists have long concentrated on using the variational method for weather forecasting by adjusting some initial state of the model. Recently, Derber (1985) and Lewis and Derber (1985) have applied the variational-adjoint technique to multi-level quasi-geostrophic forecasting models. Their results promised the feasibility and usefulness the variational-adjoint technique to multi-level quasi-geostrophic forecasting models. Their results promised the feasibility and usefulness of this approach. Le Dimet and Talagrand (1986) have presented a

general formalism for the use of the variational-adjoint analysis in the context of data assimilation. The numerical experiments which were performed on a one-dimensional nonlinear equation showed satisfactory convergence of the optimization process. This technique was also utilized by Hoffman (1986), Talagrand and Courtier (1987), Courtier and Talagrand (1987), and Lorenc (1988) for four-dimensional data assimilation. For a review of the development of the variational and optimization methods in meteorology, see Lorenc (1986) and Le Dimet and Navon (1989).

The last few years have also seen more investigations in utilizing the variational method in oceanography. Thacker and Long (1988) have described the variational-adjoint method for a model fitting process in a simple equatorial wave model to calculate optimal initial conditions in order to fit simulated observations. Fitting a model to inadequate data is also discussed in Thacker (1988). Wunsch (1987, 1988) has analyzed the oceanic transient-tracers data by an optimal control method, and has demonstrated its advantages in the case of time-dependent data. Panchang and O'Brien (1988) have determined the bottom stress in a channel flow problem by the variational-adjoint equations approach. Smedstad (1989) has implemented an efficient conjugate-gradient algorithm to improve the convergence rate of the variational method. He used island sea level data to determine the basic stratification of the equatorial Pacific ocean. Tziperman and Thacker (1989) have presented adding error estimations to an adjoint formalism. This information has been shown to be very valuable for critically evaluating how well the data

determine the parameter's values.

Given the measurements of meteorological and oceanographic fields, the practical procedure of variational data assimilation starts by defining a cost function which measures the distance between model results and observations. The cost function is therefore a function of both the observations and the control parameters (here, wind stress and eddy viscosity). In the present study, given the initial guesses for the control parameters, the Ekman model is used to calculate the value of the cost function. The corresponding adjoint model is then used to calculate the gradient of the cost function with respect to the control parameters. Next, a conjugate gradient descent algorithm uses the gradient information to obtain a new estimate for the parameters, reducing the value of the cost function. Several iterations of the conjugate gradient descent algorithm are required to obtain the minimum value of the cost function, where model results and observations are as close as allowed by the level of measurements noise. The optimal estimate for the parameters, in a least-squares sense, is that corresponding to the minimum value of the cost function.

The methodology is given in Section 2, including the modified Ekman model description, the variational formalism and the numerical scheme. Section 3 presents the results obtained using the variational algorithm. The identical twin experiments are used in which the "observations" are created by the Ekman model. In Section 4, the solutions obtained by applying the variational technique to observed observations are created by the Ekman model. In Section 4, the solutions obtained by applying the variational technique to observed meteorological and oceanographic fields are discussed. The data used

are taken from the LOTUS-3 (Long-Term Upper Ocean Study-3) records. A summary of the results and the conclusions are given in Section 5.

The derivation of the finite difference version of the equations is included in an appendix.

2. Problem Specification

2.1 Model Dynamics

Consider a continuously stratified and horizontally unbounded ocean surface layer with depth H . Take the z -axis vertically upwards, with $z = 0$ at the surface, and $z = -H$ at the bottom. The ocean is rotating about the z -axis, and the Coriolis parameter, f , is taken to be constant. Neglecting the changes of the ocean surface, the modified Ekman model is

$$\frac{\partial w}{\partial t} + i f w = \frac{\partial}{\partial z} \left(A \frac{\partial w}{\partial z} \right) \quad (1.1)$$

where horizontal velocity components u and v (u positive to the east, v positive to the north) are combined into one complex vector $w = u + iv$; the eddy viscosity $A(z)$, which is the parameter to be calculated, is a function of depth.

Surface and bottom boundary conditions for this upper ocean are as follows:

At the surface, the forcing is given by the wind stress $\tau = \tau^x + i\tau^y$ (τ^x in the x -direction, τ^y in the y -direction), that is

$$\rho_w A \frac{\partial w}{\partial z} = \tau \quad \text{at } z = 0 \quad (1.2)$$

$$\rho_w A \frac{\partial w}{\partial z} = \tau \quad \text{at } z = 0 \quad (1.2)$$

where $\rho_w = 1.025 \times 10^3 \text{ kg/m}^3$ is the density of water.

The wind stress is calculated from

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

where $\rho_a = 1.2 \text{ kg/m}^3$ is the density of air; w_a is the complex vector of wind speed; c_D is the drag coefficient. Besides estimating the eddy viscosity $A(z)$, we estimate the drag coefficient, c_D , i.e., we also adjust the wind stress by fitting the current observations in an optimal way.

At some depth, we assume no momentum flux; that is

$$A \frac{\partial w}{\partial z} = 0 \quad \text{at } z = -H \quad (1.3)$$

H is not the bottom depth of the ocean surface layer but some depth of the computational domain chosen by the given ocean data.

The initial condition for this dynamic system is

$$w = w_0 \quad \text{at } t = 0 \quad (1.4)$$

We are going to determine two parameters, one is the wind stress, which enters in the upper boundary condition in the dynamic model; the

We are going to determine two parameters, one is the wind stress, which enters in the upper boundary condition in the dynamic model; the other one is the eddy viscosity profile, which represents a physical

property of the flow and varies with depth. For convenience, we use a nondimensional form of the model.

Since the motions in the upper ocean are dominated by the inertial oscillations, the inertial period $O(\frac{1}{f})$ is chosen as the time scale. We introduce the following nondimensional variables into the model and the boundary and initial conditions.

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

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

The typical values of eddy viscosity and drag coefficient are $s_a = 5 \times 10^{-2} \text{ m}^2/\text{sec}$ and $s_c = 1.2 \times 10^{-3}$ respectively.

The nondimensional problem takes the form (after dropping all primes)

$$\frac{\partial w}{\partial t} + i w = \frac{\partial}{\partial z} \left(A \frac{\partial w}{\partial z} \right) \quad (2.1)$$

with

with

$$A \frac{\partial w}{\partial z} = c_D | w_a | w_a \quad \text{for } z = 0 \quad (2.2)$$

$$A \frac{\partial w}{\partial z} = 0 \quad \text{for } z = -\frac{H}{D} \quad (2.3)$$

and

$$w = w_0 \quad \text{for } t = 0 \quad (2.4)$$

2.2 Variational Analysis

We now apply the variational method to this model problem. A cost function is constructed first which measures the misfit between the model results and the data. Our objective is to find the current field generated by the model such that the departure from the observations is minimized. Considering the linear dynamics of our model, we choose a least squares fitting for the cost function. The cost function is defined as

$$J(w, A, c_D) = \frac{1}{2} K_m \int_t \int_z (w - \hat{w})^2 d\zeta d\tau + \frac{1}{2} K_a T \int_z (A - \hat{A})^2 d\zeta + \frac{1}{2} K_c T H (c_D - \hat{c}_D)^2 \quad (3)$$

where the carrot denotes the observations or estimates; the coefficients K_m , K_a and K_c are the Gauss precision moduli controlling the best fits where the carrot denotes the observations or estimates, the coefficients K_m , K_a and K_c are the Gauss precision moduli controlling the best fits

for each type of data; The nondimensional parameters T and H represent the total integration time and depth in the model, respectively.

The cost function (3) is composed of three terms. The first one is called the data misfit which is the squared difference between the model solution and the observations. The last two terms measure the closeness of the estimated parameters to the previous guess. Since parameter estimation is our main interest in this paper, the added terms represent prior information about the parameters which increase the chance that the cost function will be convex and therefore lead to a unique solution. Thus the model solution resulting from minimizing the cost function will best agree with the observed data and the new estimate of the parameters will not deviate far from the previous values. In this sense, the parameters' initial guess should be as reasonable as possible so that the optimization process can perform efficiently.

The minimum of the cost function generally is not expected to be equal to zero, because the observations have errors, and the model is ideal and, thus its solution is not exactly compatible with the observations. In the particular case, where a simulated field is used, the exact consistency between the "observations" and the model dynamics will make the optimal value of the cost function vanish. The model solution is expected to completely satisfy the characteristics of the "observed" field in these test runs called identical twin experiments.

Our next task is to find the correct parameters from the model equation that optimizes the cost function. The dynamic system described

Our next task is to find the correct parameters from the model equation that optimizes the cost function. The dynamic system described by the linear Ekman equation, relates the control parameters, i.e., the

wind forcing term, to the current field of the system. This defines one set of constraints. The problem of directly inferring the wind stress from the measurements of ocean currents is identified as the mathematical problem of selecting the external forcing so that the resulting model solution satisfies the functional constraints and minimizes the cost function. It is a standard procedure to introduce the Augmented Lagrange function associated with this constrained problem, given as

$$L(w, A, c_D, \lambda) = J + \int_t \int_z \{ \lambda, (\frac{\partial w}{\partial t} + i f w - \frac{\partial}{\partial z}(A \frac{\partial w}{\partial z})) \} d\zeta d\tau \quad (4)$$

where $\{ , \}$ is the inner product of two vectors; $\lambda = \lambda_u + i\lambda_v$ is the complex vector of the undefined Lagrange multipliers. (λ_u for the u-component model equation, λ_v for the v-component). The model equation represents a strong constraint according to Sasaki (1970), that needs to be satisfied exactly. The constrained optimization problem is now replaced by a series of unconstrained problem with respect to the variables w , A , c_D and λ . But not all model variables are independent. The Augmented Lagrange function allows the independent variables A and c_D to vary while the subsequent variations of the dependent current field are given by the model equations. The boundary condition does not appear as a constraint with its own Lagrange multiplier because it enters into the model equation through the forcing term in the finite-difference formalism.

The problem of minimizing the cost function subject to the model equation becomes a problem of finding the stationary points of the Augmented Lagrange function corresponding to all variables w , A , c_D and λ . This is equivalent to the determination of w , A , c_D and λ under the condition that the gradient of the Augmented Lagrange function vanishes, which yields the following set of equations:

$$\frac{\partial L(w, A, c_D, \lambda)}{\partial \lambda} = 0 \quad (5)$$

$$\frac{\partial L(w, A, c_D, \lambda)}{\partial w} = 0 \quad (6)$$

$$\frac{\partial L(w, A, c_D, \lambda)}{\partial A} = 0 \quad (7)$$

$$\frac{\partial L(w, A, c_D, \lambda)}{\partial c_D} = 0 \quad (8)$$

Equation (5) recovers the original model equation, while (6) results in the adjoint equation, given by

$$\frac{\partial \lambda}{\partial t} + i\lambda + \frac{\partial}{\partial z} \left(A \frac{\partial \lambda}{\partial z} \right) = K_m (w - \hat{w}) \quad (9)$$

Note that we have used the following natural boundary and initial

conditions when deriving the adjoint equation: the net flux of the adjoint variable λ is required to be zero both at the top and at the bottom of the chosen computational domain; that is

$$\frac{\partial \lambda}{\partial z} = 0 \quad \text{at } z = 0 \text{ and } z = -\frac{H}{D}$$

the initial condition is

$$\lambda = 0 \quad \text{at } t = T$$

Comparing with the original model equation, it is worth noting that the adjoint equation has a similar form to the model equation, except for two important features. The friction term in the adjoint equation has the opposite sign than in the model equation. The stability of the well-posed problem thus requires the integration of the adjoint equation to be backward in time. In addition, the driving factor for the adjoint equation is the square root of the data misfit. The Lagrange multipliers carry the information about the data back to the initial time to influence the reconstruction of the model state.

Through (7) and (8), the gradient of the Augmented Lagrange function L with respect to A and c_D yields the equations

$$c_D = \hat{c}_D + \frac{1}{K_c T H} \int_t (|w_a| u_a \lambda_{u z=0} + |w_a| v_a \lambda_{v z=0}) d\tau \quad (10)$$

$$c_D = \hat{c}_D + \frac{1}{K_c T H} \int_t (|w_a| u_a \lambda_{u z=0} + |w_a| v_a \lambda_{v z=0}) d\tau \quad (10)$$

$$A = \hat{A} + \frac{1}{K_a T} \int_t (\frac{\partial u}{\partial z} \frac{\partial \lambda_u}{\partial z} + \frac{\partial v}{\partial z} \frac{\partial \lambda_v}{\partial z}) d\tau \quad (11)$$

These are the equations for the control parameters.

Equations (9) - (11) can be simplified if we define

$$\frac{\lambda}{K_m} = \lambda', \quad \frac{K_c}{K_m} = K_c', \quad \frac{K_a}{K_m} = K_a'$$

the three Gauss precision moduli K_c , K_a and K_m are replaced by K_c' and K_a' . In general, if one has n precision moduli, the number can be reduced to $n-1$ by scaling. Rewriting equations (9) - (11) (dropping all the primes)

$$\frac{\partial \lambda}{\partial t} + i\lambda + \frac{\partial}{\partial z} (A \frac{\partial \lambda}{\partial z}) = (w - \hat{w}) \quad (9')$$

$$c_D = \hat{c}_D + \frac{1}{K_c T H} \int_t (|w_a| u_a \lambda_{u z=0} + |w_a| v_a \lambda_{v z=0}) d\tau \quad (10')$$

$$A = \hat{A} + \frac{1}{K_a T} \int_t (\frac{\partial u}{\partial z} \frac{\partial \lambda_u}{\partial z} + \frac{\partial v}{\partial z} \frac{\partial \lambda_v}{\partial z}) d\tau \quad (11')$$

~ t

There are six unknowns $u(z,t)$, $v(z,t)$, $A(z)$, c_D , $\lambda_u(z,t)$ and $\lambda_v(z,t)$, and six equations (2), (9'), (10') and (11'), so the system is closed. The parameters $\hat{A}(z)$, \hat{c}_D , T , H , K_a and K_c must be specified. The numerical scheme used to solve this optimal control problem will be described in the following subsection.

2.3 Numerical Method

In solving equations (2) and (9') - (11') the numerical model is formulated using a finite difference discretization on a grid with spatial increment Δz and temporal increment Δt . Its vertical structure is schematically shown in Figure 1. The w points are staggered in space with the A points. The equations are integrated in time using an implicit scheme (O'Brien, 1986).

The finite-difference form of (2) is

$$\begin{aligned} \frac{w_j^{n+1} - w_j^n}{\Delta t} + i \frac{w_j^{n+1} + w_j^n}{2} - \frac{1}{2\Delta z} (A_j \frac{w_{j-1}^n - w_j^n}{\Delta z} + A_j \frac{w_{j+1}^{n+1} - w_j^{n+1}}{\Delta z} \\ - A_{j+1} \frac{w_j^n - w_{j+1}^n}{\Delta z} - A_{j+1} \frac{w_j^{n+1} - w_{j+1}^{n+1}}{\Delta z}) = 0 \end{aligned} \quad (12.1)$$

for all $j = 1, 2, \dots, J$ and $n = 0, 1, \dots, N - 1$.

for all $j = 1, 2, \dots, J$ and $n = 0, 1, \dots, N - 1$.

with boundary conditions

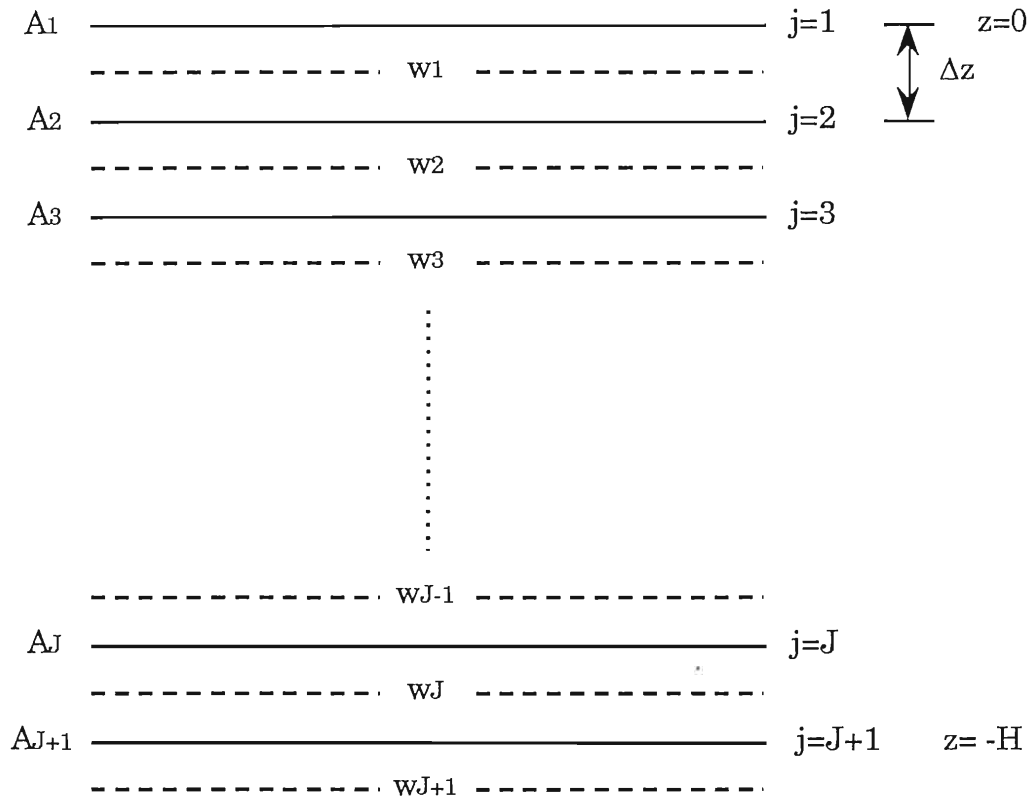


Figure 1 Diagram for the vertical structure of the numerical model

$$A_1 \frac{w_0^n - w_1^n}{\Delta z} = c_D |w_a^n| w_a^n \quad (12.2)$$

$$A_{J+1} \frac{w_J^n - w_{J+1}^n}{\Delta z} = 0 \quad (12.3)$$

for $n = 1, 2, \dots, N - 1$.

and initial condition

$$w_j^0 = w_{0j} \quad (12.4)$$

for $j = 1, 2, \dots, J$

where w_j^n is the complex vector of the current velocity at grid $z = -j\Delta z$ and $t = n\Delta t$.

