Proc. Indian Acad. Sci. (Earth Planet. Sci.), Vol. 102, No. 4, December 1993, pp. 521–536.

The variational technique of data assimilation using adjoint equations in a shallow water model

SUJIT BASU, V SUBRAMANIAN and P C PANDEY

Meteorology and Oceanography Division, Space Applications Centre, Ahmedabad 380053, India

MS received 7 December 1992; revised 29 March 1993

Abstract. The variational technique of data assimilation using adjoint equations has been illustrated using a nonlinear oceanographic shallow water model. The technique consists of minimizing a cost function representing the misfit between the model and the data subject to the model equations acting as constraints. The problem has been transformed into an unconstrained one by the use of Lagrange multipliers. Particular emphasis has been laid on finite difference formulation of the algorithm. Several numerical experiments have been conducted using simulated data obtained from a control run of the model. Implications of this technique for assimilating asynoptic satellite altimeter data into ocean models have been discussed.

Keywords. Data assimilation; variational technique; adjoint equations; nonlinear minimization.

1. Introduction

Although data assimilation has long been recognized as an integral component of numerical weather prediction, it has become a subject of interest to the ocean modellers only in recent years. One principal difference between atmospheric and oceanic prediction is that there is no oceanic counterpart to the extensive global synoptic meteorological observing network (Thacker and Long 1988). With the availability, however, of good quality data from ocean-observing satellites such as the recently launched ERS-1 of European Space Agency and Topex/Poseidon, a joint venture of USA and France, the situation promises to be good. However, observations like those of sea surface wind (by satellite scatterometer) and sea surface height (by satellite altimeter) are asynoptic and confined to the surface of the ocean. Ocean models and sophisticated data assimilation techniques will thus have to play a crucial role in successfully extrapolating these surface observations to the ocean interior (Holland and Malanotte-Rizzoli 1989) and even in obtaining estimates of unobserved variables from the observations of other variables. What we mean by extrapolation is that the model equations e.g., those for a multi-level ocean model, will ensure that the impact of the surface observations after their assimilation is felt by lower levels. Further, as we hope to show in this paper, the variational data assimilation technique essentially consists of providing the best possible estimate of the model initial conditions from observations of certain variables like sea surface height. Thus, although some variables like currents may never be observed or may be observed at only a few locations, the data assimilation technique seeks to provide an estimate of these unobserved variables also, so as to improve the predictive capability of the model.

With the launching of various space missions such as IRS series and INSAT satellites, the Indian Space Research Organisation (ISRO) is actively planning future environmental satellites which may have ocean-observing sensors, a fundamentally important one being an altimeter which is basically an active microwave instrument used for ocean research. It measures the height of the sea surface relative to the marine geoid and also the ocean surface wind speed and significant wave height (Basu and Pandey 1991). It is thus important to study how to assimilate the vast amount of asynoptic satellite data into various numerical ocean models. The present paper can be reckoned as a modest beginning in this direction.

Briefly speaking, a data assimilation technique is one which seeks to optimally combine the measurements of one or more dynamical variables with the corresponding model predicted values to obtain a better estimate of the current state of the model and hence to improve its forecast capability. Climatology can also provide a good first guess for the initial conditions. Several such techniques are currently available (Ghil and Malanotte-Rizzoli 1991), a recent and powerful one being a variational technique using adjoint equations (Lewis 1985; Le Dimet and Talagrand 1986; Talagrand and Courtier 1987; Thacker and Long 1988). The present work attempts to study the impact of this technique on the predictive capability of a simple ocean model based on one-dimensional nonlinear shallow water equations without any forcing. Adjoint data assimilation can be formulated using the concepts of functional analysis such as Hilbert Space, linear operators etc. This has been the route taken by most investigators (Lewis 1985; Le Dimet and Talagrand 1986; Talagrand and Courtier 1987; Schröter 1989). However, as pointed out by Thacker and Long (1988), these formulations tend to get bogged down in unnecessary details of functional analysis. Also, the adjoint equations are derived in the form of partial differential equations whose discretization along with the application of proper boundary conditions remain somewhat ambiguous and arbitrary. However, discretization of these equations has to be in conformity with that of original model equations. To bring out this fact we take the alternative and straightforward approach of simple partial differentiation with the use of Lagrange multipliers. The advantage here is that the discretized adjoint equations will be directly derived leaving no scope for any ambiguities. In actual practice, one has to ultimately solve the discretized equations. Thus there does not seem to be any point in describing the technique using the functional-analytic approach.

2. Model formulation

The model is based on the following nonlinear shallow water equations in a none-dimensional cyclic domain

$$(\partial v/\partial t) + v(\partial v/\partial x) - A_M(\partial^2 v/\partial x^2) + (\partial \phi/\partial x) = 0, \tag{1}$$

$$(\partial \phi/\partial t) + \partial(v\phi)/\partial x = 0. \tag{2}$$

Here v is the velocity, A_M the viscosity and ϕ the potential ($\phi = gh$, h being the water depth). Equation (1) is the usual momentum equation in the absence of external forcing and equation (2) is the continuity equation. To start with, we non-dimensionalize equations (1) and (2)

$$(\partial v'/\partial t') + v'(\partial v'/\partial x') - A'_M(\partial^2 v'/\partial x'^2) + (\partial \phi'/\partial x') = 0, \tag{3}$$

$$(\partial \phi'/\partial t') + \partial(v'\phi')/\partial x' = 0, \tag{4}$$

where the primed quantities are non-dimensional variables. The space variable has been scaled by length L_x of the cyclic domain, time has been scaled by (L_x/v_0) and v and ϕ have been scaled by v_0 and ϕ_0 respectively, these being the average initial values. Further, $A'_M(=A_M/L_x v_0)$ denotes the non-dimensionalized viscosity. We have also taken v_0 equal to $\sqrt{\phi_0}$ for simplicity. Since after non-dimensionalization, L_x does not appear in the model equations, its numerical value is not of any importance. However, typically L_x is of the order of few thousand kms, and v_0 is of the order of a few cms/sec. Primes denoting nondimensional variables are dropped hereafter for convenience.

