Variational Four-Dimensional Analysis Using Quasi-Geostrophic Constraints

JOHN C. DERBER*

Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison, Madison, WI 53706

(Manuscript received 13 November 1985, in final form 17 November 1986)

ABSTRACT

A variational four-dimensional analysis technique using quasi-geostrophic models as constraints is examined using gridded fields as data. The analysis method uses a standard iterative nonlinear minimization technique to find the solution to the constraining forecast model which best fits the data as measured by a predefined functional. The minimization algorithm uses the derivative of the functional with respect to each of the initial condition values. This derivative vector is found by inserting the weighted differences between the model solution and the inserted data into a backwards integrating adjoint model.

The four-dimensional analysis system was examined by applying it to fields created from a primitive equations model forecast and to fields created from satellite retrievals. The results show that the technique has several interesting characteristics not found in more traditional four-dimensional assimilation techniques. These features include a close fit of the model solution to the observations throughout the analysis interval and an insensitivity to the frequency of data insertion or the amount of data. The four-dimensional analysis technique is very versatile and can be extended to more complex problems with little theoretical difficulty.

1. Introduction

The complete and accurate specification of the four-dimensional structure of the atmospheric state is a basic problem in meteorology. The conventional observational network alone is not capable of providing this complete description of the atmospheric state, since the observational density is not high enough in either space or time. The development of new data sources such as satellite and ground-based atmospheric profilers may improve the situation but will not eliminate the problem. For this reason it is desirable to include additional information (e.g., statistical or dynamical information) in the four-dimensional analysis technique.

The most common technique for combining observations distributed in time is through the use of a four-dimensional data assimilation system. These assimilation systems introduce the observations into an integrating forecast model as the observation time is reached or at some intermittent interval. (See Bengtsson, 1975, and McPherson, 1975, for a review of these techniques and Bengtsson et al., 1981 for a collection of more recent work.) While such a system is useful for the initialization of forecast models in an operational environment, they do not provide a four-dimensional analysis which is consistent with the dynamics of the forecast model. To find a four-dimensional analysis which fits the data as well as possible and satisfies the dynamics of the forecast model, variational techniques can be used. The variational problem is to minimize some functional, which measures the fit of the four-dimensional analysis to the data, while using the forecast model as a strong constraint.

The traditional meteorological method for solving the variational problem involves the direct solution of the associated Euler–Lagrange equations for each grid point in space and time, using the model as a strong constraint (Sasaki, 1958, 1970). However, this method of solution requires either prohibitively large computing resources, a simple forecast model (e.g., Thompson, 1969; Lewis, 1982) or approximations in the solution (e.g., Bloom, 1983). If one notes that the specification of the initial conditions to the constraining forecast model completely specifies the solution in time, the magnitude of the problem can be reduced considerably by adjusting only the initial fields. Hoffman (1986) has used this idea to fit a very simple model variationally. However, his method is still too inefficient to be usable on a realistic model. LeDimet and Talagrand (1986) suggested a method based on optimal control theory (Lions, 1971) to perform the variational four-dimensional analysis adjusting only the initial fields. This approach was employed by Lewis and Derber (1985) with simple two-dimensional constraints and will be used here with much more complicated four-dimensional quasi-geostrophic constraints.

Since the forecast model is applied as a strong constraint, the quality of the final four-dimensional analysis will depend not only on the amount and error structure of the data, but also the accuracy of the forecast model. The effect of the modeling errors will be

* Present affiliation: Geophysical Fluid Dynamics Program, Princeton University, Princeton, NJ 08542.

examined by inserting analyses created from a primitive equations model forecast into the four-dimensional analysis technique using a quasi-geostrophic model as a constraint (section 3). The effect of the error structure on the final four-dimensional analysis will be demonstrated by inserting analyses created from satellite retrievals which appear to contain temporally correlated errors (section 4). The results show that the variational four-dimensional analysis technique is well behaved and is able to produce a reasonably accurate representation of the atmospheric state despite the approximations in the dynamics of the constraining models. However, the four-dimensional analysis technique cannot remove temporally correlated errors. Also, while the computational costs are high for this method, they are substantially less than other variational analysis methods. It appears to be feasible to apply this method on realistic models in the research mode with current computing resources. With the next generation of computers, it should be feasible for operational applications.