Rearranging the terms in equation (12.1) and applying the boundary condition properly to the equation, the evolution of the new state w over the time step Δt is given by

$$\begin{aligned} -A_j' w_{j-1}^{n+1} + (1 + A_j' + A_{j+1}') w_j^{n+1} - A_{j+1}' w_{j+1}^{n+1} + i\alpha w_j^{n+1} \\ = A_j' w_{j-1}^n + (1 - A_j' - A_{j+1}') w_j^n + A_{j+1}' w_{j+1}^n - i\alpha w_j^n \end{aligned} \quad (13.1)$$

For $j = 1$, the equation takes the form

For $j = 1$, the equation takes the form

$$\begin{aligned}
& (1 + A_2') w_1^{n+1} - A_2' w_2^{n+1} + i\alpha w_1^{n+1} \\
& = (1 - A_2') w_1^n + A_2' w_2^n - i\alpha w_1^n + c_D' (|w_a^n| w_a^n + |w_a^{n+1}| w_a^{n+1})
\end{aligned}
\tag{13.2}$$

For $j = J$, the equation is

$$\begin{aligned}
& -A_J' w_{J-1}^{n+1} + (1 + A_J') w_J^{n+1} + i\alpha w_J^{n+1} \\
& = A_J' w_{J-1}^n + (1 - A_J') w_J^n - i\alpha w_J^n
\end{aligned}
\tag{13.3}$$

where $A_j' = \frac{\Delta t}{2\Delta z^2} A_j$, for $j = 1, 2, \dots, J$; $\alpha = \frac{\Delta t}{2}$ and $c_D' = \frac{\Delta t}{2\Delta z} c_D$.

Separating equation (13) into two components u and v and combining them to form the matrix

$$B_m q^{n+1} = C_m q^n + F^n \tag{14}$$

where the column vector q is comprised of all the u and v components; the column vector F represents the forcing term and the $2J \times 2J$ matrices B_m and C_m are constant with time. Their respective forms are as follows

m m

$$q^n = \begin{bmatrix} u_1^n \\ u_2^n \\ \vdots \\ \vdots \\ u_J^n \\ v_1^n \\ v_2^n \\ \vdots \\ \vdots \\ v_J^n \end{bmatrix} \quad F^n = \begin{bmatrix} c_D(|w_a^n|u_a^n + |w_a^{n+1}|u_a^{n+1}) \\ 0 \\ \vdots \\ \vdots \\ 0 \\ c_D(|w_a^n|v_a^n + |w_a^{n+1}|v_a^{n+1}) \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$$

$$B_m = \begin{bmatrix} \mathfrak{R}_+ & -\alpha I \\ \alpha I & \mathfrak{R}_+ \end{bmatrix}$$

$$C_m = \begin{bmatrix} \mathfrak{R}_- & \alpha I \\ -\alpha I & \mathfrak{R}_- \end{bmatrix}$$

where I is an identity matrix. Here \mathfrak{R}_+ in B_m and \mathfrak{R}_- in C_m are given by

$$\mathfrak{R}_+ = \begin{bmatrix} 1+A_2' & -A_2' & 0 & \dots & \dots & 0 \\ -A_2' & 1+A_2'+A_3' & -A_3' & 0 & & \\ 0 & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & -A_j' & 1+A_j'+A_{j+1}' & -A_{j+1}' & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots \\ & & & & & & 0 \\ & & & & & 0 & -A_{J-1}' & 1+A_{J-1}'+A_J' & -A_J' \\ 0 & \dots & \dots & & & 0 & -A_J' & 1+A_J' \\ 0 & \dots & \dots & & & 0 & -A_{J-1}' & 1+A_{J-1}'+A_J' & -A_J' \\ & & & & & 0 & -A_J' & 1+A_J' \end{bmatrix}$$

$$\mathfrak{K}_- = \begin{bmatrix} 1-A_2' & A_2' & 0 & \dots & \dots & 0 \\ A_2' & 1-A_2'-A_3' & A_3' & 0 & & \\ 0 & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & A_j' & 1-A_j'-A_{j+1}' & A_{j+1}' & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots \\ 0 & & \dots & \dots & & A_{J-1}' & 1-A_{J-1}'-A_J' & A_J' \\ & & & & & 0 & A_J' & 1-A_J' \end{bmatrix}$$

Equation (14) is a linear dynamic model with the state vector q , the forcing vector F and the dynamical operators B_m and C_m . It is clear that through the finite-difference form the boundary condition is connected to the current field w_j^n at any point j and n . Therefore, the forcing term serves as an updating of the current field and its own value is updated at each iteration also.

The cost function (3) in the discretized version can be written as

$$\begin{aligned} J(w_j^n, A_j, c_D) &= \frac{1}{2} K_m \sum_{n=1}^N \sum_{j=1}^J (w_j^n - \hat{w}_j^n)^2 \Delta z \Delta t + \\ &+ \frac{1}{2} K_a T \sum_{j=1}^J (A_j - \hat{A}_j)^2 \Delta z + \frac{1}{2} K_c TH (c_D - \hat{c}_D)^2 \quad (15) \\ &+ \frac{1}{2} K_a T \sum_{j=1}^J (A_j - \ddot{A}_j)^2 \Delta z + \frac{1}{2} K_c TH (c_D - \hat{c}_D)^2 \quad (15) \end{aligned}$$

Applying the finite-difference scheme to the adjoint model (9) and following the same procedure as for the dynamical model, the adjoint model is transformed into the matrix form

$$B_\lambda \Lambda^n = C_\lambda \Lambda^{n+1} - D^{n+1} \quad (16)$$

where the column vector Λ consists of all the Lagrange multipliers λ_u and λ_v ; the column vector D stands for the data misfit; the $2J \times 2J$ matrices B_λ and C_λ are the transpose of B_m and C_m . They can be expressed as:

$$\Lambda^n = \begin{bmatrix} \lambda_{u1}^n \\ \lambda_{u2}^n \\ \vdots \\ \lambda_{uJ}^n \\ \lambda_{v1}^n \\ \lambda_{v2}^n \\ \vdots \\ \lambda_{vJ}^n \end{bmatrix} \quad D^n = \begin{bmatrix} \Delta t(u_1^n - \hat{u}_1^n) \\ \Delta t(u_2^n - \hat{u}_2^n) \\ \vdots \\ \Delta t(u_J^n - \hat{u}_J^n) \\ \Delta t(v_1^n - \hat{v}_1^n) \\ \Delta t(v_2^n - \hat{v}_2^n) \\ \vdots \\ \Delta t(v_J^n - \hat{v}_J^n) \end{bmatrix}$$

$$B_\lambda = \begin{bmatrix} \mathfrak{R}_+ & \alpha I \\ -\alpha I & \mathfrak{R}_+ \end{bmatrix} \quad C_\lambda = \begin{bmatrix} \mathfrak{R}_- & -\alpha I \\ \alpha I & \mathfrak{R}_- \end{bmatrix}$$

The discretized form of the equations for calculating the control parameters, A and c_D , are given in (17) and (18)

The discretized form of the equations for calculating the control parameters, A and c_D , are given in (17) and (18)

$$c_D = \hat{c}_D - \frac{1}{K_c T_H} \sum_{n=1}^{N-1} -\frac{1}{2} \{ (\lambda_{u1}^n | w_a^n | u_a^n + | w_a^{n+1} | u_a^{n+1}) \\ + (\lambda_{v1}^n | w_a^n | v_a^n + | w_a^{n+1} | v_a^{n+1}) \} \Delta t \quad (17)$$

$$A_j = \hat{A}_j - \frac{1}{K_a T} \frac{1}{2\Delta z} \sum_{n=1}^{N-1} \frac{1}{2} \{ (\lambda_{uj-1}^n - \lambda_{uj}^n) ((u_{j-1}^n - u_j^n) + (u_{j-1}^{n+1} - u_j^{n+1})) \\ + (\lambda_{vj-1}^n - \lambda_{vj}^n) ((v_{j-1}^n - v_j^n) + (v_{j-1}^{n+1} - v_j^{n+1})) \} \Delta t \quad (18)$$

With the equations (14) - (18), our procedure for using the variational method to solve this system can be fully described as follows:

- (1) begin with a best initial estimate for the control parameters A and c_D .
- (2) integrate the model equation (14) forward in time and calculate the value of the cost function.
- (3) compute the data misfits $(w - \hat{w})$.
- (4) integrate the adjoint equation (16) backward in time.
- (5) use equations (17) and (18) to calculate the components of ∇J (the gradients of the cost function) corresponding to A and c_D with solutions for λ and w from steps (2) and (4).
- (6) with the gradient information, apply the unconstrained minimization descent algorithm to obtain the new values of A
- (6) with the gradient information, apply the unconstrained minimization descent algorithm to obtain the new values of A

and c_D simultaneously.

(7) check if the convergence criterion

$$|\nabla J|/|\nabla J_0| < 10^{-2}$$

for the minimization process is satisfied, where ∇J_0 is the value at the initial iteration.

(8) return to step (2) if the optimal solution is not found.

This minimization determines the best fit of the data when the optimal solution is approached. Many different minimization methods are available (Gill et al., 1981). The method we used is the limited-memory quasi-Newton conjugate-gradient method, which is implemented in the Shanno and Phua's (1980) CONMIN algorithm. The conjugate gradient algorithm is superior to the steepest descent algorithm, as has been shown in many papers. When dealing with the well-conditioned problem, the conjugate gradient method provides fast functional reduction within the first few iterations. For linear dynamics, its convergence should be achieved in at most M iterations, where M is the number of the control variables. In fact, the rate of convergence depends to a large extent on the quality of the observations. Noisy observations poorly reflect the model dynamics, so the conjugate gradient method will converge slowly and all M iterations will be needed to obtain the required accuracy.

3. The Variational Analysis Technique Applied to a Simulated Current Field

In this section, four identical twin experiments are going to be discussed to demonstrate the feasibility of utilizing the variational analysis to estimate the wind forcing, as well as, the eddy viscosity distribution.

The model parameters in the dimensional form are chosen as

$$\Delta z = 10 \text{ m}$$

$$\Delta t = 30 \text{ minutes}$$

$$f = 10^{-4} \text{ s}^{-1}$$

$$\text{total model integration time } T = 10 \text{ days}$$

$$\text{data extension depth } H = -100 \text{ m}$$

The scaling parameters D and U are calculated from (2) by using these values

$$\rho_a = 1.2 \text{ kg/m}^3$$

$$\rho_w = 1.025 \times 10^3 \text{ kg/m}^3$$

$$s_a = 5.0 \times 10^{-2} \text{ m}^2/\text{sec}$$

$$s_a = 5.0 \times 10^{-2} \text{ m}^2/\text{sec}$$

$$s_c = 1.2 \times 10^{-3}$$

$$U_a = 10 \text{ m/sec}$$

The identical twin experiments obtain "observations" from the forward run of the original dynamical model. Thus, the data produced correspond exactly to the model dynamics. Figure 2 shows the eddy viscosity profile used in the model integration to create the "observations". The initial state of the system is at rest. Three sets of wind data are employed to simulate the current field, a constant westerly wind, a sinusoid wind and a random wind observation.

In order to test the variational algorithm, experiments 1 and 2 start by estimating only the eddy viscosity (assuming that the wind forcing is known exactly). The third experiment calculates the wind stress and the unknown eddy viscosity distribution simultaneously. This is a more complicated but more realistic scenario. Finally, we make use of the wind observation to recover the wind stress along with the viscosity from the identical twin data by the optimization technique.

3.1 *Experiment 1: Wind is uniformly eastward, estimate only the eddy viscosity profile*

In this experiment, the wind stress is assumed to be known exactly. The identical twin data is generated by a constant westerly wind with a magnitude of 10 m/sec. The unscaled drag coefficient c_D is set to exactly. The identical twin data is generated by a constant westerly wind with a magnitude of 10 m/sec. The unscaled drag coefficient c_D is set to

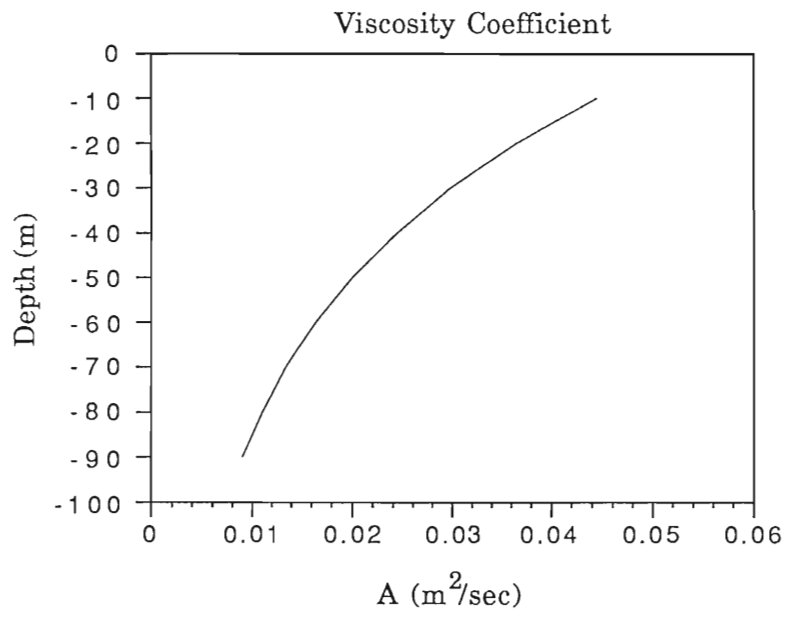
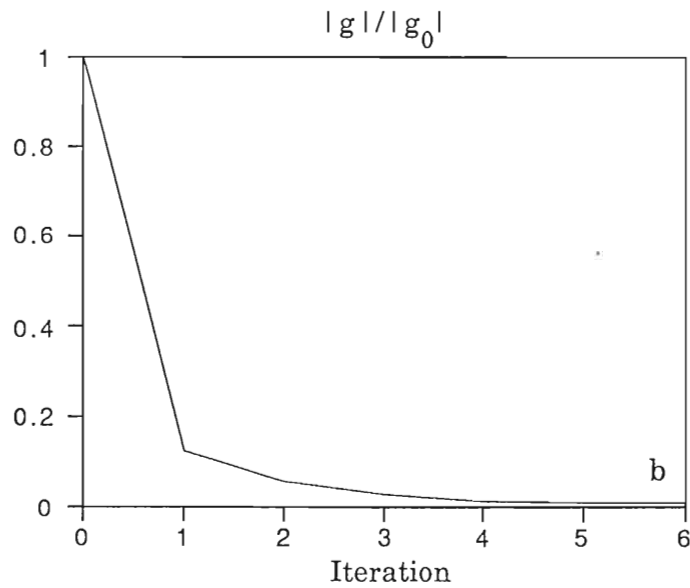
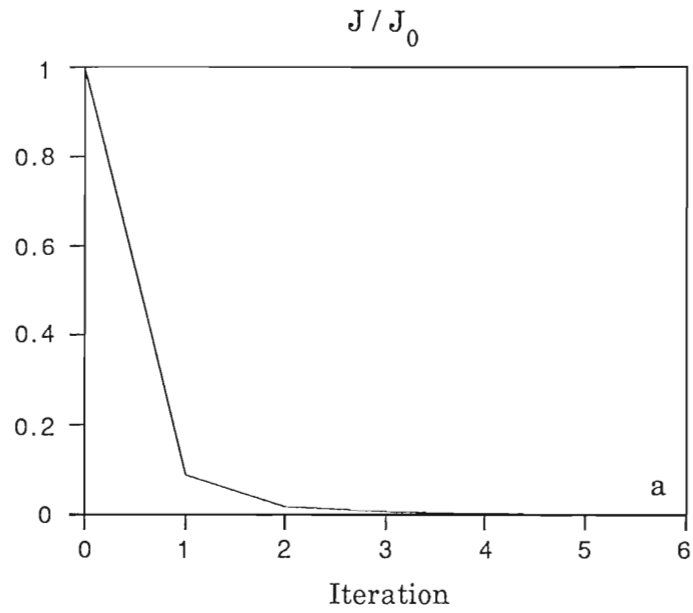


Figure 2 The eddy viscosity profile used to create the "observations"

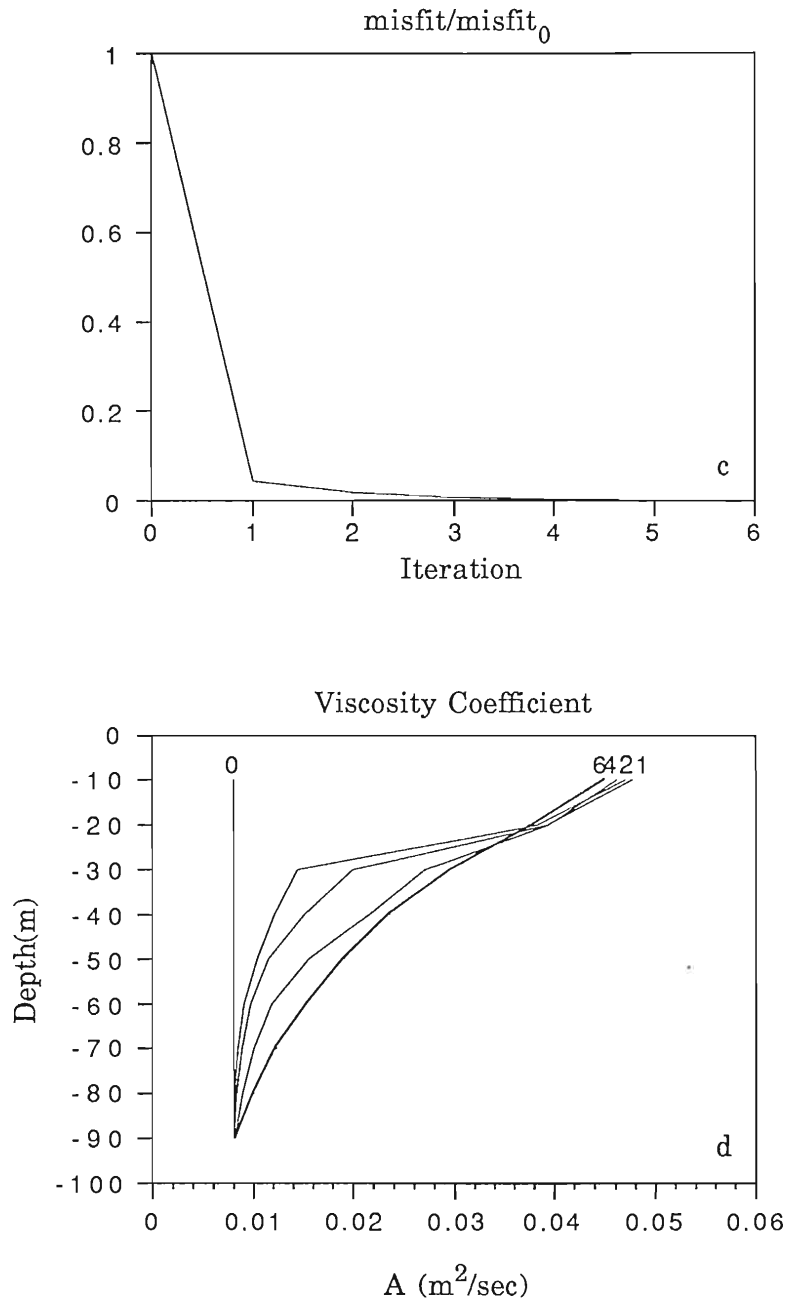
1.2×10^{-3} .