3. Variational data assimilation using adjoint equations

The approach consists of reconstructing the present model state by fitting the forecast model to all data gathered in the recent past. The phase space trajectory of the model is fitted to the available data in an optimal manner. Usually the optimal fitting is done in a least squares sense with the model equations acting as constraints. Optimization is achieved by minimizing a cost function representing the misfit between the model and the data. Minimization is done with respect to a set of control variables. In our case, the control variables are nothing but the model initial conditions. The problem is thus one of optimal control, or constrained optimization (constraints being the model equations). Since the number of initial conditions is generally quite large, a large-scale minimization technique such as the conjugate gradient method (Navon and Legler 1987) has to be usually employed. The method is iterative and requires the computation of the components of the gradient of the function to be minimized (cost function) at each iteration. Because of the large number of components involved ordinary finite difference methods are prohibitive from the computational point of view. The elegance of the adjoint approach lies in that instead of finite difference computation of the gradient, one computes the gradient by solving a set of equations, adjoint to the linearized version of the model equations and forced by the model data misfit. Of course, the original model equations have to be integrated first. The technique also requires a backward integration of the adjoint equations. The values of the adjoint variables at the model initial time contain the gradient information to be used by the minimization routine. In what follows, we will illustrate the ideas sketched above in the concrete case of the one-dimensional ocean model represented by equations (3) and (4).

4. Finite difference formulation

The model equations are discretized using a staggered grid and leap-frog time differencing scheme in the following manner.

$$v_i^n = v_i^{n-2} - \mathrm{A}[(v_{i-1}^{n-1} + v_{i+1}^{n-1} + 2v_i^{n-1})(v_{i+1}^{n-1} - v_{i-1}^{n-1})] + \mathrm{B}(v_{i+1}^{n-2} + v_{i-1}^{n-2} - 2v_i^{n-2}) - \mathrm{C}(\phi_i^{n-1} - \phi_{i-1}^{n-1}) \tag{5}$$

$$\phi_i^n = \phi_i^{n-2} - (\mathrm{C}/2)[v_{i+1}^{n-1}(\phi_{i+1}^{n-1} + \phi_i^{n-1}) - v_i^{n-1}(\phi_{i-1}^{n-1} + \phi_i^{n-1})]. \tag{6}$$

In equations (5) and (6) i and n denote spatial and temporal indices. While i ranges from 1 to I $(=41$ in our case), n varies from 2 to N (the total number of time steps of integration). We have used $N=200$ in the present study. Further, the constants

A, B, C are related to the time step τ, grid interval d and viscosity A_M in the following manner

$$A = (\tau/4d), \quad B = (2\tau A_M/d^2), \quad C = 2\tau/d.$$

We have used $\tau = 3{\cdot}125 \times 10^{-3}$ and $A_M = 0{\cdot}01$. The value of d is 0·025 since the spatial extent is normalized to unity. It can be seen that a backward-differencing has been employed for the diffusion term as is the common practice. The particular spatial differencing of the advection term in equation (5) ensures numerical stability. To apply equations (5) and (6) at the boundaries $i = 1$ and $i = I$ we use cyclic boundary conditions. In other words, the two boundaries are identified and consequently $i = -1$ is identified as $i = I - 1$ and $i = I + 1$ is the same as $i = 2$. Equations (5) and (6) can thus be applied at internal points as well as at the boundaries.

To start the integration, i.e. to compute the values after the first time step, we use forward differencing.

$$v_i^1 = v_i^0 - (A/2)[(v_{i+1}^0 + v_{i-1}^0 + 2v_i^0)(v_{i+1}^0 - v_{i-1}^0)] - (C/2)(\phi_i^0 - \phi_{i-1}^0) \quad (7)$$

$$\phi_i^1 = \phi_i^0 - (C/4)[v_{i+1}^0(\phi_{i+1}^0 + \phi_i^0) - v_i^0(\phi_{i-1}^0 + \phi_i^0)] \quad (8)$$

where $\{v_i^0\}$ and $\{\phi_i^0\}$ are the initial conditions.

We now come to the derivation of adjoint equations. As was mentioned earlier, the main idea is to minimize the cost function J representing the misfit between the model and the data subject to the model equations (5)–(8) acting as constraints. For the time being, we do not specify the functional form of the cost function. Thus the approach to be outlined will be valid for arbitrary cost function. Only at the end of our discussion we will use the actual form of the cost function.

To begin with, in each of equations (5)–(8) we transfer all quantities to the left hand side and symbolically represent these equations by $E_i^n = 0$, $F_i^n = 0$ $(n = 2, 3, \ldots, N)$ and $E_i^1 = 0$, $F_i^1 = 0$. These are the constraints to be satisfied while minimizing J.

We thus form the customary Lagrangian by appending the constraints (multiplied by corresponding Lagrange multipliers) to the cost function J

$$L = J + \sum_{n=2}^{N} \sum_i (\lambda_i^n E_i^n + \mu_i^n F_i^n) + \sum_i (\lambda_i^1 E_i^1 + \mu_i^1 F_i^1). \quad (9)$$

It is well known that the stationary point of the Lagrangian coincides with the minimum of the cost function. Thus one has simply to find the partial derivatives of the Lagrangian with respect to all of its variables and set them to zero. The constrained problem has thus been reduced to an unconstrained one. It is elementary to see that by differentiating L with respect to the Lagrange multipliers ($\lambda - s$ and $\mu - s$) one recovers the original model equations (5)–(8). Differentiation with respect to the model variables v_i^n, ϕ_i^n etc. (for $n = 1, 2, \ldots, N$) leads to the equations

$$\lambda_i^N = -(\partial J/\partial v_i^N) \quad (10)$$

$$\mu_i^N = -(\partial J/\partial \phi_i^N) \quad (11)$$

$$\lambda_i^k = \lambda_i^{k+2} + 2A[v_{i+1}^k(\lambda_{i+1}^{k+1} - \lambda_i^{k+1}) + v_{i-1}^k(\lambda_i^{k+1} - \lambda_{i-1}^{k+1}) + v_i^k(\lambda_{i+1}^{k+1} - \lambda_{i-1}^{k+1})] + B(\lambda_{i+1}^{k+2} + \lambda_{i-1}^{k+2} - 2\lambda_i^{k+2}) + (C/2)[(\phi_i^k + \phi_{i-1}^k)(\mu_i^{k+1} - \mu_{i-1}^{k+1})] - (\partial J/\partial v_i^k) \quad (12)$$

$$\mu_i^k = \mu_i^{k+2} + C(\lambda_{i+1}^{k+1} - \lambda_i^{k+1}) + (C/2)[v_{i+1}^k(\mu_{i+1}^{k+1} - \mu_i^{k+1}) + v_i^k(\mu_i^{k+1} - \mu_{i-1}^{k+1})] - (\partial J/\partial \phi_i^k) \quad (13)$$

In equations (12) and (13) k ranges from 1 to $N-1$. In applying these equations for $k=N-1$, it has, however, to be kept in mind that $\lambda_i^{N+1}=\mu_i^{N+1}=0$ for all i. These are just dummy variables, introduced in order to keep the form of these equations same for all k, including $k=N-1$.