2. The variational four-dimensional analysis technique

The variational four-dimensional analysis technique employed in this paper is essentially the same as that used in Lewis and Derber (1985) except it is applied with more complicated and realistic constraints. The objective is to find the solution to the constraining forecast model which will fit a series of fields best over some time interval. This fit is defined through a scalar known as the functional and given in this case by

$$I=\frac{1}{2}\sum_{p=1}^{P}\langle(\Psi(t_p)-\tilde{\Psi}(t_p)),\mathbf{W}(t_p)(\Psi(t_p)-\tilde{\Psi}(t_p))\rangle, \quad (1)$$

where P is the number of observed static fields available in the assimilation interval, $\Psi(t_p)$ is an N-component vector containing the forecast geostrophic streamfunction values defined over all interior grid points at time t_p, $\tilde{\Psi}(t_p)$ is a static observed N-component geostrophic streamfunction field at time t_p, and $\langle a, b\rangle$ denotes the inner product over all interior grid points, given by

$$\sum_{n=1}^{N} a_n b_n.$$

It is assumed that the observed fields are ordered such that $t_1 \leqslant t_2 \leqslant \cdots \leqslant t_P$. The $\mathbf{W}(t_p)$ matrices are $N \times N$, positive-definite, symmetric matrices which can be used to incorporate information on the spatial quality of the observed fields.

The time evolution of the geostrophic streamfunction values ($\Psi(t)$) is constrained to satisfy a forecast model. Two different constraining forecast models will be used. In section 3, a five-level quasi-geostrophic model will be used to evaluate the assimilation technique. In section 4, a ten-level quasi-geostrophic model will be applied to fields created from temperature profiles which are, in turn, retrieved from satellite-measured radiances.

Both of the constraining forecast models are second-order finite difference approximations to the differential equation

$$\frac{\partial q}{\partial t}=-m^2J(\Psi,q)-m^2\beta\frac{\partial\Psi}{\partial x} \quad (2)$$

where

$$q=m^2\nabla^2\Psi+\frac{\partial}{\partial p}\frac{f_0^2}{\sigma}\frac{\partial\Psi}{\partial p},$$

β is the north–south derivative of the Coriolis parameter, σ is the mean static stability, m is a map factor and the Jacobian, $J(a, b)$, is defined as

$$\frac{\partial a\partial b}{\partial x\partial y}-\frac{\partial b\partial a}{\partial x\partial y}.$$

The equations are integrated on a 23×28 Lambert conformal grid using an iterative implicit time integration scheme. In the vertical, the five-level version has equally spaced levels at 100, 300, 500, 700 and 900 mb. The vertical levels in the ten-level version are not equally spaced, using the mandatory levels at 100, 150, 200, 250, 300, 400, 500, 700, 850 and 1000 mb. The vertical boundary conditions force ω ($=\partial p/\partial t$) to vanish at $p = 0$ and $p = 1000$ mb and do not include orographic effects. In the horizontal, the geostrophic streamfunction is specified for two bounding grid points on each edge. Further details on these models can be found in Derber (1985).

To achieve the objective of minimizing the functional, many different standard nonlinear optimization algorithms could be used (Polak, 1971, or Gill et al., 1981). For this study, a conjugate gradient algorithm was chosen since it represents a good compromise in convergence rates and computer memory requirements between simpler and more complex methods. Simpler methods, such as steepest descent algorithms, will usually converge slower than conjugate gradient algorithms. More complex methods, such as quasi-Newton methods, usually converge faster than a conjugate gradient algorithm but they require storage of $N \times N$ matrices while the conjugate gradient algorithms only require storing vectors of length N, where N is the number of initial values to the model. For large forecast models, N easily can be greater than 10^5, thus it may be impractical to use the more complex, faster converging methods.

The conjugate gradient algorithm is an iterative method which produces a better approximation to the optimal solution with each iteration. Within an iteration, an estimate is made of the best way to change each grid point of the initial conditions to produce the maximum reduction of the functional. This is done by first finding a search direction. This search direction is an estimate of the relative change to each initial grid point value to produce the maximum reduction in the functional. Unfortunately, the search direction only gives the relative changes to the initial conditions, not

the absolute magnitude. To find the absolute changes, the optimal stepsize must be estimated. This optimal stepsize is a positive scalar which specifies how much the initial conditions should be changed in the search direction to produce the maximum reduction in the functional. Thus, the revised initial conditions after each iteration of the conjugate gradient algorithm are given by the old initial conditions plus the stepsize times the search direction.