The data assimilation follows the procedure described in section 2.3. The iterative process begins with the constant initial guess of $0.8 \times 10^{-2} \text{ m}^2/\text{sec}$ for the eddy viscosity. Figures 3a-3c show the values of the cost function, its gradient and the data misfit ($w - \hat{w}$) versus the number of iterations in the minimization procedure. All the values have been normalized by their own initial values to allow a direct comparison of the convergence rate. As can be seen in Figure 3a, the normalized cost function drops dramatically in the first couple of iterations. After 4 iterations, J/J_0 has been reduced to 10^{-3} . The norm of the gradient (Figure 3b) also experiences a sharp decrease in these two iterations. $|g|/|g_0|$ ($|g|$ represents $|\nabla J|$ thereafter) reaches the convergence criterion after 6 iterations. Figure 3d displays the variation of the eddy viscosity profile during the iterative process. The bold line is the converged solution. It shows a good approximation to its true value (see Figure 2). The value of A in the upper 30 meters experiences a strong correction in the first iteration. This is because the chosen initial estimation for the eddy viscosity is relatively small compared with its true value. The current field generated with this value produces a large data misfit. The Lagrange multipliers computed from the adjoint model, are being driven by this big data misfit, and, therefore, have a great effect on the calculation of the gradient of the cost function. Hence, a better estimation is obtained. Physically, for a given forcing, the value great effect on the calculation of the gradient of the cost function. Hence, a better estimation is obtained. Physically, for a given forcing, the value of the eddy viscosity has to be large enough to transfer the input



Figures 3a-3b Results from experiment 1. The variation of (a) the cost function, (b) the norm of the gradient with the number of iterations (scaled by their own initial values)

function, (b) the norm of the gradient with the number of iterations (scaled by their own initial values)



Figures 3c-3d Results from experiment 1. (c) The variation of the data misfit ($w - \hat{w}$) with the number of iterations (scaled by its initial value), and (d) The variation of the eddy viscosity profile during the iterative process (the number denotes the iteration)

momentum downward. The profile changes slowly for the rest of the iterative process. Since the data misfit is very small after the first iteration, this indicates that the Lagrange multiplier propagates with very little information about the model data misfit thereafter, so little correction is given to the eddy viscosity profile.

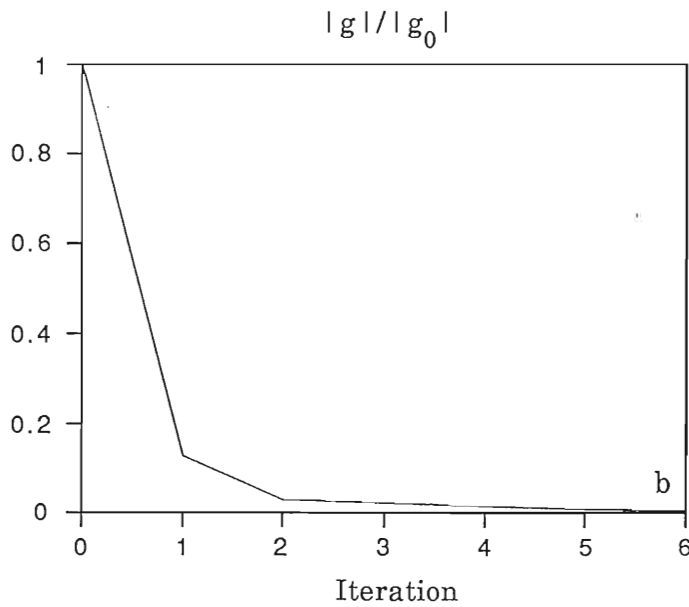
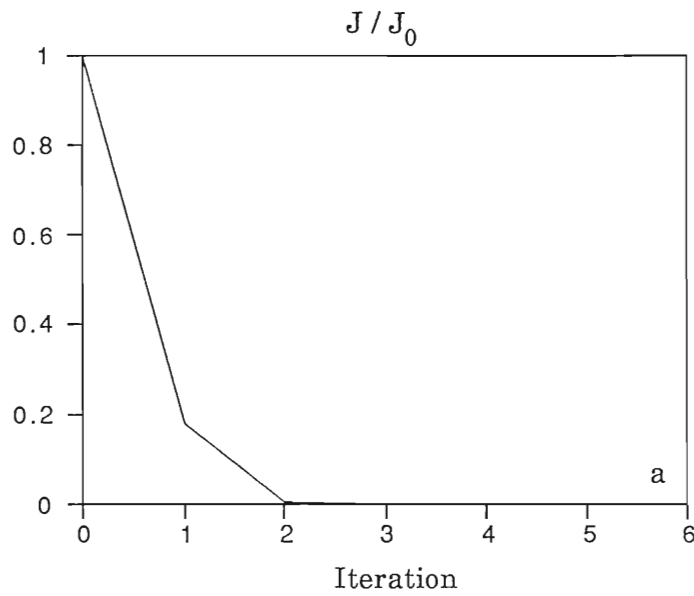
3.2 *Experiment 2: Wind is given as a sinusoid pattern and the eddy viscosity profile only is determined*

Assume again that the forcing is known and only the eddy viscosity distribution is needed to be determined. The simulated observation is obtained by running the linear dynamical model forward, using a sinusoid wind pattern $w = 10 \sin\left(\frac{2\pi t}{T_0}\right)$ (m/sec), where $T_0 = 10$ hours.

The convergence of the cost function to the optimal value is illustrated in Figure 4a. Its gradient and data misfit are shown in Figures 4b-4c. It is obvious that the variational algorithm quickly improves the estimate for the eddy viscosity in the first few iterations. The data misfit is corrected to about 0.01 in only one iteration. Within 6 iterations, the cost function is at its minimum. Correspondingly, the profile of the eddy viscosity in Figure 4d shows rapid approach to its true solution.

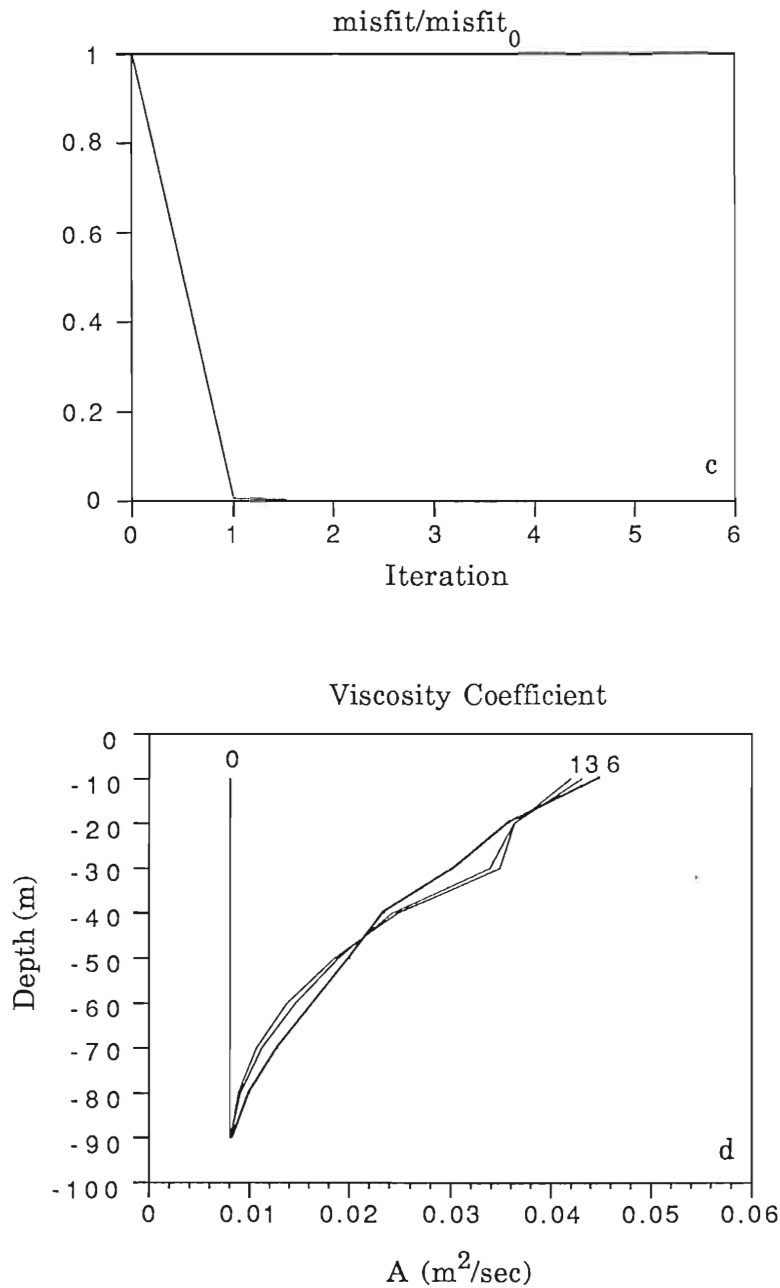
It is important to note that the choice of the initial value for the unknown parameter is arbitrary, however, the solution is independent

It is important to note that the choice of the initial value for the unknown parameter is arbitrary, however, the solution is independent of the initial guess. The fast convergence rate in these two experiments



Figures 4a-4b Results from experiment 2. The variation of (a) the cost function, (b) the norm of the gradient with the number of iterations

Figures 4a-4b Results from experiment 2. The variation of (a) the cost function, (b) the norm of the gradient with the number of iterations



Figures 4c-4d Results from experiment 2. (c) The variation of the data misfit with the number of iterations, and (d) The variation of the eddy viscosity profile during the iterative process

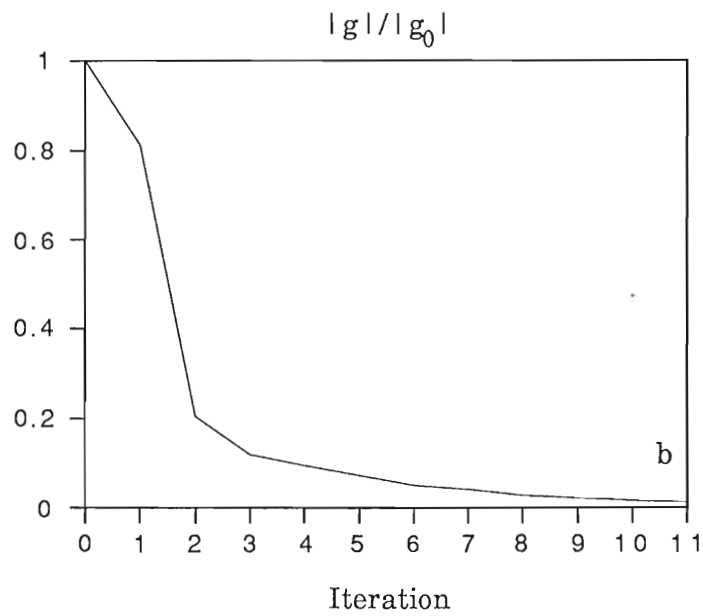
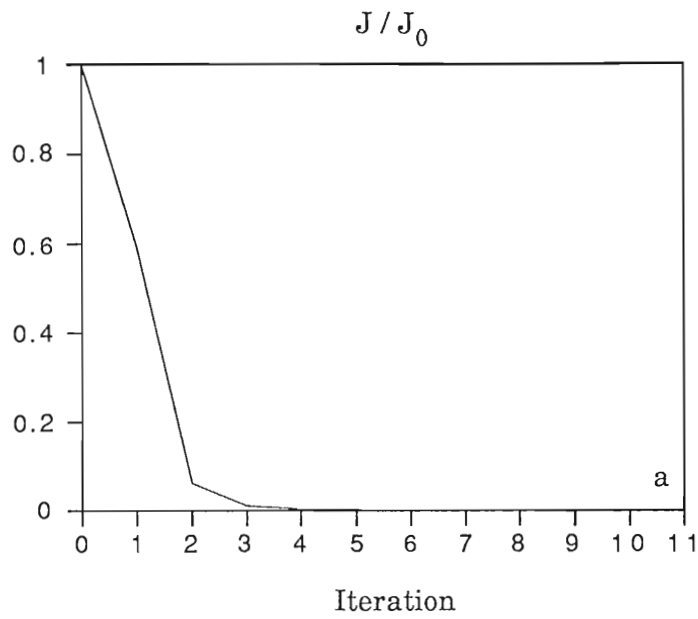
is very encouraging. It shows that the algorithm works successfully at least in variationally performing the data assimilation by adjusting only the eddy viscosity.

3.3 *Experiment 3: Using the identical twin data in experiment 2 to estimate the wind stress parameter and the eddy viscosity profile simultaneously*

This experiment begins to treat both the wind parameter and the eddy viscosity as unknowns and to recover them simultaneously by the optimization. We use the identical twin data from experiment 2 to be the "observations", which is created by a sinusoid wind.

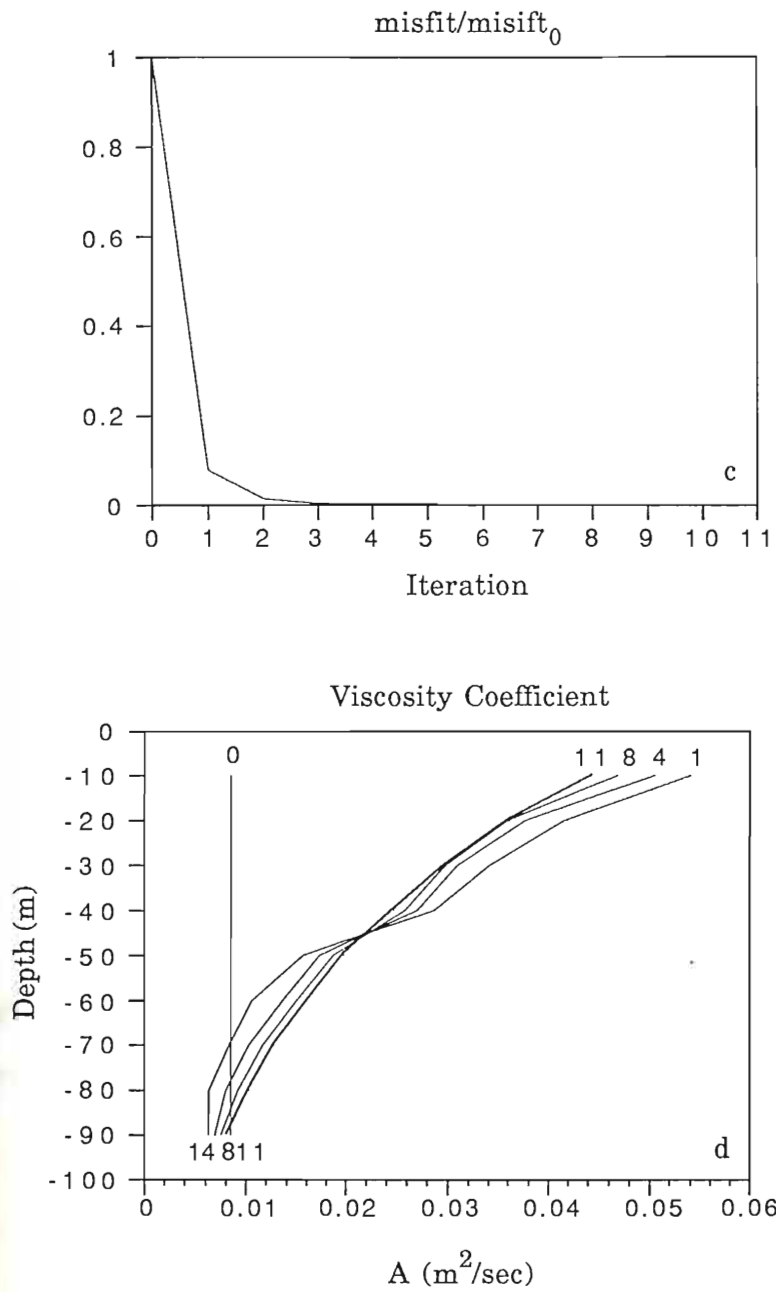
Relating the unknown boundary condition to the current observations is advantageous to the data assimilation. The simultaneous improvement of the prior estimate of the wind stress along with the eddy viscosity allows a flexible and efficient approach in combining data with the observations.