Finally, differentiating the Lagrangian with respect to the model initial conditions, we obtain

$$\begin{aligned}&-\lambda_i^1-\lambda_i^2-\mathrm{A}[v_{i+1}^0(\lambda_{i+1}^1-\lambda_i^1)+v_{i-1}^0(\lambda_i^1-\lambda_{i-1}^1)\\&\quad+v_i^0(\lambda_{i+1}^1-\lambda_{i-1}^1)]-\mathrm{B}(\lambda_{i+1}^2+\lambda_{i-1}^2-2\lambda_i^2)\\&-(\mathrm{C}/4)[(\phi_i^0+\phi_{i-1}^0)(\mu_i^1-\mu_{i-1}^1)]+(\partial J/\partial v_i^0)=0\end{aligned}\tag{14}$$

$$\begin{aligned}&-\mu_i^1-\mu_i^2-(\mathrm{C}/2)(\lambda_{i+1}^1-\lambda_i^1)-(\mathrm{C}/4)[v_{i+1}^0(\mu_{i+1}^1-\mu_i^1)+v_i^0(\mu_i^1-\mu_{i-1}^1)]\\&\quad+(\partial J/\partial\phi_i^0)=0\end{aligned}\tag{15}$$

The left sides of (14) and (15) represent the components of the gradient of the cost function with respect to the model initial conditions. Had we introduced dummy variables λ_i^0 and μ_i^0, the left sides would have been equal to these variables with their signs changed. This proves our earlier statement that the adjoint variables at the model initial time contain the required gradient informations. Equations (10)–(13) are generally referred to as the adjoint equations. They are so named because the corresponding partial differential equations can be shown to be adjoint to the linearized version of the model equations. Hence the equations themselves are linear in the adjoint variables λ and μ as can be readily verified. It can also be easily observed that integrations of the adjoint equations are carried out backward in time and that the equations are forced by model-data misfit represented by the partial derivatives of the cost function with respect to the model variables. If this misfit vanishes then the corresponding adjoint variables remain zero throughout the entire integration. It is then seen from equations (14) and (15) that their left hand sides vanish too and we are at the minimum of the cost function, as we should be.

In general, however, the model-data misfit terms do not vanish and the adjoint variables at the end of integration are non-zero. A minimization routine has thus to be used. The minimization routine employed by us was the ZXCGR routine of International Mathematical and Statistical Library (IMSL). This routine uses conjugate gradient algorithm (Navon and Legler 1987) for searching for the minimum. The algorithm starts from a guess value for the minimum. In our case, these are the guess values of initial conditions for starting the forward run of the basic model. With these guess values, the shallow water model is integrated forward in time using equations (5)–(8). The adjoint equations (10)–(13) are then integrated backward in time. At the end of adjoint integration we use (14) and (15) to obtain values of the components of gradient of the cost function. The gradient information is utilized by the routine to search for new initial conditions for starting a new iteration. Each iteration thus consists of a forward run of the original model, a backward run of the adjoint model followed by a search for a new set of initial conditions. Iterations are stopped when a specific convergence criterion is met. In figure 1 we present a flow chart of the variational data assimilation technique described above.

In our presentation so far, we have refrained from specifying the form of the cost function. Any concrete application of the adjoint technique requires, however, besides the basic model in which data are to be assimilated, a specific cost function representing misfit between the model results and the data. Following standard practice, we have

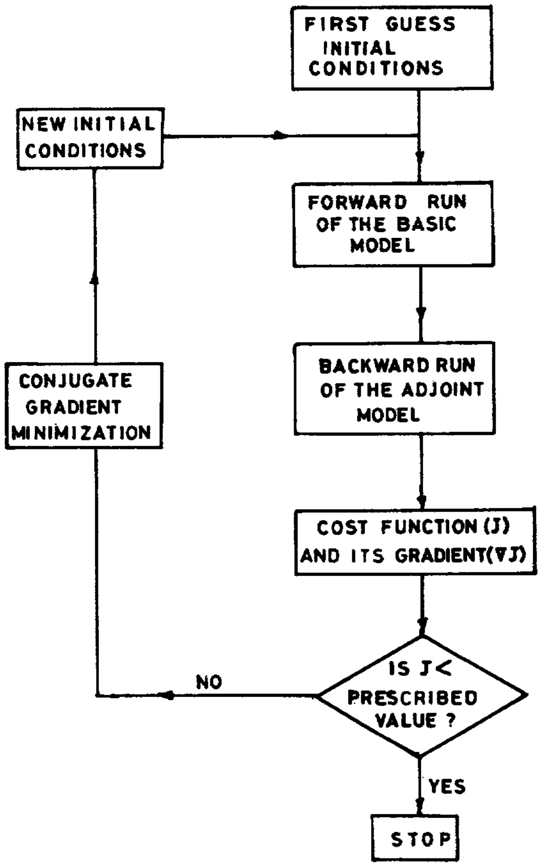

Figure 1. Flow chart of the variational data assimilation technique.

chosen a quadratic cost function given by

$$J = (1/2) \sum_{n=0}^{N} \sum_{i} [(v_i^n - vd_i^n)^2 C_i^n + (\phi_i^n - \phi d_i^n)^2 D_i^n], \tag{16}$$

where vd_i^n and ϕd_i^n are current and potential data at grid point i and at time n. Since there may not be observations at all grid points and at all times, we have introduced two reliability matrices C_i^n and D_i^n. The corresponding matrix elements will take on either of the two values 1 or 0, depending on availability or lack of observation at the corresponding space-time point.