The most difficult part of applying this algorithm is the determination of the search direction. Many different conjugate gradient algorithms have been proposed with slight differences in the search directions (Polak, 1971). For this study, a search direction which attempts to account for errors in the estimate of the optimal stepsize was chosen (Shanno, 1978). Note that this type of conjugate gradient algorithm is sometimes referred to as a one step, limited-memory, quasi-Newton method. The determination of this search direction requires knowledge of the derivative of the functional with respect to each of the initial condition values (the gradient of the functional). This gradient is combined with information from previous iterations, according to Eq. 20 of Shanno (1978), to produce the search direction.

To find the gradient of the functional, the adjoint equation technique introduced to meteorological applications by LeDimet and Talagrand (1986) will be used. This technique involves the calculation of the gradient of the functional through the use of an adjoint model. The adjoint model is defined by inserting the model forecast equations into the functional to eliminate all unknowns except the initial conditions. The functional is then differentiated with respect to the initial conditions and the results algebraically manipulated to create the adjoint model. Further details can be found in Lewis and Derber (1985) or LeDimet and Talagrand (1986).

For the functional and quasi-geostrophic model given by (1) and (2), the adjoint model equations are an approximation to the differential equation given by

$$\frac{\partial \xi}{\partial t} = \left[m^2 \nabla^2 + \frac{\partial}{\partial p}\left(\frac{f_0^2}{\sigma} \frac{\partial}{\partial p} \right) \right] J(\Psi, s) + J(s, q) - \beta \frac{\partial s}{\partial x} \quad (3)$$

where s is found by solving the three dimensional Helmholtz equation given by

$$\xi = \left(m^2 \nabla^2 + \frac{\partial}{\partial p} \frac{f_0^2}{\sigma} \frac{\partial}{\partial p} \right)(s/m^2),$$

Ψ and q are values from the previous quasi-geostrophic forecast integration, and ξ is the adjoint variable. The gradient of the functional is found by first initializing at time t_P to the weighted difference between the quasi-geostrophic forecast and the observed field [i.e., $\xi(t_P) = \mathbf{W}(t_P)(\Psi(t_P) - \tilde{\Psi}(t_P)]$. The adjoint variable at earlier timesteps is then found by integrating the adjoint Eq. (3) backwards in time and adding the weighted difference between the forecast and observed fields $[\mathbf{W}(t_p)(\Psi(t_p) - \tilde{\Psi}(t_p)]$, as each observation time (t_p) is reached in the backwards integration. The result at the initial time $[\xi(t_1)]$ is then equal to the exact gradient of the functional.

Note that it is essential to use the finite difference form of the adjoint equations which is derived from the finite difference form of the forecast equation rather than an arbitrarily chosen finite difference form of Eq. (3). Otherwise, differences between the finite differencing may result in the failure of the minimization algorithm. Also note that the operations required in integrating the adjoint model and the forecast model are very similar. Thus, the integration of the adjoint model from t_P to t_1 requires about the same amount of computation as an integration of the forecast model from t_1 to t_P.

The optimal stepsize is estimated by assuming the forecast solutions vary linearly in the search direction, where the linear change is based on the difference between the current solution and a solution produced from an initial field incremented in the search direction by some guess stepsize (ρ_g). The estimate of the optimal stepsize is then given by

$$\rho = \frac{\rho_g \sum_{p=1}^{P} \langle \mathbf{W}(t_p)(\Psi(t_p) - \tilde{\Psi}(t_p)), \Psi'(t_p) \rangle}{\sum_{p=1}^{P} \langle \mathbf{W}(t_p)\Psi'(t_p), \Psi'(t_p) \rangle} \quad (4)$$

where $\Psi'(t_p)$ is the difference between the current solution and the guess solution at analysis time t_p. For all reasonable guess stepsizes (ρ_g), this simple estimate of the stepsize has worked extremely well for the quasi-geostrophic constraint. However, if nonlinearities become *more* important in a more complex model, a more sophisticated estimate may be necessary.

Using these procedures, each iteration of the algorithm to minimize the functional can be completely defined:

1) Integrate the forecast model from t_1 to t_P using the current best estimate of the optimal initial conditions.

2) Find the gradient of the functional with respect to the initial conditions by adding the weighted difference between the current solution from step 1 and the observed streamfunction field, $[\mathbf{W}(t_p)(\Psi(t_p) - \tilde{\Psi}(t_p))]$, to the adjoint variable ξ as each analysis time (t_p) is reached in the integration of the adjoint model from t_p to t_1.