The results of the iterative process are shown in Figures 5a-5e. The algorithm attains its convergence criterion after 11 iterations (Figure 5b). As before, the cost function decreases greatly during the first few iterations (Figure 5a). It is interesting to see the plots in Figures 5d-5e which show the evolution of the eddy viscosity distribution and the drag coefficient during the optimization. The initial estimates for A and c_D are given as $0.8 \times 10^{-2} \text{ m}^2/\text{sec}$ and 0.72×10^{-3} (unscaled), and the drag coefficient during the optimization. The initial estimates for A and c_D are given as $0.8 \times 10^{-2} \text{ m}^2/\text{sec}$ and 0.72×10^{-3} (unscaled), respectively. These values are smaller than their true values. The



Figures 5a-5b Results from experiment 3. The variation of (a) the cost function, (b) the norm of the gradient with the number of iterations

function, (b) the norm of the gradient with the number of iterations



Figures 5c-5d Results from experiment 3. (c) The variation of the data misfit with the number of iterations, and (d) The variation of the eddy viscosity profile during the iterative process

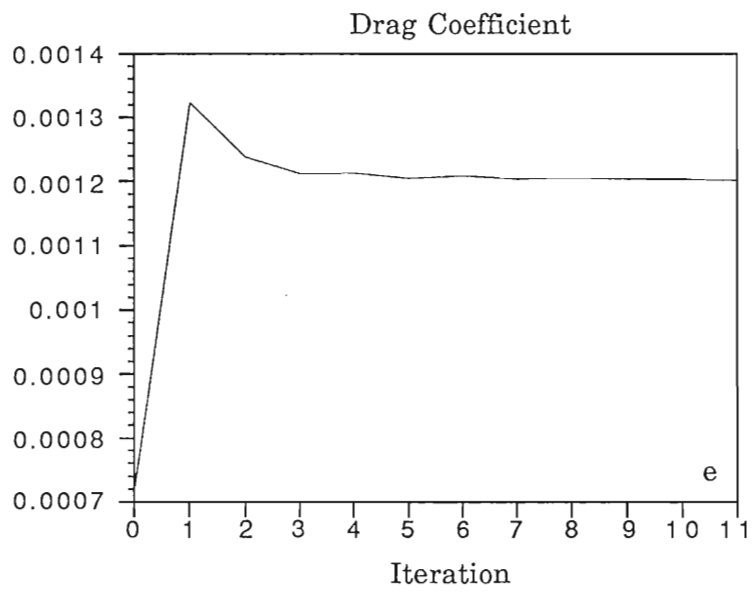


Figure 5e Results from experiment 3. The variation of drag coefficient with the number of iterations

Lagrange multiplier containing the information about this large data misfit makes a big correction to the previous estimates. Figure 5e shows there is a sharp increase for the drag coefficient during the first iteration, but it overshoots. Corresponding to this strong forcing, the eddy viscosity in about upper 50 meters has a relatively large value in comparison with its true solution. The drag coefficient tends towards its correct value after the first iteration and the eddy viscosity is also adjusted gradually. At iteration 11, the profile of the eddy viscosity looks identical to the one used to create the "observations" (Figure 2).

This experiment indicates that the variation method makes it possible to determine the model unknown parameter and the external forcing simultaneously. The forcing term is updated at each iteration. Its new evolution gives a new model state which determines the closeness of the model data to the observation. This controls the information the Lagrange multiplier carries, which in turn has a great influence on the new estimation for the wind stress and drag coefficient. The algorithm permits the boundary condition and model to adjust simultaneously and approach the best model data fit efficiently.

3.4 Experiment 4: The simulated current field is generated by a real wind, the wind forcing and the eddy viscosity profile are simultaneously recovered

As in experiment 3, the forcing and the eddy viscosity are all *simultaneously recovered*

As in experiment 3, the forcing and the eddy viscosity are all regarded to be unknown. The current field is created from a real wind

observation, instead of the given wind pattern we used in the previous experiments. The purpose is to test the performance of our variational algorithm in handling the irregular wind stress. The wind data is part of the observations which were acquired from a surface mooring set in the western Sargasso Sea, the details of which will be described in the next section. Figure 6 gives the 10 days wind observations. We use the same parameters as before to create the "observations".

The results of using this set of data are displayed in Figures 7a-7e. It is not surprising to see that there is still a fast reduction in the cost function at the outset of the assimilation. The cost function reaches its optimal value within 11 iterations. The adjustment of the drag coefficient seesaws during the first few iterations (Figure 7e). The eddy viscosity then oscillates around the true solution.

These experiments have been restricted to identical twin runs, i.e., the "observations" are results from the model so they are perfect without any noise. The fast convergence rate and the excellent results hold the promise that the variational method is a feasible approach to solve a problem in which there is an unknown boundary condition, as well as, an unknown viscosity profile. In the following section, we will apply this technique to a real, observed current field.

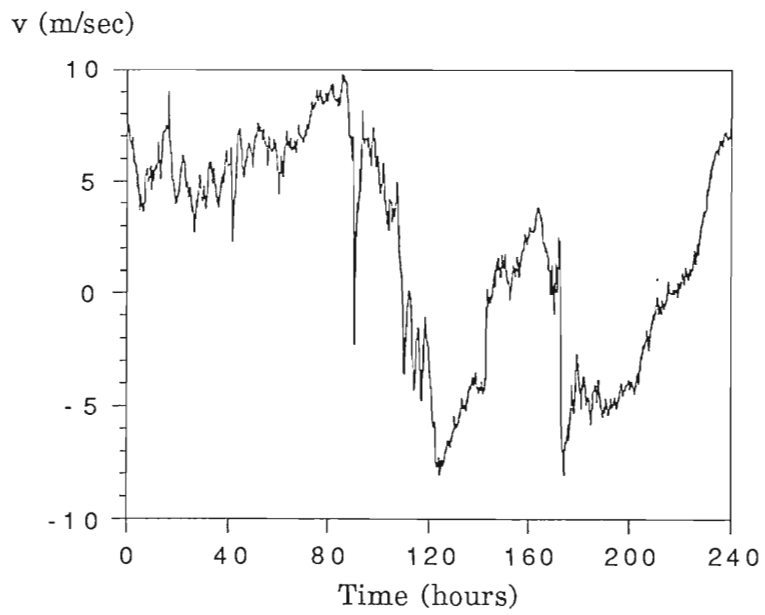
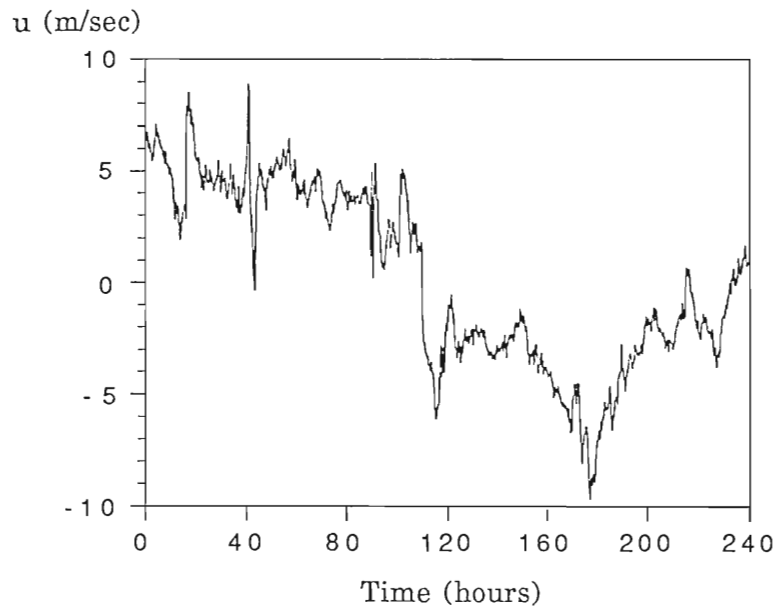
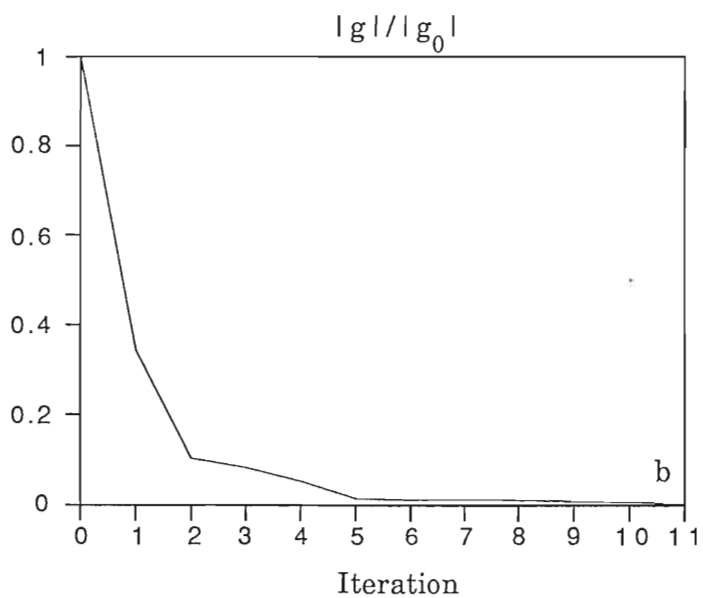
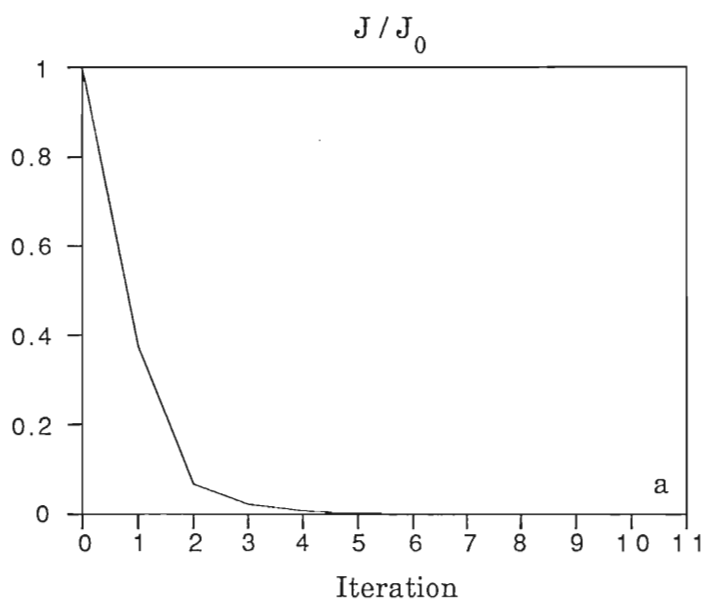


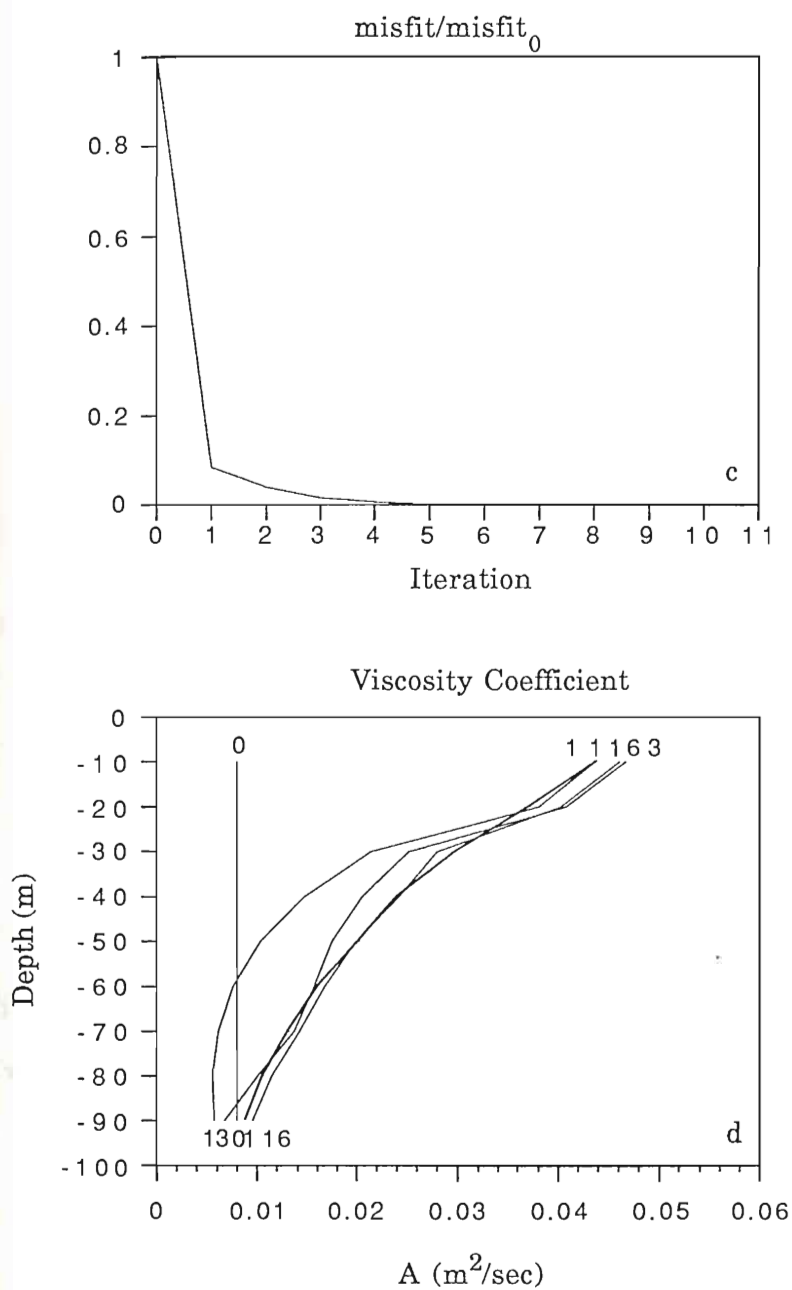
Figure 6 Wind observations

Figure 6 Wind observations



Figures 7a-7b Results from experiment 4. The variation of (a) the cost function, (b) the norm of the gradient with the number of iterations

unction, (b) the norm of the gradient with the number of iterations



Figures 7c-7d Results from experiment 4. (c) The variation of the data misfit with the number of iterations, and (d) The variation of the eddy viscosity profile during the iterative process

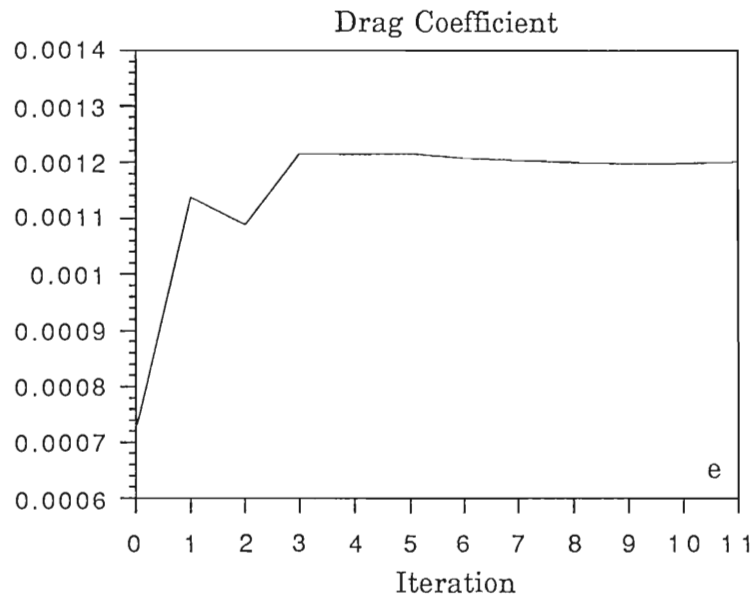


Figure 7e Results from experiment 4. The variation of drag coefficient with the number of iterations

4. The Variational Analysis Technique Applied to an Observed Field

4.1 Data

The oceanographic and meteorological field observations used for our analysis were acquired from the Long-Term Upper Ocean Study -3 (LOTUS-3) buoy by WHOI. This deployment was located in the northwestern Sargasso Sea (34° N, 70° W) during the summer of 1982. In situ current measurements were made by Vector Measuring Current Meters (VMCMs) fixed at depths of 5, 10, 15, 20, 25, 35, 50, 65, 75, and 100 meters. The wind speeds were measured by the Vector-Averaging Wind Recorder (VAWR) which was mounted on the tower of LOTUS-3. The sample interval was 15 minutes. These data were kindly supplied by Briscoe, Price and Weller from WHOI.