In principle, one may assign to C_i^n and D_i^n, other values lying between 0 and 1 to signify the reliability of the data. In the present study, however, the data are generated from a control run of the model (identical twin experiment). Hence we have used only 0 or 1 for the values of the matrix elements.

5. Results

To generate the synthetic data set, we have used a control run of the model starting from the following initial conditions (written here in the analytical form, discretization being obvious)

$$v_0(x) = 1 + \sum_{j=1}^{2} a_j \sin(k_j x + \alpha_{v,j})$$

$$\phi_0(x) = 1 + \sum_{j=1}^{2} a_j \sin(k_j x + \alpha_{\phi,j})$$

where the wave amplitudes and wave numbers are given by $a_1 = 0{\cdot}15$, $a_2 = 0{\cdot}12$, $k_1 = 6\pi, k_2 = 4\pi$. The phase angles are $\alpha_{v,1} = \pi/2, \alpha_{v,2} = \pi/4, \alpha_{\phi,1} = 7\pi/6, \alpha_{\phi,2} = 13\pi/12$.

These initial conditions are represented in figures 2 and 3 by dots and are referred to as "true" initial conditions. In both these figures, the solid line indicates the initial conditions, retrieved by applying the adjoint technique. As a first guess initial condition, we took $v_0(x) = \phi_0(x) = 1$ identically for all x. We have tried other sets of

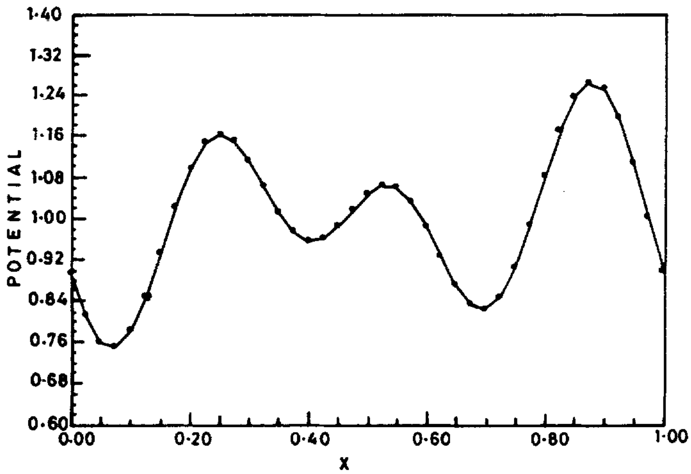

Figure 2. "True" initial conditions (dots) and retrieved solution for initial conditions (solid line) for the potential as a result of assimilating perfect potential data at the initial and final times.

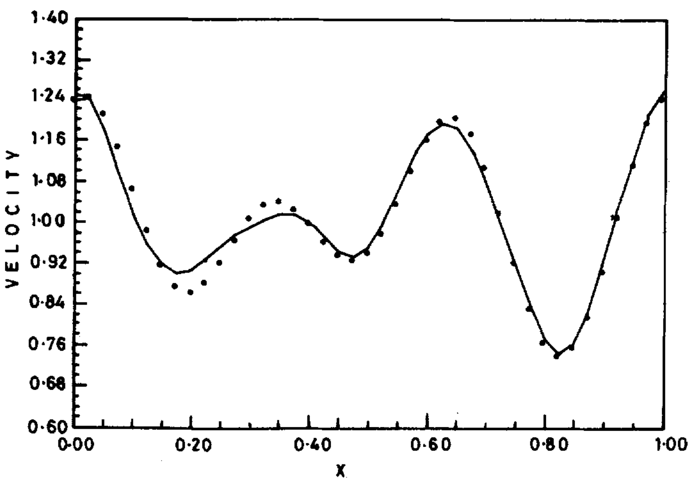

Figure 3. Same as in figure 2 except for velocity. No velocity data was assimilated.

guess values with means, not too different from unity. They also lead to the same retrieved set. Data assimilated in the model were taken from the result of the previous control run. However, only potential data at the initial and final time were assimilated. No current data was used. This was done to mimic the assimilation of altimeter data into ocean models. This is because one can calculate the instantaneous ocean depth (ϕ/g) from altimeter measurements, once the mean ocean depth is known. Thus the assimilation of altimeter sea level data is equivalent to assimilation of potential data in our model, although the similarity is quite crude, since our model is not a good representative of a realistic ocean model. Coming back to the assimilation we see that the number of observations is equal to the number of initial conditions to be retrieved. Thus, in principle, we should be able to retrieve the initial conditions unambiguously, unless there is redundancy in the data. It is seen from figure 2 that retrieval for potential is excellent (which is not surprising) while that for current is not so good.

In figures 4–7 we show the states of the model after 100 and 200 time steps starting

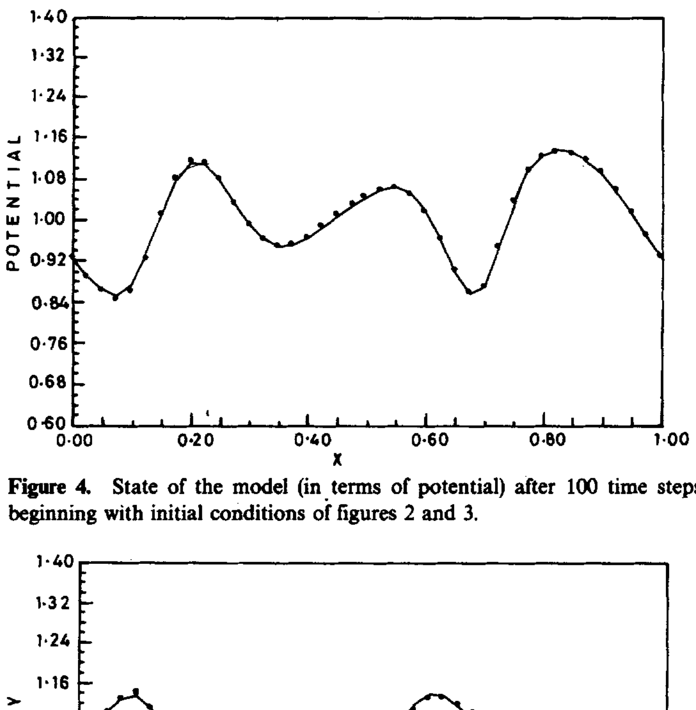

Figure 4. State of the model (in terms of potential) after 100 time steps of integration beginning with initial conditions of figures 2 and 3.