3) Calculate the search direction using equation 20, Shanno (1978).

4) Estimate the optimal stepsize using (4).

5) Update the initial conditions by adding the optimal stepsize times the search direction to the current estimate of the initial conditions.

6) Check to see if the convergence criterion is sat-

isfied. For this study, the four-dimensional analysis was assumed to have converged when the maximum value of the gradient over the entire grid was less than g/f_0. This results in the maximum change in the initial streamfunction field which corresponds to a change in the initial height field of less than (and often much less than) one-half meter.

The minimization algorithm is not inexpensive to run, requiring approximately three forecast model integrations per iteration (one to determine the solution, one to determine the stepsize, and one integration of the adjoint model to determine the gradient of the functional). This is one of the reasons for the choice of a simpler quasi-geostrophic model over a p.e. model. There is no theoretical difficulty in using a more complex forecast model (e.g., a primitive equations model), if the solution remains differentiable with respect to the initial conditions, but in practical terms the storage requirements and computational expense would be very high. The demand on computational resources has also forced us to limit the assimilation intervals in the following two sections to be less than or equal to 12 hours.

3. The four-dimensional analysis technique applied to fields created from a primitive equations model

Fields created from a primitive equations (p.e.) model forecast were inserted into the variational four-dimensional analysis technique. By treating these p.e. height fields as truth, the results from the four-dimensional analysis can then be used to evaluate the performance of the four-dimensional analysis technique and the effects of using an imperfect forecast model on the solution. The p.e. height fields were created using the subsynoptic scale model described in Mills and Hayden (1983). This model was run with a horizontal resolution of 67.6 km and ten vertical sigma levels for 15 hours after being initialized at 1200 UTC 26 April 1982. Only the last 12 hours of this forecast were used to produce height fields since the amplitude of the nonphysical gravity modes tend to be much larger in the first 3 hours of the forecast. During this 12 hour period, the eastern half of the forecast region is dominated by an eastward moving and slowly decaying upper air trough, associated with a surface low which decays and moves east in the first 6 hours but becomes stationary and reintensifies in the last 6 hours. An upper air ridge exists over the region's western half, along with a strong, nearly unchanging surface high over South Dakota. The convention of referring to the time as the number of hours after 1500 UTC 26 April will be employed in this section. Thus, the 0, 6 and 12 hour forecasts are from 1500 and 2100 UTC 26 April and 0300 UTC 27 April, respectively.

The forecast height field at each hour of this twelve hour period was then interpolated from the p.e. model's coordinate system to the quasi-geostrophic model's coordinate system. This interpolation involved halving the horizontal resolution (to 135.2 km) and vertically interpolating the heights to the pressure levels of the five and ten level models. For those points below the p.e. model topography, a standard atmospheric lapse rate of 6.5°C km^{-1} was assumed to extrapolate the height fields. These p.e. fields are used in this section to specify the horizontal boundary conditions, as input observation fields and as verification fields.

Assuming equal weighting at each grid point [$\mathbf{W}(t_p) = \mathbf{I}$, the identity matrix], the fields were inserted into the four-dimensional analysis technique over two intervals, the first 6 hours and the complete 12 hours of the analysis period. Within each analysis interval, the fields were inserted at various frequencies. In the 12 hour interval, the fields were inserted every 12 hours (two insertion times), every 6 hours (three insertion times), every 3 hours (five insertion times), and every hour (13 insertion times). For the 6 hour assimilation, the fields were inserted at 6 hour, 3 hour and 1 hour periods using two, three and seven insertion times, respectively.

For the insertion of data at 0 and 12 hours, the root-mean-square (rms) height differences between the p.e. fields and the five-level quasi-geostrophic model forecasts are shown in Fig. 1 for the unadjusted forecast, the first five iterations and the converged forecast (iteration 17). The four-dimensional analysis can be seen to be converging smoothly to the solution with most of the adjustment occurring in the first five iterations. Also, note that despite the fact that information was inserted only at 0 and 12 hours, the quasi-geostrophic