Ten days data are chosen from June 30 to July 9, 1982. The time series measurements of wind speed are already plotted in Figure 6. The current observations at 5, 25, 50, and 75m are shown in Figures 8a-8e. Inertial oscillations are dominant at 5 and 25m. At 50m the inertial signal is evident but obviously it is incoherent with the motion in the upper layer. The same thing occurs at 75m. This inconsistency is, perhaps, due to other physical phenomena (e.g. the diurnal tides, the internal waves) or the observation errors.

The measured currents contain the pressure-driven currents (e.g. internal waves) or the observation errors.

The measured currents contain the pressure-driven currents (e.g.

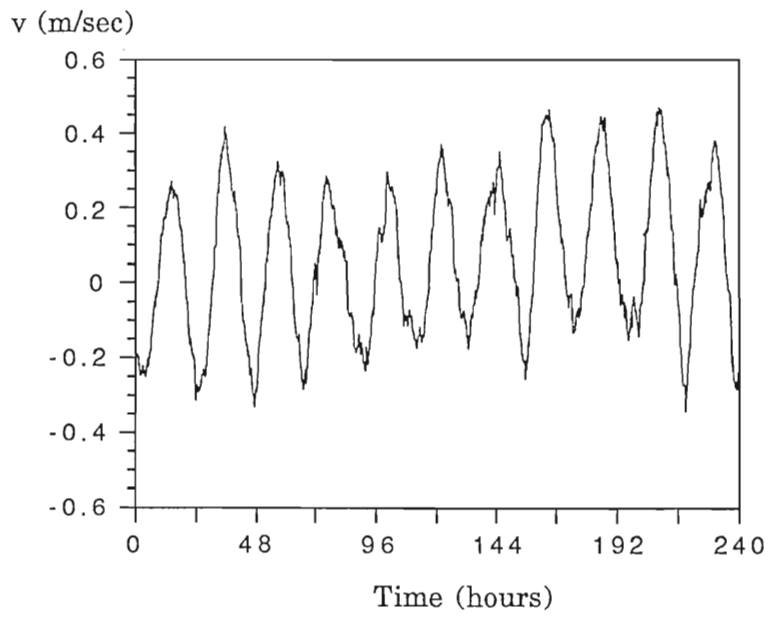
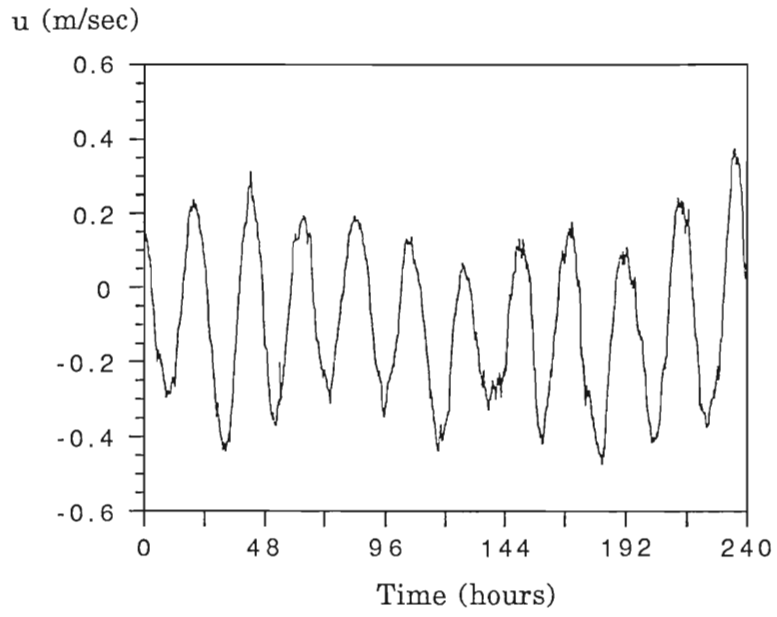


Figure 8a Current observations at 5m

Figure 8a Current observations at 5m

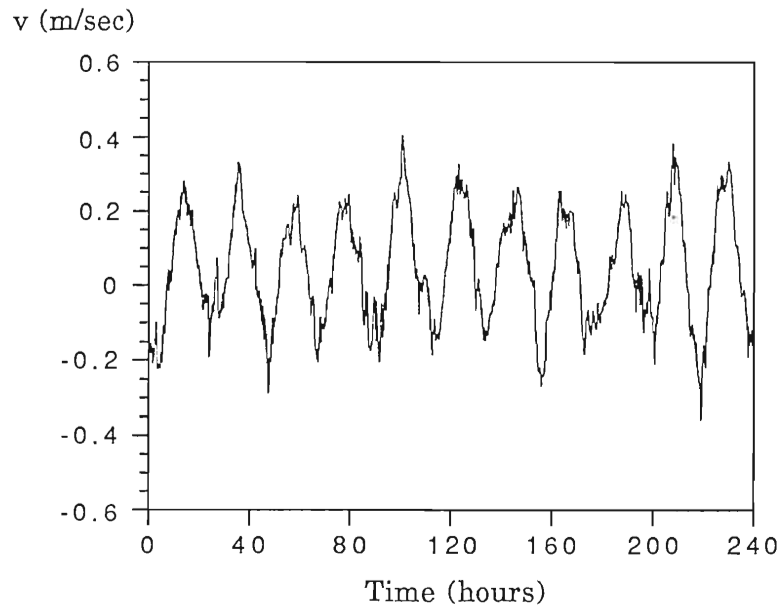
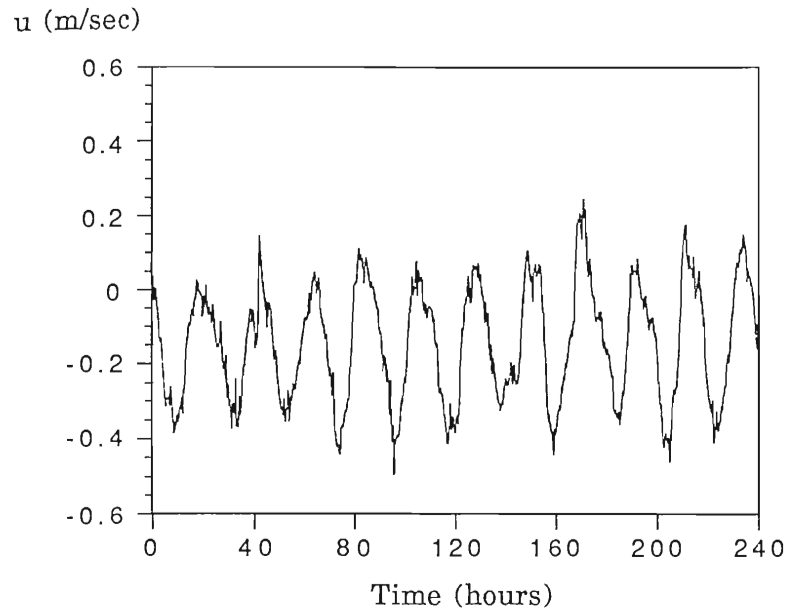


Figure 8b Current observations at 25m

Figure 8b Current observations at 25m

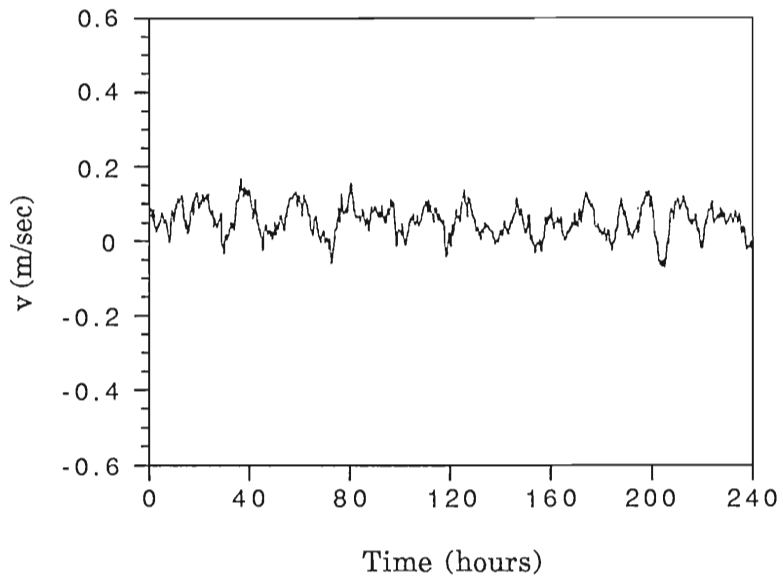
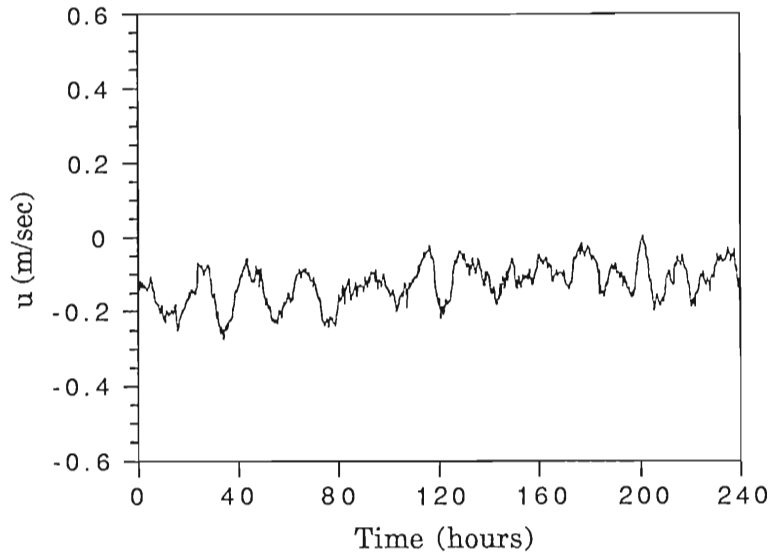


Figure 8c Current observations at 50m

Figure 8c Current observations at 50m

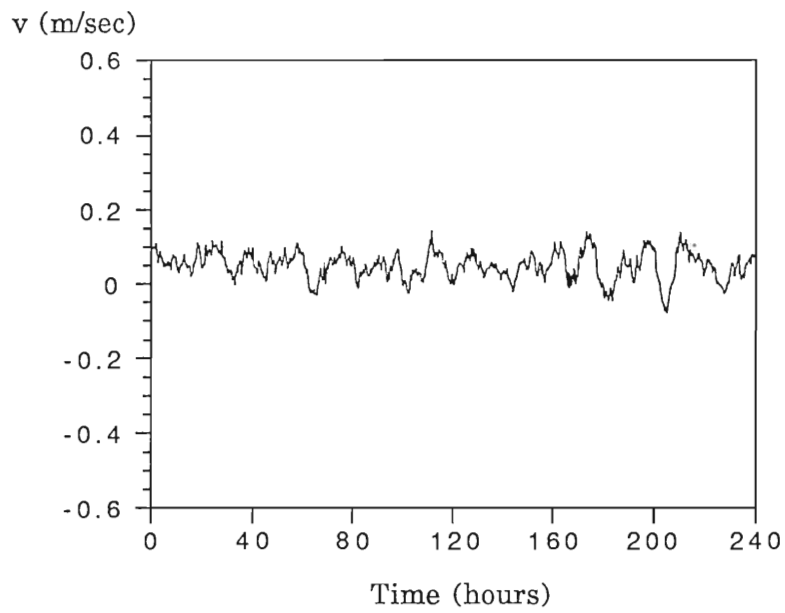
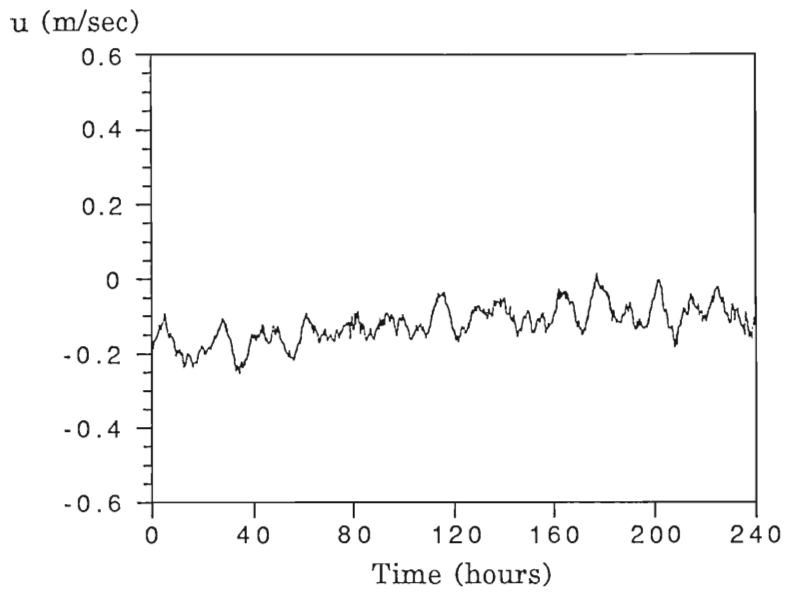


Figure 8d Current observations at 75m

Figure 8d Current observations at 75m

tides, geostrophic motions) besides the locally wind-driven currents. The mean wind-driven current has an amplitude of about 0.05m/sec, while the pressure-driven current has a root mean square value about five times larger (Price et al.,1987). Because we have observations at only one station, it is impossible to compute the pressure-driven motions from the observations. Considering that the time scale of geostrophic motion is much longer than the period of inertial oscillations, we process the data at each depth by removing its trend to filter out the geostrophic components.

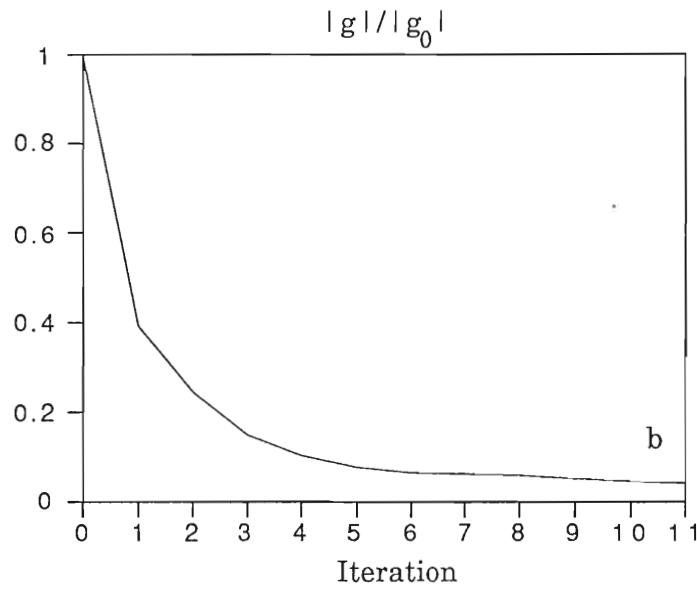
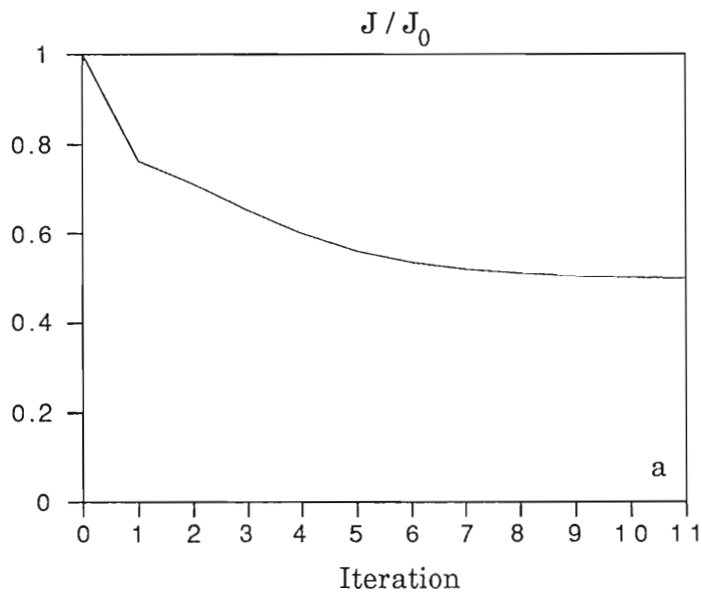
Our numerical model has equally spaced grid points in the vertical, but the data are not available at some depths. Linear interpolation is used to fill the gaps. However, in principle, we do not need additional data at grid points where there are no measurements.

4.2 Results

The observed field at starting time (June 30,1982) is taken as the initial model state. Initial guesses for A and c_D of $0.01 \times 10^{-3} \text{ m}^2/\text{sec}$ and 1.34×10^{-3} , respectively, are chosen. The model is first run forward for 10 days with a time step of 15 minutes which is the same as the sample interval. The data assimilation follows the same procedure as described in the previous section.