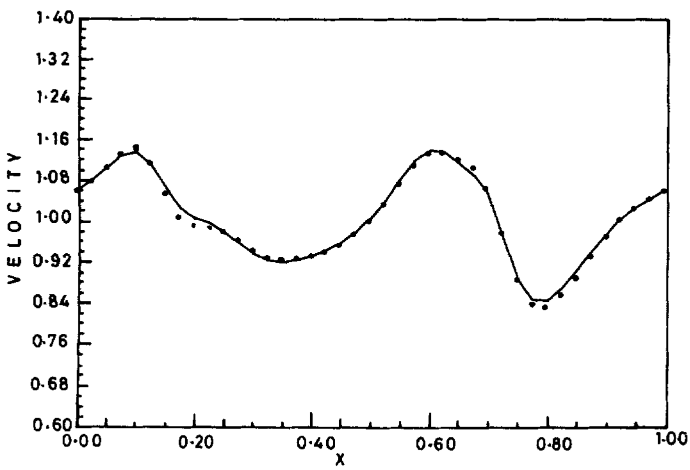

Figure 5. Same as in figure 4, except for velocity.

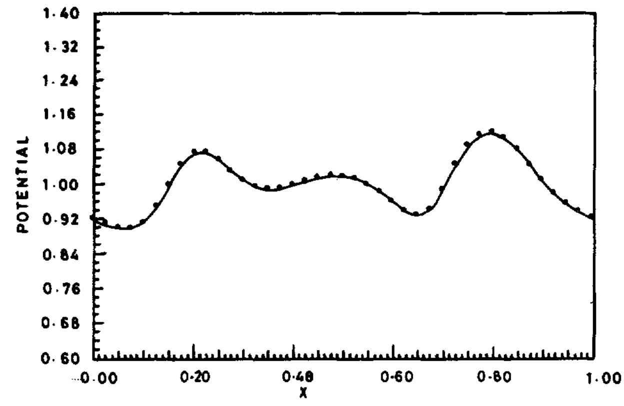

Figure 6. Same as in figure 4, but after 200 time steps.

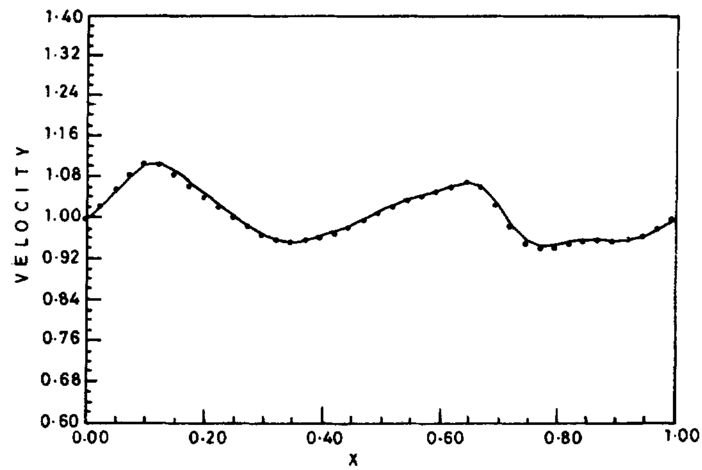

Figure 7. Same as in figure 5, but after 200 time steps.

from "true" initial conditions (the dots) as well as from the retrieved solution (the solid lines). It is seen that with the passage of time the agreement between the "true" current and retrieved current becomes better and better. This may be due to the dissipation present in the model. Inclusion of potential data at other times only improves the retrieval.

To get a feeling of the performance of the nonlinear optimization routine ZXCGR of the IMSL, in figure 8 we show the evolution of the cost function and the absolute value of the gradient as a function of the number of iterations. The functions have been scaled by their values at the zeroth iteration and have been expressed in logarithmic units. After about 20 iterations, the retrieval is quite satisfactory since the gradient has fallen below 10^{-4} times its original value and the cost function has been reduced by more than a factor of 10^{-3}. The convergence is fairly rapid within the first few iterations. The rate of reduction becomes slower afterwards.

We have so far studied the assimilation of perfect, i.e., noise-free data. This is seldom the case in practice. Observational data are inevitably noisy. To study this case, we simulate noisy data by adding white Gaussian noise to the "true" data.

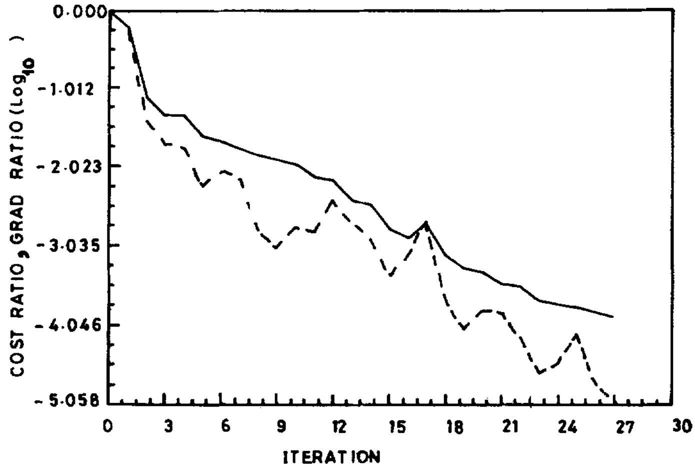

Figure 8. Cost reduction factor and gradient reduction factor versus number of iterations completed for the case studied. The solid line represents cost reduction factor and the dashed line represents reduction factor for the absolute value of the gradient.

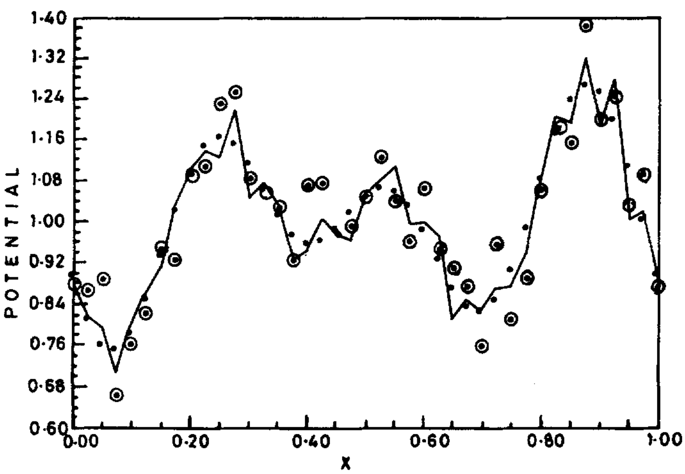

Figure 9. Same as in figure 2, except for assimilation of noisy potential data available at all space-time points. Noisy data are represented by encircled dots.