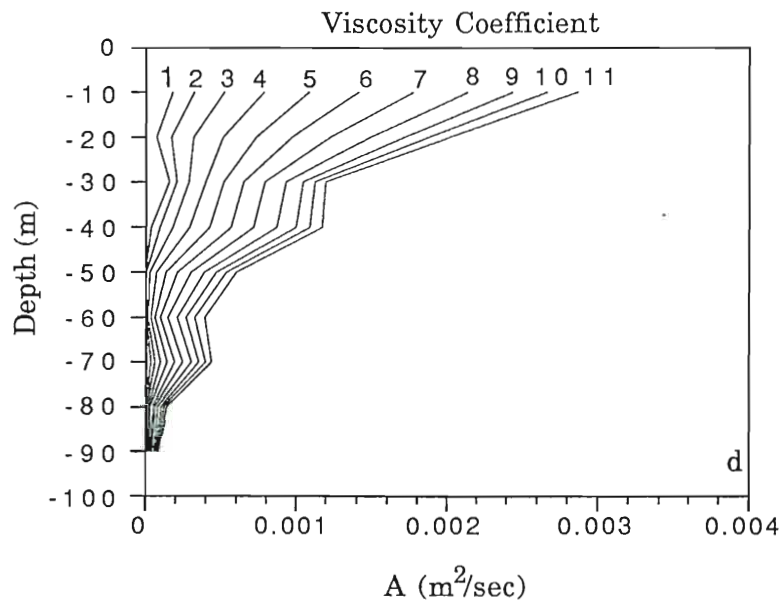
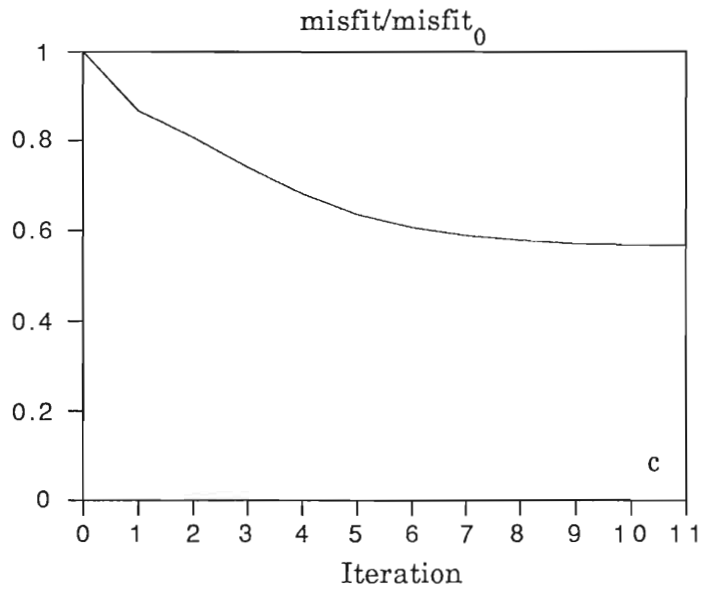
The variation of the cost function with the number of iterations is plotted in Figure 9a. The cost function decreases to about 53% of its

The variation of the cost function with the number of iterations is plotted in Figure 9a. The cost function decreases to about 53% of its initial value after 11 iterations. Figure 9b shows the gradient of the cost



Figures 9a-9b The variation of (a) the cost function, (b) the norm of the gradient with the number of iterations

Figures 9a-9b The variation of (a) the cost function, (b) the norm of the gradient with the number of iterations



Figures 9c-9d (c) The variation of the data misfit with the number of iterations, and (d) The variation of the eddy viscosity profile during the iterative process

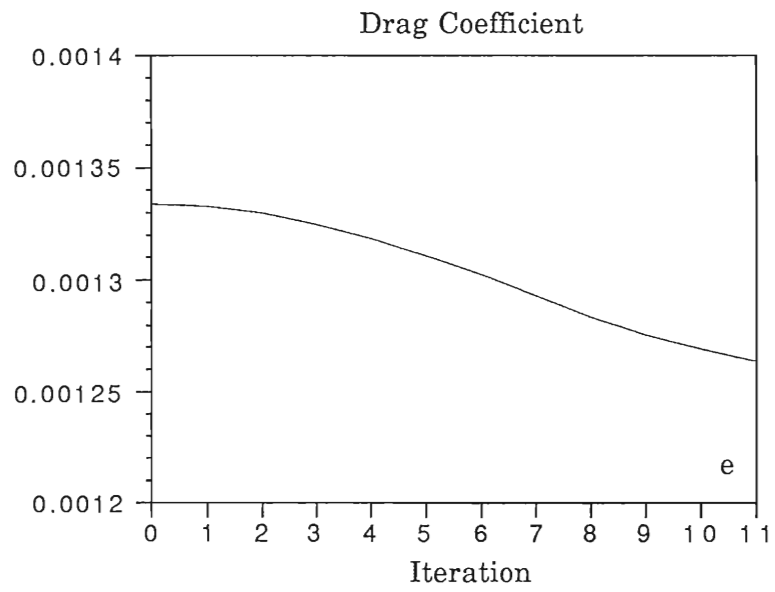


Figure 9e The variation of the drag coefficient with the number of iterations

function. As in the previous experiments, the gradient experiences a rapid decrease during the first few iterations. After iteration 11, it drops to 5% of its initial value and reaches a steady state. Further iterations do not change the results. Since the minimum of a quadratic function of M variables should be found within M iterations when using the conjugate-gradient algorithm, we terminate the minimization process after 11 iterations (there are 11 control variables in our problem). At iteration 11, the eddy viscosity profile (Figure 9d) has a maximum value of $2.9 \times 10^{-3} \text{ m}^2/\text{sec}$ at the surface and decreases with depth. The drag coefficient (Figure 9e) experiences a reduction in its value gradually and reaches 1.26×10^{-3} finally. All these values are very reasonable.

The seasonal thermocline is between 20 to 50 meters during this period. Figure 9d shows that the eddy viscosity decreases greatly within these depths. This is because the stratification suppresses the turbulent mixing in the thermocline, so the degree of the turbulence, and therefore the eddy viscosity, is much smaller than in the mixed surface layer. Obviously, the stratification is well reflected in the eddy viscosity profile.

Consider the correlation coefficient which is defined as

$$r = \frac{\sum_t q q'}{(\sum_t q^2 \sum_t q'^2)^{\frac{1}{2}}}$$

where q and q' are the model results and the observations, respectively. The mean values have been removed from q and q' . The correlation coefficient defined here represents the degree of fitting between the model results and the observations at different depths. Figure 10 is a plot of the correlation coefficient as a function of iterations at depths 5, 25, 75, and 95m. It shows that the correlation coefficient has an increase at all these depths, with the biggest at 5m and the smallest at 95m, but at 75m the correlation coefficient has the lowest value. This can be understood because the motions in the upper 50m are dominant by the inertial oscillations, while below this level, other dynamical processes and the observation errors superimposed on the weak inertial oscillations make the observed current fields very complicated. Our linear dynamical system successfully reproduces the motions in the upper 50m, but is only able to describe a portion of the motions below 50m due to the noisy data. The comparisons of the time series of the modelled and observed current fields are displayed in Figures 11a-11f. Obviously, there is a very good agreement between model results and observations at 5, 15, 25, and 35m. But the erratic changes of amplitude and phase of the observations at 65 and 75m are not reproduced by the simple dynamical model. Undoubtedly, most of the residual data misfits (Figure 9c) come from the lower 50 meters. Since both the cost function and its gradient have a big decrease during the iterative process and reach a steady state after 11 iterations, and the estimates of A and c_D indeed improve the model results, we conclude that the solutions of A and c_D at iteration 11 are the best ones we can derive from these observations.

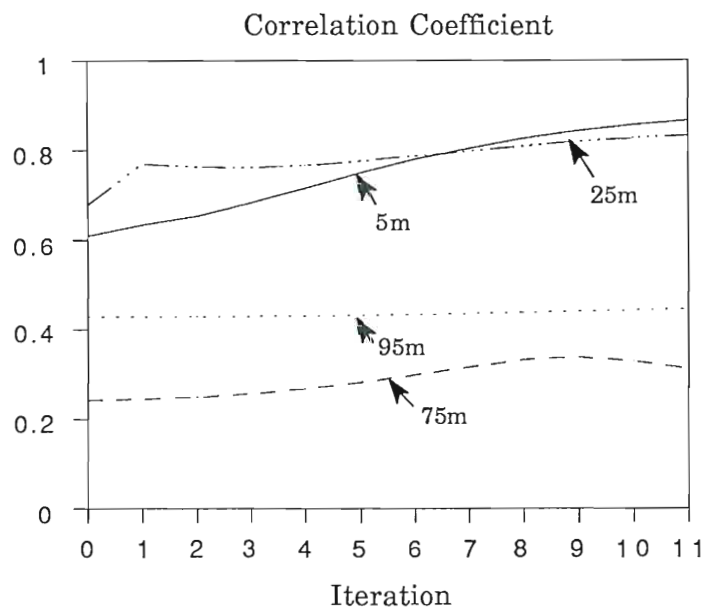


Figure 10 The variation of the correlation coefficient at different depths with the number of iterations

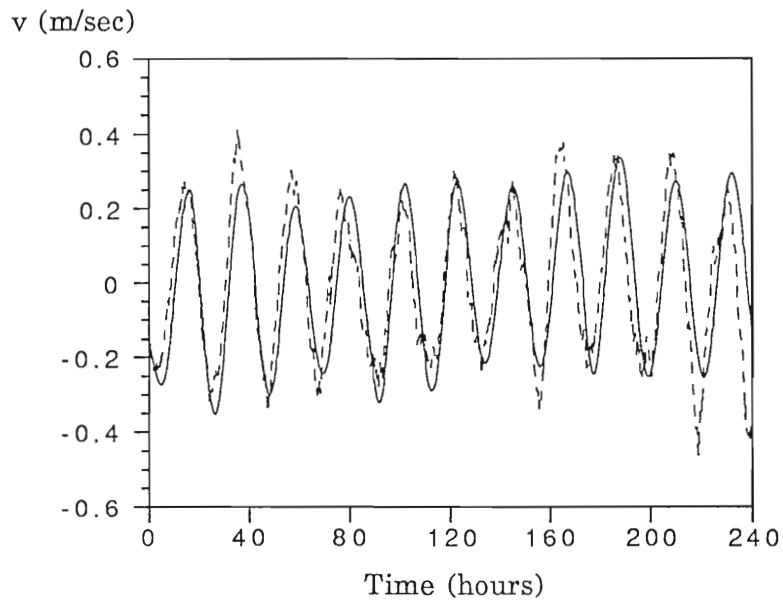
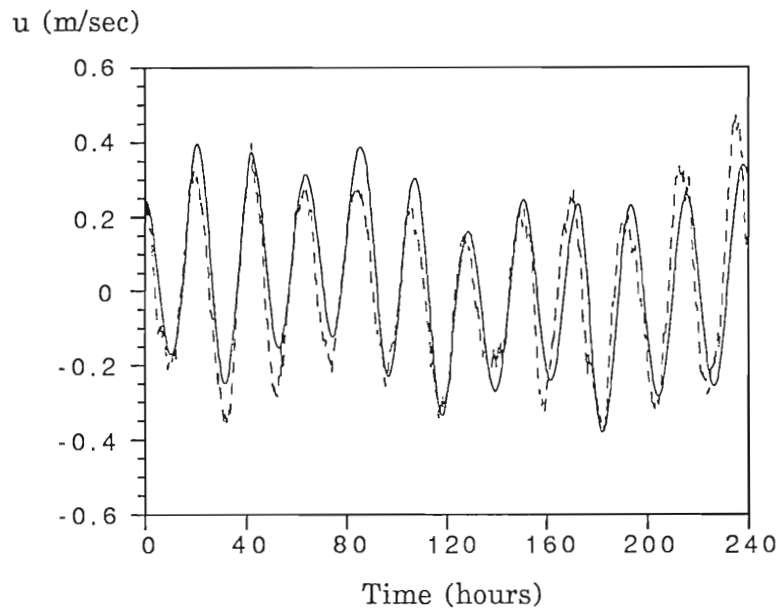


Figure 11a Comparison of modelled(solid) and observed(dashed) current speeds at 5m

Figure 11a Comparison of modelled(solid) and observed(dashed) current speeds at 5m

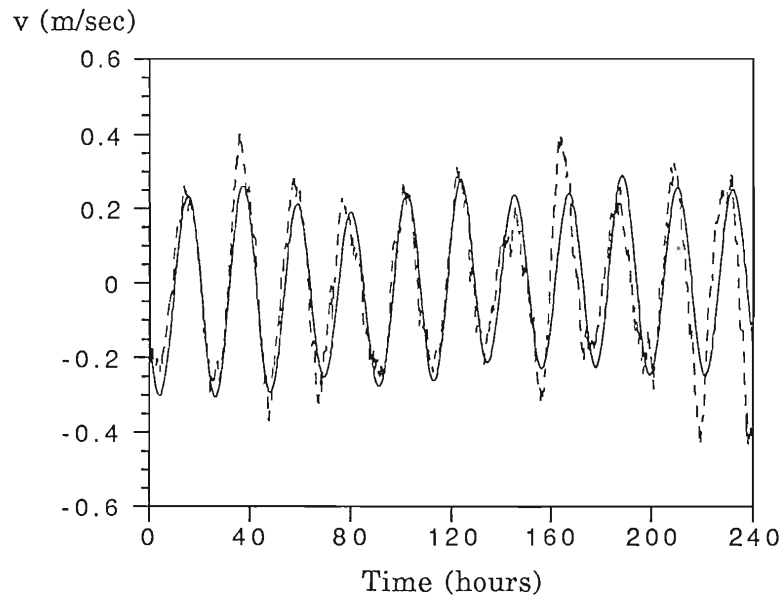
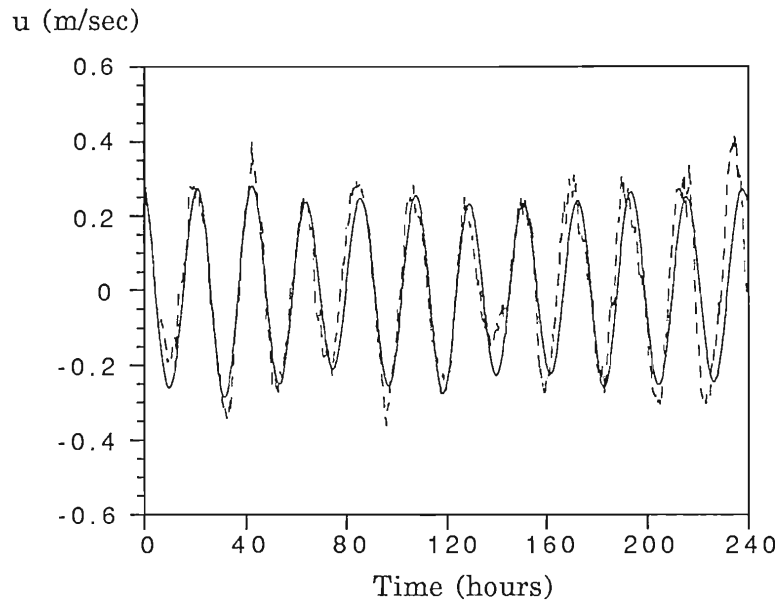


Figure 11b Comparison of modelled(solid) and observed(dashed) current speeds at 15m

Figure 11b Comparison of modelled(solid) and observed(dashed) current speeds at 15m

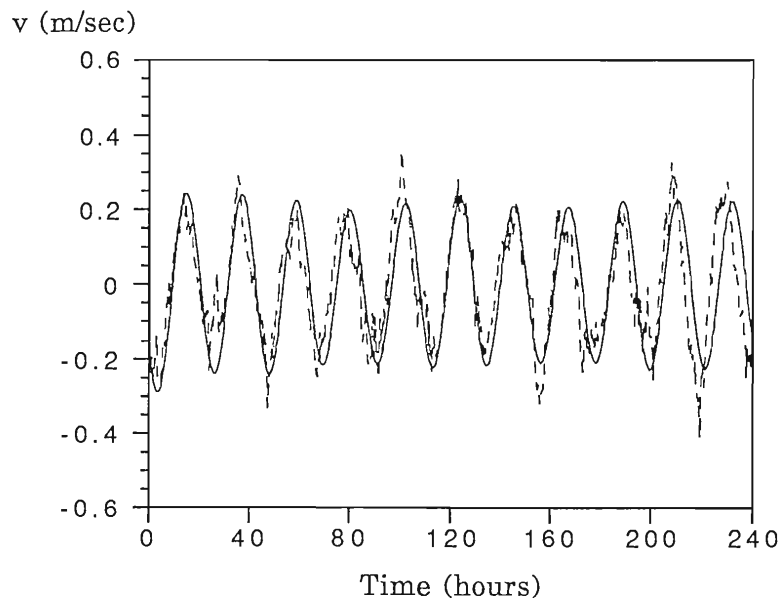
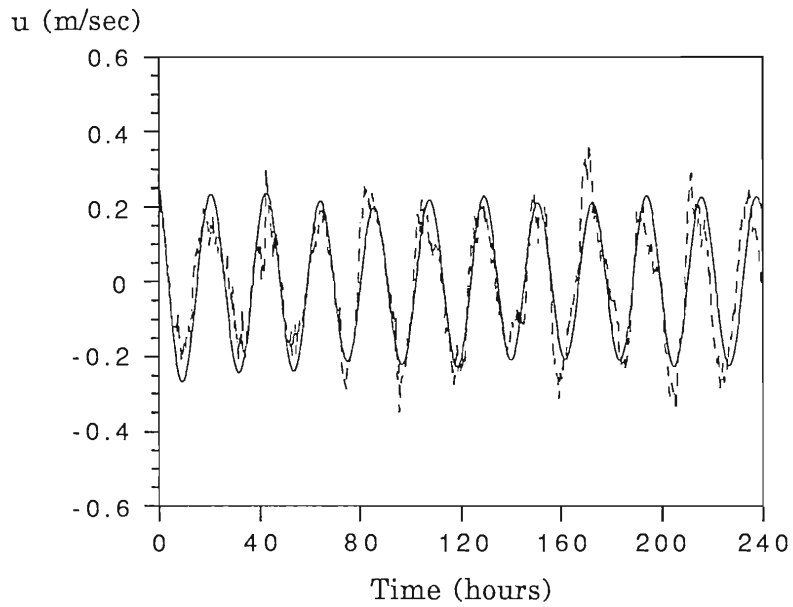


Figure 11c Comparison of the modelled(solid) and observed(dashed) current speeds at 25m

Figure 11c Comparison of the modelled(solid) and observed(dashed) current speeds at 25m

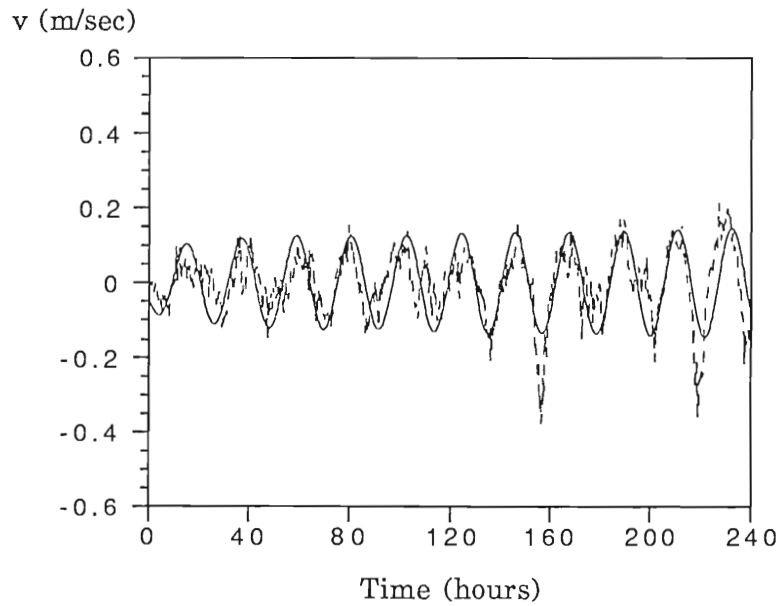
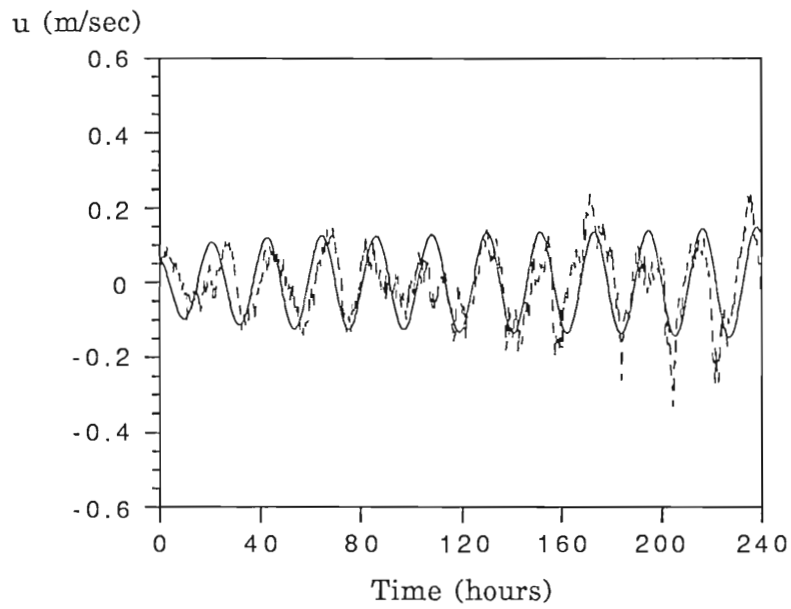


Figure 11d Comparison of the modelled(solid) and observed(dashed) current speeds at 35m

Figure 11d Comparison of the modelled(solid) and observed(dashed) current speeds at 35m

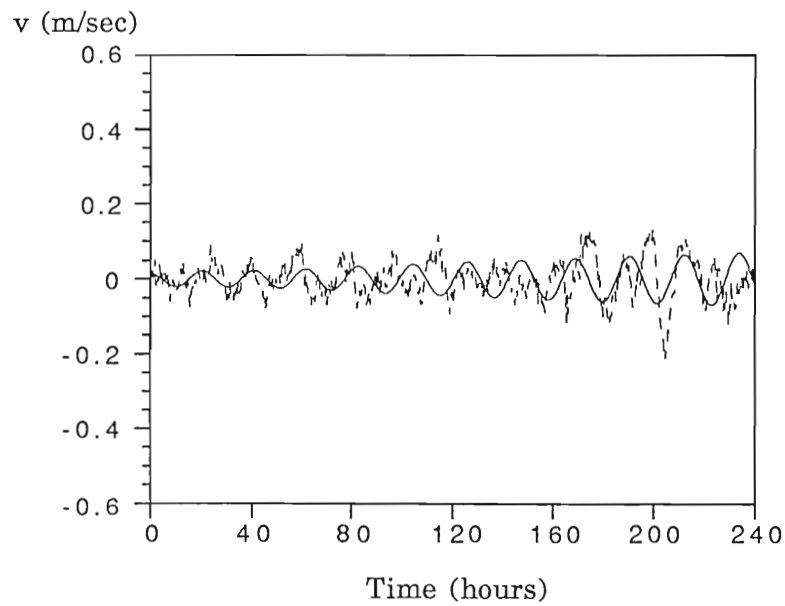
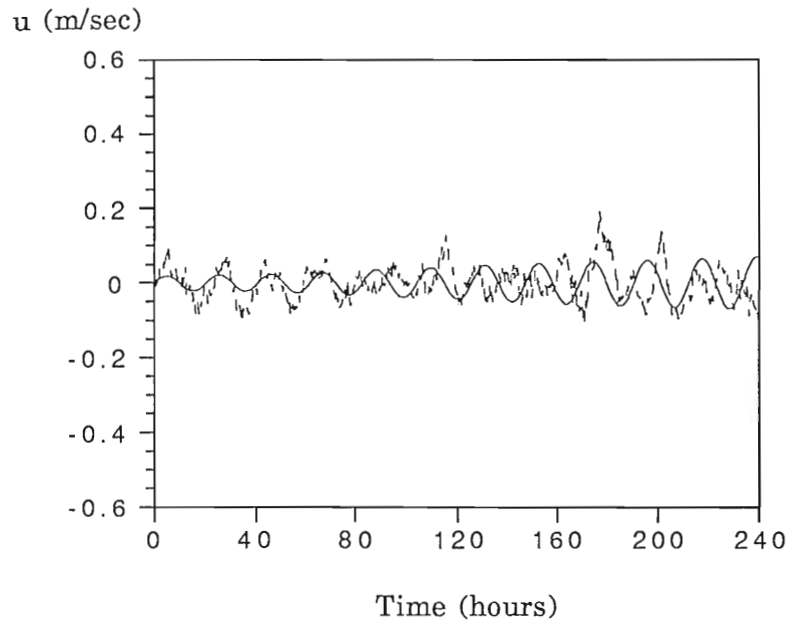


Figure 11e Comparison of the modelled(solid) and observed(dashed) current speeds at 65m

Figure 11e Comparison of the modelled(solid) and observed(dashed) current speeds at 65m

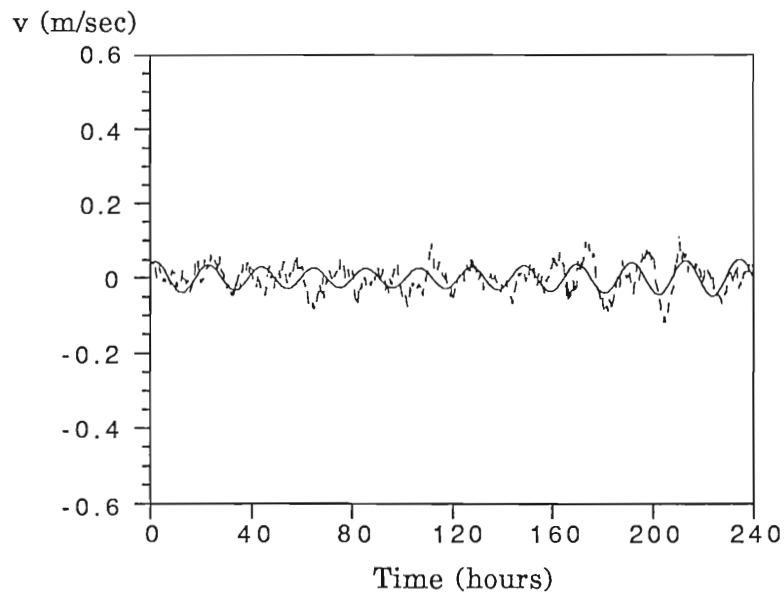
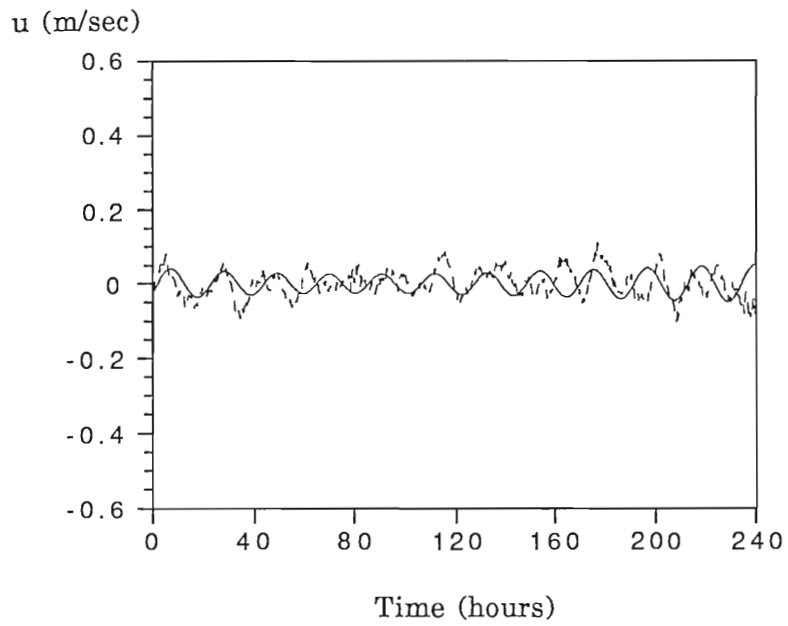


Figure 11f Comparison of the modelled(solid) and observed(dashed) current speeds at 75m

Figure 11f Comparison of the modelled(solid) and observed(dashed) current speeds at 75m

5. Summary and Conclusion

We have demonstrated the utility of a variational-adjoint analysis technique to analyze an observed current data with the aid of a modified upper ocean Ekman model. The wind stress and viscosity profile have been determined simultaneously by fitting model results to observations. The optimal estimates of the model fields have been obtained as well.

The wind stress represents the upper boundary condition in our model, and enters in the model equation as a forcing term in the numerical formalism. The simultaneous adjustment of the wind and eddy viscosity distribution is a positive advantage of the variational analysis for data assimilation. It allows all the dynamics, boundary conditions and observations to influence the model solution and thus, provides a flexible approach in combining the model results with the observations.

The variational-adjoint algorithm has been tested using both identical twin data and real observed data. In identical twin test runs, the "observations" are created by the Ekman model. Three sets of wind patterns are applied to the current field, a constant westerly wind, a sinusoid wind and, a random wind observation, respectively. The first two experiments are considered a simple scenario in which only the eddy viscosity distribution is needed to be determined, assuming that the wind stress is known exactly (constant or sinusiod). The results eddy viscosity distribution is needed to be determined, assuming that the wind stress is known exactly (constant or sinusiod). The results

obtained show that the algorithm is computationally efficient, and for both experiments, convergence is achieved in 6 iterations. It is important to note that the model solution is independent of the choice of the initial guess of the eddy viscosity.

In the following two experiments, the wind stress and the eddy viscosity are adjusted simultaneously by variationally performing the data assimilation. During the iterative process, the wind stress serves as an updating of the current field and its own value is updated at each iteration also. The results illustrate that, whatever for the sinusoid wind (experiment 3) or the real wind observations (experiment 4), the adjoint model algorithm allows the model to fit simultaneously both wind data and current data. The best model-data fit can be obtained within 11 iterations. The very good approximations and the fast convergence obtained are encouraging. They show a promise that the variational-adjoint method is able to solve a problem in which there is an unknown boundary condition and an unknown eddy viscosity.

Real observations from LOTUS-3 records are also assimilated. The wind stress parameter and the eddy viscosity estimated are reasonable. The optimal current field obtained successfully reproduce the observed fields in upper 50m, while due to the noisy data, they are only able to describe part of the motions below 50m. No model can fully describe all the phenomena occurring in the ocean. The observed data can always be divided into three parts, those which are consistent with the model dynamics, those which are inconsistent with the model can always be divided into three parts, those which are consistent with the model dynamics, those which are inconsistent with the model dynamics and, those which are due to the observation errors. Clearly,

the important characteristics of the variational-adjoint analysis is that it is capable of extracting from the available observations that part which are consistent with the model dynamics and adjusting the final model state to the intrinsic dynamics.

The variational-adjoint analysis is conceptually simple and internally consistent. The adjoint model introduced makes the computation of the gradient of the cost function more efficient than the direct perturbation method does. It is clear that the systematic and quantitative approach that the variational-adjoint analysis technique provided can be utilized for a wide variety of diverse problems. It can adjust not only the initial conditions of the model, as most meteorologists do in numerical weather prediction, but also the lateral boundary conditions (Le Dimet and Nouailler, 1985), in the case of a limited area model, the upper boundary conditions, such as the surface forcing by momentum and heat fluxes, and the various physical and numerical parameters which enter the definition of the model.

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 Technical Report*, WHOI-86-5, 51pp.
- Courtier, P. and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part II: Numerical results. *Q. J. R. Meteorol. Soc.*, 113, 1329-1347.
- Derber, J. C., 1985: The variational four-dimensional assimilation of analyses using filtered models as constraints. Ph. D. dissertation, University of Wisconsin, Madison, Wisconsin, USA.
- Derber, J. C., 1987: Variational four-dimensional analysis using quasi-geostrophic constraints. *Mon. Wea. Rev.*, 115, 998-1008.
- Gaspar, P. and C. Wunsch, 1989: Estimates from altimeter data of barotropic Rossby waves in the northwestern Atlantic ocean. *J. Phys. Oceanogr.*, 19, 1821-1844.
- Gill, P. E., W. Murray and M. H. Wright, 1981: *Practical Optimization*. Springer-Verlag, 377pp.
- Hoffman, R. N., 1986: A four-dimensional analysis exactly satisfying equations of motion. *Mon. Wea. Rev.*, 114, 388-397.
- Hopkins, T. S. and L. A. Slatest, 1986: The vertical eddy viscosity in the presence of baroclinic flow in coastal waters. *J. Geophys. Res.*, 91,
- Hopkins, T. S. and L. A. Slatest, 1986: The vertical eddy viscosity in the presence of baroclinic flow in coastal waters. *J. Geophys. Res.*, 91,

14269-14280.

- Le Dimet, F. and I. M. Navon, 1989: Variational and optimization methods in meteorology - a review. Submitted to *Mon. Wea. Rev.*.
- Le Dimet, F. and A. Nouailler, 1985: Assimilation of dynamical data in a limited area model. *Variational Methods in Geosciences*, Y. Sasaki ed., Elsevier, New York, 181-185.
- 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 adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, 37A, 309-322.
- Lions, J. L., 1971: *Optimal control of systems governed by partial differential equations*. Springer-Verlag, 369pp.
- Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 112, 1177-1194.
- Lorenc, A. C., 1988: Optimal nonlinear objective analysis. *Q. J. R. Meteorol. Soc.*, 114, 205-240.
- Luenberger, D. C., 1984: *Linear and nonlinear programming*. Addison-Wesley, 491pp.
- Navon, I. M. and D. Legler, 1987: Conjugate-gradient methods for large-scale minimization in meteorology. *Mon. Wea. Rev.*, 115, 1479-1502.
- O'Brien, J. J., 1986: The diffusive problem. *Advanced Physical Oceanographic Numerical Modelling*, J. J. O'Brien ed., D. Reidel
- O'Brien, J. J., 1986: The diffusive problem. *Advanced Physical Oceanographic Numerical Modelling*, J. J. O'Brien ed., D. Reidel Publ., 127-144.

- Panchang, V. G. and J. J. O'Brien, 1988: On the determination of hydraulic model parameters using the adjoint state formulation. *Modeling Marine System*, I, CRC Press, 5-18.
- Pedlosky, J., 1979: *Geophysical Fluid Dynamics*. Springer-Verlag, 624pp.
- Price, J. F., R. A. Weller and R. Pinkel, 1986: Diurnal cycling: Observations and models of the upper ocean response to diurnal heating, cooling, and wind mixing. *J. Geophys. Res.*, 91, 8411-8427.
- Price, J. F., R. A. Weller and R. R. Schudlich, 1987: Wind-driven ocean currents and Ekman transport. *Science*, 238, 1534-1538.
- Sasaki, Y., 1970: Some basic formations in numerical variational analysis. *Mon. Wea. Rev.*, 98, 875-883.
- Smedstad, O. M., 1989: Data assimilation and parameter estimation in oceanographic models. Ph. D. dissertation, Florida State University, Tallahassee, Florida, USA.
- Talagrand, O. and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Q. J. R. Meteorol. Soc.*, 113, 1311-1328.
- Tarbell, S. A., N. J. Pennington and M. G. Briscoe, 1984: A compilation of moored current meter and wind recorder data. Volume XXXV, Long-Term Upper Ocean Study (LOTUS) (Moorings 764, 765, 766, 767, 770), May 1982-April 1983. *Woods Hole Oceanographic Institution Technical Report*, WHOI-84-36, 154pp.
- Thacker, W. C., 1988: Fitting models to inadequate data by enforcing *Technical Report*, WHOI-84-36, 154pp.
- Thacker, W. C., 1988: Fitting models to inadequate data by enforcing spatial and temporal smoothness. *J. Geophys. Res.*, 93, 10655-10665.