Noise variance is taken to be one-fourth of the variance of the initial fields. The noise is generated using IMSL Gaussian pseudorandom number generation routine.

In figures 9 and 10 we show the retrievals of initial potentials and velocities. Only noisy potential data were made available at all space time points. Velocity data were withheld. Measurement error for potential and residual error after assimilation (in the rms sense) were both found to be about 0·07. Thus the model fits the data to within measurement error. In figures 11 and 12 we show the state of the model at the final time. It is interesting to note that residual error after assimilation for the velocity was also found to be the same as the corresponding measurement error, although no velocity data were assimilated. We have also conducted assimilation by including noisy current data at all space-time points alongwith the potential data.

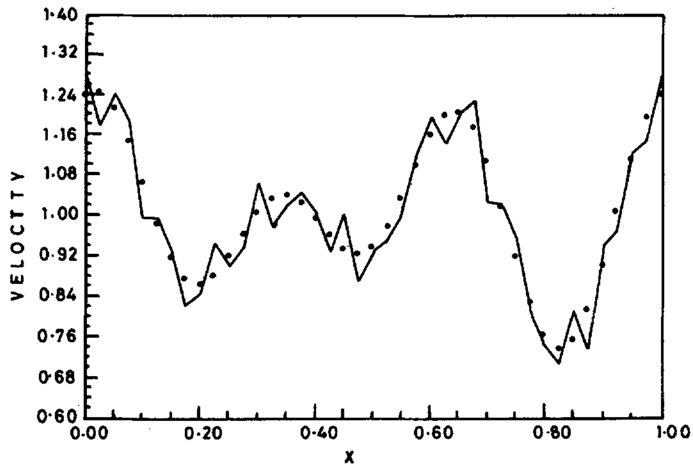

Figure 10. Same as in figure 9, except for velocity. It is to be stressed again that velocity data were withheld.

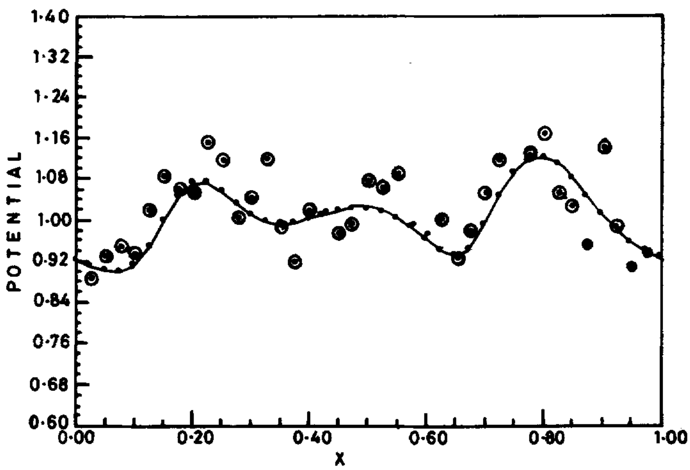

Figure 11. Same as in figure 6, but with initial conditions of figures 9 and 10.

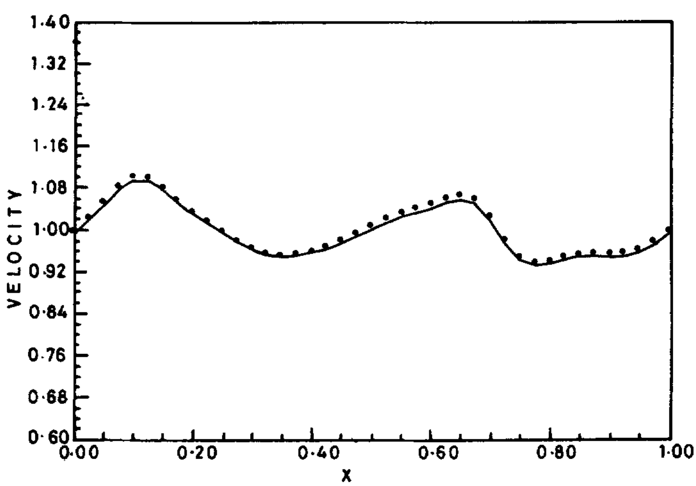

Figure 12. Same as in figure 11, except for velocity.

The residual error after assimilation is still of the order of measurement error. We are thus led to the conclusion that only potential measurements (in a sufficiently large number) are enough for retrieving the model state with accuracy, limited only by the accuracy of the data. By the coupling of the adjoint variables λ and μ, the adjoint equations force the model trajectory to remain close to the "true" trajectory, in the entire space-time domain.

To see the effect of the data reduction on the results of our data assimilation procedure we have conducted a number of experiments using a variety of data distribution. For the economy of space, we present only two examples. The first example is a repeat of the previous example, except for the fact that potential data are supplied not at all the grid points, but are confined to a part of the model domain from $x = 0{\cdot}25$ to $x = 0{\cdot}5$. In figure 13 we see that the initial retrieved field for the potential fits the measurement reasonably well in the measurement region. The same

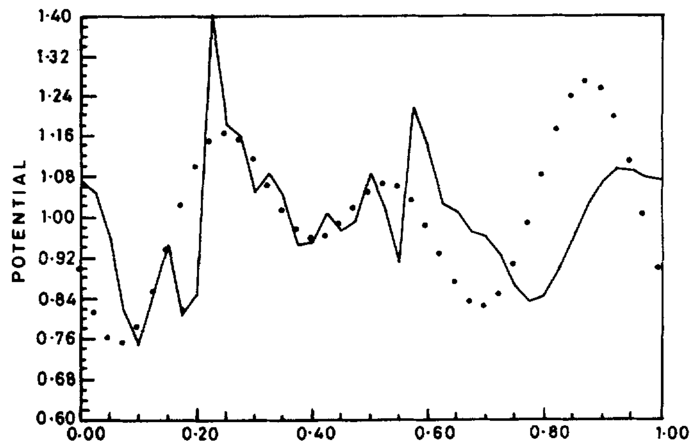

Figure 13. Same as in figure 9, but for the fact that noisy potential data are given only between $x = 0{\cdot}25$ and $x = 0{\cdot}5$ and for all time steps. Data are not shown. They are the same as those used in the previous experiment.

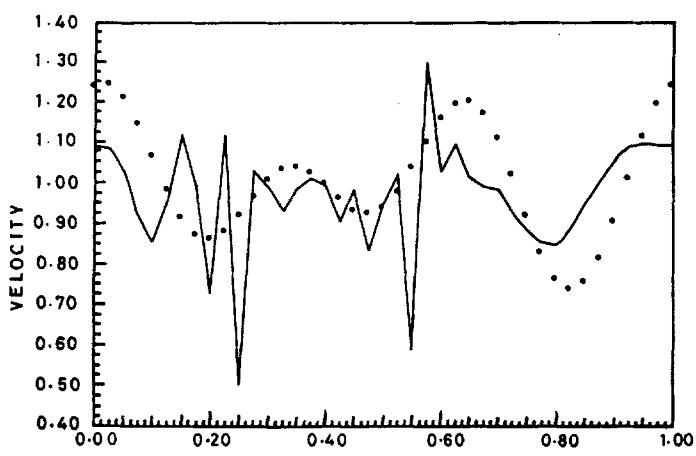

Figure 14. Same as in figure 13, except for velocity.

is true for the velocity (figure 14) although the fit is not as good. Outside the observation region, retrieved fields are quite arbitrary. However, even with these retrieved initial conditions, at the end of integration time, the fit for the potential is quite good in the region of measurements (figure 15). The solution for velocity is, however, very poor (figure 16). Thus the model trajectory, although close to the potential observations at all times, is not everywhere close to the "true" values. Thus the potential data in part of the domain turn out to be insufficient for a good reconstruction of the state of the model. We infer that some current observations are essential.

Our next example again involves continuous data in part domain. This time, however, current observations are made available at all time steps in the region from 0·25 to 0·5, while potential data are given at all grid points but for time steps $t = 190$ to $t = 200$ only. Figures 17 and 18 show the results of retrieving initial fields. Now the fit is seen to be good. We also show the potential and velocity fields at the end

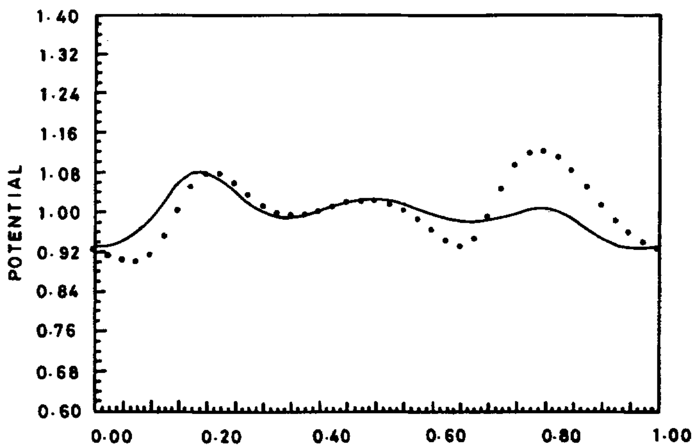

Figure 15. Same as in figure 11, but with initial conditions of figures 13 and 14.

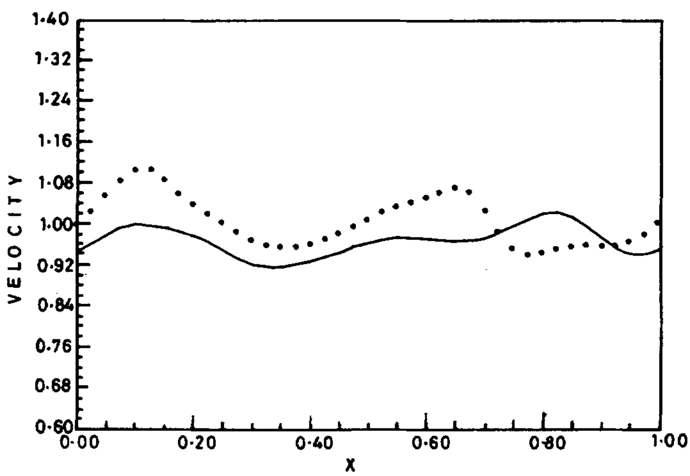

Figure 16. Same as in figure 15, except for velocity.

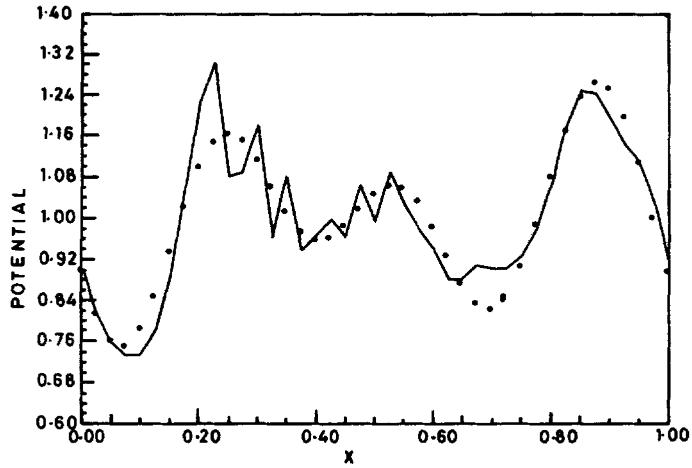

Figure 17. Same as in figure 13, but for assimilation of noisy velocity data between $x = 0{\cdot}25$ and $x = 0{\cdot}5$ at all time steps alongwith noisy potential data at all grid points, but for time steps from 190 to 200 only. Data set is again the same as used in the earlier experiments.