- Thacker, W. C. and R. B. Long, 1988: Fitting dynamics to data. *J. Geophys. Res.*, 93, 1227-1240.
- Tziperman, E. and W. C. Thacker, 1989: An optimal-control/adjoint-equations approach to studying the oceanic general circulation. *J. Phys. Oceanogr.*, 19, 1471-1485.
- Wunsch, C, 1987: Using transient tracers: the regularization problem. *Tellus*, 39B, 477-492.
- Wunsch, C, 1988: Transient tracers as a problem in control theory. *J. Geophys. Res.*, 93, 8099-9110.

Appendix Derivation of the finite difference equations

The finite difference discretized equation for the Augmented Lagrange functional in (4) can be written as

$$L(u_j^n, v_j^n, A_j, c_D, \lambda_{uj}^n, \lambda_{vj}^n) =$$

$$\sum_{n=1}^{N-1} \sum_{j=1}^J \lambda_{uj}^n \left\{ \frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{v_j^{n+1} + v_j^n}{2} - \frac{1}{2\Delta z} (A_j \frac{u_{j-1}^n - u_j^n}{\Delta z} + A_j \frac{u_{j-1}^{n+1} - u_j^{n+1}}{\Delta z} \right. \\ \left. - A_{j+1} \frac{u_j^n - u_{j+1}^n}{\Delta z} - A_{j+1} \frac{u_j^{n+1} - u_{j+1}^{n+1}}{\Delta z} \right\} \Delta z \Delta t$$

$$+ \sum_{n=1}^{N-1} \sum_{j=1}^J \lambda_{vj}^n \left\{ \frac{v_j^{n+1} - v_j^n}{\Delta t} + \frac{u_j^{n+1} + u_j^n}{2} - \frac{1}{2\Delta z} (A_j \frac{v_{j-1}^n - v_j^n}{\Delta z} + A_j \frac{v_{j-1}^{n+1} - v_j^{n+1}}{\Delta z} \right. \\ \left. - A_{j+1} \frac{v_j^n - v_{j+1}^n}{\Delta z} - A_{j+1} \frac{v_j^{n+1} - v_{j+1}^{n+1}}{\Delta z} \right\} \Delta z \Delta t$$

$$+ \frac{1}{2} K_a T \sum_{j=1}^J (A_j - \hat{A}_j)^2 \Delta z + \frac{1}{2} K_c TH (c_D - \hat{c}_D)^2$$

- j=1 - - - - -

$$+ \frac{1}{2} K_m \sum_{n=1}^N \sum_{j=1}^J \{ (u_j^n - \hat{u}_j^n)^2 + (v_j^n - \hat{v}_j^n)^2 \} \Delta z \Delta t \quad (\text{B - 1})$$

where superscript n denotes the time level and subscript j represents the position; λ_{uj}^n and λ_{vj}^n are the Lagrange multipliers for u_j^n and v_j^n , respectively; K_a , K_c and K_m are the Gauss precision moduli.

In the discretized version, the Equations (5) - (8), which are found under the condition that the gradient of the Lagrange function vanishes, take the form

$$\frac{\partial L(u_j^n, v_j^n, A_j, c_D, \lambda_{uj}^n, \lambda_{vj}^n)}{\partial \lambda_{uj}^n} = 0 \quad (\text{B - 2})$$

$$\frac{\partial L(u_j^n, v_j^n, A_j, c_D, \lambda_{uj}^n, \lambda_{vj}^n)}{\partial \lambda_{vj}^n} = 0 \quad (\text{B - 3})$$

$$\frac{\partial L(u_j^n, v_j^n, A_j, c_D, \lambda_{uj}^n, \lambda_{vj}^n)}{\partial u_j^n} = 0 \quad (\text{B - 4})$$

$$\frac{\partial L(u_j^n, v_j^n, A_j, c_D, \lambda_{uj}^n, \lambda_{vj}^n)}{\partial A_j} = 0 \quad (\text{B - 5})$$

$$\frac{\partial L(u_j^n, v_j^n, A_j, c_D, \lambda_{uj}^n, \lambda_{vj}^n)}{\partial v_j^n} = 0 \quad (\text{B - 5})$$

$$\frac{\partial L(u_j^n, v_j^n, A_j, c_D, \lambda_{uj}^n, \lambda_{vj}^n)}{\partial A_j} = 0 \quad (\text{B - 6})$$

$$\frac{\partial L(u_j^n, v_j^n, A_j, c_D, \lambda_{uj}^n, \lambda_{vj}^n)}{\partial c_D} = 0 \quad (\text{B - 7})$$

(B - 2) and (B - 3) recover the original Ekman model equations (12). Using the boundary conditions, the equations of u-component can be expressed as

$$\begin{aligned} & \frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{v_j^{n+1} + v_j^n}{2} - \frac{1}{2\Delta z} \left(A_j \frac{u_{j-1}^n - u_j^n}{\Delta z} + A_j \frac{u_{j-1}^{n+1} - u_j^{n+1}}{\Delta z} \right. \\ & \left. - A_{j+1} \frac{u_j^n - u_{j+1}^n}{\Delta z} - A_{j+1} \frac{u_j^{n+1} - u_{j+1}^{n+1}}{\Delta z} \right) = 0 \end{aligned} \quad (\text{B - 8.1})$$

for $j = 2, 3, \dots, J-1$ and $n = 1, 2, \dots, N-1$.

For $j = 1$, the equation is

--- $j = 2, 3, \dots, J-1$ and $n = 1, 2, \dots, N-1$.

For $j = 1$, the equation is

$$\frac{u_1^{n+1} - u_1^n}{\Delta t} - \frac{v_1^{n+1} + v_1^n}{2} - \frac{1}{2\Delta z} (c_D |w_a^n| u_a^n + c_D |w_a^{n+1}| u_a^{n+1} - A_2 \frac{u_1^n - u_2^n}{\Delta z} - A_2 \frac{u_1^{n+1} - u_2^{n+1}}{\Delta z}) = 0 \quad (\text{B - 8.2})$$

For $j = J$, the equation is

$$\frac{u_J^{n+1} - u_J^n}{\Delta t} - \frac{v_J^{n+1} + v_J^n}{2} - \frac{1}{2\Delta z} (A_J \frac{u_{J-1}^n - u_J^n}{\Delta z} + A_J \frac{u_{J-1}^{n+1} - u_J^{n+1}}{\Delta z}) = 0 \quad (\text{B - 8.3})$$

The equations for the v-component have the similar forms. For simplicity, we are not going to give the explicit expressions.

(B - 4) and (B - 5) result in the adjoint models which the Lagrange multipliers must be satisfied. Let's derive (B - 4) which considers the gradient of L with respect to u_j^n .

$$\frac{\partial L}{\partial u_j^n} = \sum_{n=1}^{N-1} \sum_{j=1}^J \frac{\partial}{\partial u_j^n} \{ \lambda_{uj}^n [\frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{1}{2\Delta z} (A_j \frac{u_{j-1}^n - u_j^n}{\Delta z} + A_j \frac{u_{j-1}^{n+1} - u_j^{n+1}}{\Delta z})] \}$$

$$\frac{\partial L}{\partial u_j^n} = \sum_{n=1}^N \sum_{j=1}^J \frac{\partial}{\partial u_j^n} \{ \lambda_{uj}^n [\frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{1}{2\Delta z} (A_j \frac{u_{j-1}^n - u_j^n}{\Delta z} + A_j \frac{u_{j-1}^{n+1} - u_j^{n+1}}{\Delta z})] \}$$

$$\begin{aligned}
& - A_{j+1} \frac{u_j^n - u_{j+1}^n}{\Delta z} - A_{j+1} \frac{u_j^{n+1} - u_{j+1}^{n+1}}{\Delta z} \Big) \Big] \Delta z \Delta t \\
& + \sum_{n=1}^{N-1} \sum_{j=1}^J \frac{\partial}{\partial u_j^n} \left\{ \lambda_{vj}^n \frac{u_j^{n+1} + u_j^n}{2} \right\} \Delta z \Delta t + \frac{1}{2} K_m \sum_{n=1}^N \sum_{j=1}^J \frac{\partial}{\partial u_j^n} (u_j^n - \hat{u}_j^n)^2 \Delta z \Delta t \\
& = \sum_{n=2}^N \sum_{j=1}^J \left\{ \frac{\lambda_{uj}^{n-1} - \lambda_{uj}^n}{\Delta t} - \frac{1}{2\Delta z} \left(\frac{A_{j+1} \lambda_{uj+1}^n - A_j \lambda_{uj}^n}{\Delta z} + \frac{A_{j+1} \lambda_{uj+1}^{n-1} - A_j \lambda_{uj}^{n-1}}{\Delta z} \right. \right. \\
& \quad \left. \left. - \frac{A_{j+1} \lambda_{uj}^n - A_j \lambda_{uj-1}^n}{\Delta z} - \frac{A_{j+1} \lambda_{uj}^{n-1} - A_j \lambda_{uj-1}^{n-1}}{\Delta z} \right) \right\} \Delta z \Delta t \\
& + \sum_{n=2}^N \sum_{j=1}^J \left\{ \frac{\lambda_{vj}^{n-1} + \lambda_{vj}^n}{2} \right\} \Delta z \Delta t + K_m \sum_{n=1}^N \sum_{j=1}^J (u_j^n - \hat{u}_j^n) \Delta z \Delta t
\end{aligned}$$

(rearranging terms)

$$\begin{aligned}
& = \sum_{n=2}^N \sum_{j=1}^J \left\{ \frac{\lambda_{uj}^{n-1} - \lambda_{uj}^n}{\Delta t} + \frac{\lambda_{vj}^{n-1} + \lambda_{vj}^n}{2} - \frac{1}{2\Delta z} \left[A_j \frac{\lambda_{uj-1}^n - \lambda_{uj}^n}{\Delta z} \right. \right. \\
& = \sum_{n=2}^N \sum_{j=1}^J \left\{ \frac{\lambda_{uj}^{n-1} - \lambda_{uj}^n}{\Delta t} + \frac{\lambda_{vj}^{n-1} + \lambda_{vj}^n}{2} - \frac{1}{2\Delta z} \left[A_j \frac{\lambda_{uj-1}^n - \lambda_{uj}^n}{\Delta z} \right. \right.
\end{aligned}$$

$$\begin{aligned}
& + A_j \frac{\lambda_{uj-1}^{n-1} - \lambda_{uj}^{n-1}}{\Delta z} - A_{j+1} \frac{\lambda_{uj}^n - \lambda_{uj+1}^n}{\Delta z} - A_{j+1} \frac{\lambda_{uj}^{n-1} - \lambda_{uj+1}^{n-1}}{\Delta z}] \Delta z \Delta t \\
& + K_m \sum_{n=1}^N \sum_{j=1}^J (u_j^n - \hat{u}_j^n) \Delta z \Delta t \tag{B - 9}
\end{aligned}$$

With the boundary conditions

$$\frac{\lambda_{u1}^n - \lambda_{u2}^n}{\Delta z} = 0$$

$$\frac{\lambda_{uJ}^n - \lambda_{uJ+1}^n}{\Delta z} = 0$$

the condition that (B - 9) vanishes yields

$$\frac{\lambda_{uj}^{n-1} - \lambda_{uj}^n}{\Delta t} + \frac{\lambda_{vj}^{n-1} + \lambda_{vj}^n}{2} - \frac{1}{2\Delta z} \left[A_j \frac{\lambda_{uj-1}^n - \lambda_{uj}^n}{\Delta z} + A_j \frac{\lambda_{uj-1}^{n-1} - \lambda_{uj}^{n-1}}{\Delta z} \right.$$

$$\left. - A_{j+1} \frac{\lambda_{uj}^n - \lambda_{uj+1}^n}{\Delta z} - A_{j+1} \frac{\lambda_{uj}^{n-1} - \lambda_{uj+1}^{n-1}}{\Delta z} \right] + K_m (u_j^n - \hat{u}_j^n) = 0 \tag{B - 10.1}$$

$$\left. - A_{j+1} \frac{\lambda_{uj}^{n-1} - \lambda_{uj+1}^{n-1}}{\Delta z} - A_{j+1} \frac{\lambda_{uj}^n - \lambda_{uj+1}^n}{\Delta z} \right] + K_m (u_j^n - \hat{u}_j^n) = 0 \tag{B - 10.1}$$

for $j = 2, 3, \dots, J-1$ and $n = 1, 2, \dots, N$.

$$\begin{aligned} & \frac{\lambda_{u1}^{n-1} - \lambda_{u1}^n}{\Delta t} + \frac{\lambda_{v1}^{n-1} + \lambda_{v1}^n}{2} + \frac{1}{2\Delta z} \left[A_2 \frac{\lambda_{u1}^n - \lambda_{u2}^n}{\Delta z} + A_2 \frac{\lambda_{u1}^{n-1} - \lambda_{u2}^{n-1}}{\Delta z} \right] \\ & + K_m (u_j^n - \hat{u}_j^n) = 0 \end{aligned} \quad (\text{B - 10.2})$$

for $j = 1$ and $n = 1, 2, \dots, N$.

$$\begin{aligned} & \frac{\lambda_{uJ}^{n-1} - \lambda_{uJ}^n}{\Delta t} + \frac{\lambda_{vJ}^{n-1} + \lambda_{vJ}^n}{2} - \frac{1}{2\Delta z} \left[A_J \frac{\lambda_{uJ-1}^n - \lambda_{uJ}^n}{\Delta z} + A_J \frac{\lambda_{uJ-1}^{n-1} - \lambda_{uJ}^{n-1}}{\Delta z} \right] \\ & + K_m (u_J^n - \hat{u}_J^n) = 0 \end{aligned} \quad (\text{B - 10.3})$$

for $j = J$ and $n = 1, 2, \dots, N$.

Differentiation of L with respect to v_j^n yields the similar expressions as (B - 10.1) - (B - 10.3).

The differentiation of L with respect to A_j yields

$$\begin{aligned} \frac{\partial L}{\partial A_j} &= \sum_{n=1}^{N-1} \sum_{j=1}^J \frac{\partial}{\partial A_j} \left\{ \lambda_{uj}^n \left[-\frac{1}{2\Delta z} \left(A_j \frac{u_{j-1}^n - u_j^n}{\Delta z} + A_j \frac{u_{j-1}^{n+1} - u_j^{n+1}}{\Delta z} \right) \right. \right. \\ & \left. \left. - \frac{v_j^n}{\Delta z} \right] \right\} \end{aligned}$$

$$\begin{aligned}
& - A_{j+1} \frac{u_j^n - u_{j+1}^n}{\Delta z} - A_{j+1} \frac{u_j^{n+1} - u_{j+1}^{n+1}}{\Delta z} \Big] \Big] \Delta z \Delta t \\
& + \sum_{n=1}^{N-1} \sum_{j=1}^J \frac{\partial}{\partial A_j} \left\{ \lambda_{vj}^n \left[- \frac{1}{2\Delta z} \left(A_j \frac{uv_{j-1}^n - v_j^n}{\Delta z} + A_j \frac{v_{j-1}^{n+1} - v_j^{n+1}}{\Delta z} \right. \right. \right. \\
& \left. \left. \left. - A_{j+1} \frac{v_j^n - v_{j+1}^n}{\Delta z} - A_{j+1} \frac{v_j^{n+1} - v_{j+1}^{n+1}}{\Delta z} \right) \right] \right\} \Delta z \Delta t \\
& + \frac{1}{2} K_a T \sum_{j=1}^J \frac{\partial}{\partial A_j} (A_j - \hat{A}_j)^2 \Delta z \\
& = \sum_{j=2}^J \sum_{n=1}^{N-1} \left\{ - \frac{1}{2} \left[\left(\frac{u_{j-1}^n - u_j^n}{\Delta z} + \frac{u_{j-1}^{n+1} - u_j^{n+1}}{\Delta z} \right) \frac{\lambda_{uj-1}^n - \lambda_{uj}^n}{\Delta z} \right. \right. \\
& \left. \left. + \frac{1}{2} \left(\frac{v_{j-1}^n - v_j^n}{\Delta z} + \frac{v_{j-1}^{n+1} - v_j^{n+1}}{\Delta z} \right) \frac{\lambda_{vj-1}^n - \lambda_{vj}^n}{\Delta z} \right] \right\} \Delta t \Delta z \\
& \dots \\
& + K_a T \sum_{j=2}^J (A_j - \hat{A}_j) \Delta z
\end{aligned}$$

Therefore, $\frac{\partial L}{\partial A_j} = 0$ gives

$$\begin{aligned} \frac{1}{2} \sum_{n=1}^{N-1} \left\{ \left(\frac{u_{j-1}^n - u_j^n}{\Delta z} + \frac{u_{j-1}^{n+1} - u_j^{n+1}}{\Delta z} \right) \frac{\lambda_{uJ-1}^n - \lambda_{uJ}^n}{\Delta z} \right. \\ \left. + \left(\frac{v_{j-1}^n - v_j^n}{\Delta z} + \frac{v_{j-1}^{n+1} - v_j^{n+1}}{\Delta z} \right) \frac{\lambda_{vJ-1}^n - \lambda_{vJ}^n}{\Delta z} \right\} \Delta t + K_a T (A_j - \hat{A}_j) = 0 \end{aligned}$$

(B - 11)

for $j = 2, 3, \dots, J$.

At last from the differentiation of L with respect to c_D one can obtain

$$\begin{aligned} - \frac{1}{2} \sum_{n=1}^{N-1} \left\{ \lambda_{u1}^n (|w_a^n| u_a^n + |w_a^{n+1}| u_a^{n+1}) + \lambda_{v1}^n (|w_a^n| v_a^n + |w_a^{n+1}| v_a^{n+1}) \right\} \Delta t \\ + K_c T H (c_D - \hat{c}_D) = 0 \end{aligned}$$

(B - 12)