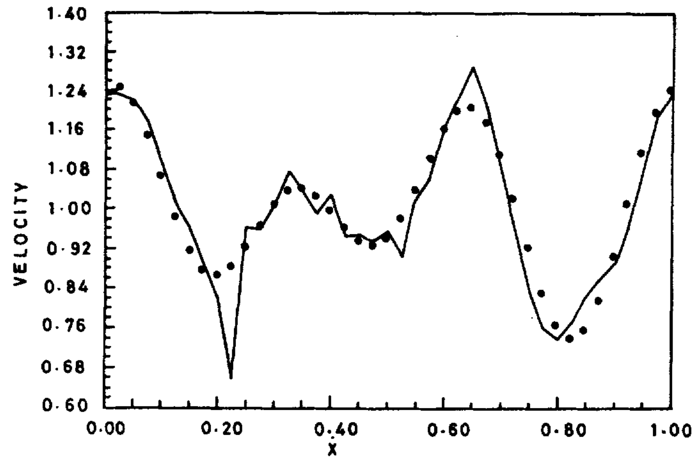

Figure 18. Same as in figure 17, except for velocity.

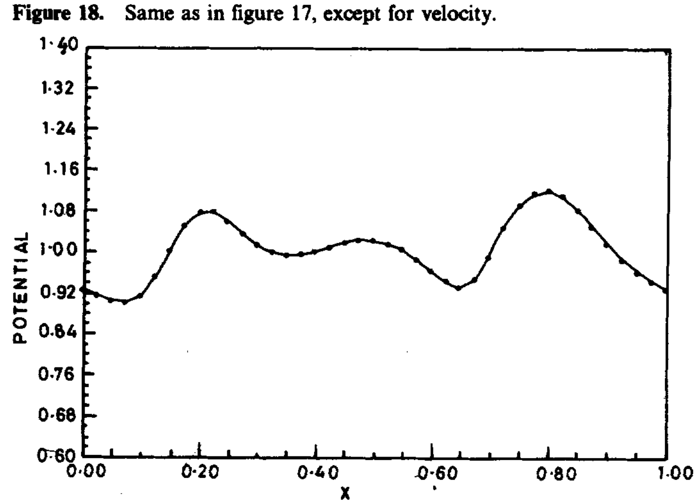

Figure 19. Same as in figure 15, but with initial conditions of figures 17 and 18.

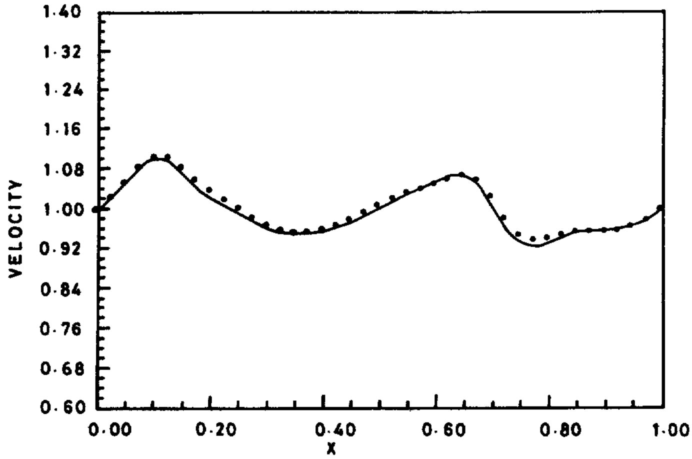

Figure 20. Same as in figure 19, except for velocity.

of integration (figures 19 and 20) to give an idea of how the solution matches the "true" trajectory.

6. Conclusions

We have described a data assimilation procedure using adjoint technique for assimilating observations into a dynamic forecast model for the purpose of improving predictive capability of the model with particular emphasis on finite difference formulation of the procedure. The method is a powerful tool particularly for assimilating asynoptic observations such as may be obtainable from sensors onboard the ocean observing satellites.

It was shown that the method is effectively equivalent to least squares fitting of the model computed variables to the data with model equations acting as hard constraints. A Lagrange multiplier approach transforms the problem of constrained optimization to an unconstrained one. The resulting large system of equations was solved by the adjoint approach coupled with a nonlinear optimization routine using conjugate gradient algorithm.

The method was illustrated in the example of a model based on one-dimensional shallow water equations in a cyclic domain. Hypothetical data consisting of potential (which can be obtained from satellite altimeter measurements provided undisturbed ocean depth is known) and, sometimes of current, were assimilated into the model.

Our results demonstrate that the state of the model can be retrieved with reasonable success (leading to improved predictive capability) by assimilating perfect potential data at the initial and final times. The same is true when noisy potential data are available at all space time points (in the sense that the model fits the data to within data error). However, potential data at a few locations are insufficient for a correct retrieval of the model state. An optimal measurement strategy seems to be that of combining continuous measurements of a variable (the current) at the same points alongwith quasisynoptic measurements of another variable (potential) at regular time intervals.

It is however to be borne in mind that the results are only illustrative as the model was a very simple one. An actual application will require the use of a more sophisticated ocean model. We have already begun the work of developing a three-dimensional equatorial ocean model to test our data assimilation technique and the results will be reported in the future. Nevertheless, the results of the experiments presented here seem to be quite encouraging and can surely make one confident that the adjoint technique is going to be an immensely powerful technique for extracting valuable information from a limited amount of, maybe asynoptic, oceanographic data.

References

Basu S and Pandey P C 1991 Numerical experiment with modelled return echo of a satellite altimeter from a rough ocean surface and a simple iterative algorithm for the estimation of significant wave height; *Proc. Indian Acad. Sci. (Earth Planet. Sci.)* **100** 155–163

Ghil M and Malanotte-Rizzoli P 1991 Data assimilation in meteorology and oceanography; *Adv. Geophys.* **33** 141–266

Holland W R and Malanotte-Rizzoli P 1989 Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies; *J. Phys. Oceanogr.* **19** 1507–1534

Le Dimet F X and Talagrand O 1986 Variational algorithms for analysis and assimilation of meteorological observations; *Tellus* **A38** 97–110

Lewis J M 1985 The use of adjoint equations to solve a variational adjustment problem with advective constraints; *Tellus* **A37** 309–322

Navon I M and Legler D 1987 Conjugate gradient methods for large scale minimization in meteorology; *Mon. Weather Rev.* **15** 1479–1502

Schröter J 1989 Driving of non-linear time dependent ocean models by observation of transient tracers – A problem of constrained optimization; In: *Ocean circulation models: Combining data and dynamics* (Dordrecht: Kluwer Academic Publishers) pp. 257–285

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

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