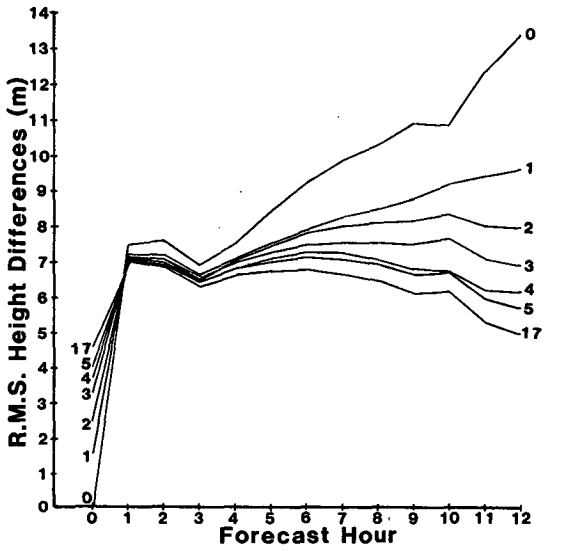

FIG. 1. Rms height differences between four-dimensional analysis and p.e. height fields for first five iterations of the analysis technique using two input height fields twelve hours apart. Also shown are the rms differences for the unadjusted forecast (iteration 0) and the converged analysis (iteration 17).

forecast after the four-dimensional analysis fits the p.e. fields better than the forecast from the unadjusted initial conditions at all times between 1 and 12 hours. This indicates that the method is finding a solution to the quasi-geostrophic model which approximates the evolution of the p.e. model solution over the entire assimilation interval rather than finding a solution which is physically unrealistic but fits the analyses well at 0 and 12 hours.

At the initial time, the rms difference between the quasi-geostrophic initial conditions and the p.e. fields must be increased since the initial conditions in the first iteration are set equal to the p.e. fields. The maximum in the rms differences found at 1 and 2 hours appear to be primarily due to the presence of noise in the p.e. forecast resulting from the initialization. By 3 hours into the interval, this noise has become damped sufficiently to become unimportant.

The convergence rate of the conjugate gradient algorithm is governed by the separation of the eigenvalues of the functional's second derivative matrix (the Hessian matrix). The more clustered the eigenvalues, the faster the convergence (Gill et al., 1981). The Hessian matrix, $\mathbf{H}$, for any iteration of the functional represented by equation (1) is given by

$$\mathbf{H} = \sum_{p=1}^{P} \mathbf{D}_p^{\mathrm{T}} \mathbf{W}(t_p) \mathbf{D}_p \tag{5}$$

where $\mathbf{D}_p$ is a matrix representing the integration of the forecast model from t_1 to t_p [i.e., $\Psi(t_p) = \mathbf{D}_p \Psi(t_1)$]. For a nonlinear model, the $\mathbf{D}_p$ matrix is formed by linearizing around the current solution. If the observed fields are equally weighted ($\mathbf{W}(t_p) = \mathbf{I}$), the expression for the Hessian reduces to

$$\mathbf{H} = \sum_{p=1}^{P} \mathbf{D}_p^{\mathrm{T}} \mathbf{D}_p. \tag{6}$$

Since the fields are equally weighted in this section, the clustering of the eigenvalues of this matrix will control the convergence rate. When unequal weighting is included in the next section, the convergence rate is reduced substantially.

When $\mathbf{D}_p^{\mathrm{T}} \mathbf{D}_p = \mathbf{I}$ for all p, then the Hessian matrix will be proportional to the identity matrix and the algorithm will converge in a single iteration. This situation occurs in the limiting case when all the observed fields are at the initial time. The $\mathbf{D}_p^{\mathrm{T}} \mathbf{D}_p$ matrix will usually deviate further from the identity matrix as the analysis interval becomes longer. Thus one expects the convergence rate to become slower as the analysis interval becomes longer.

This can be seen to occur by examining the evolution of the magnitude of the gradient vector. Since the magnitude of the gradient decays in an exponential fashion, the normalized magnitude of the gradient is displayed logarithmically in Fig. 2. The values shown in Fig. 2 are normalized by the initial magnitude of the gradient to allow direct comparison of the convergence rates. As expected, the convergence rate is dependent on the length of the analysis interval, with the 6-hour interval version converging much more rapidly than that using the twelve hour intervals after the first iteration. However, for analysis intervals of the same length, the magnitude of the gradient vector decays in approximately the same exponential fashion. This implies that the rate at which the four-dimensional analysis technique converges to the solution is controlled by the maximum separation of the insertion times and is nearly independent of the number of analyses inserted into the four-dimensional analysis technique. Also, examination of the spatial characteristics of the gradient vectors shows that the maximum values of the gradient are located near the boundaries and that these maximum values tend to alternate in sign between iterations. This indicates that the boundary conditions have a detrimental effect on the convergence rate through distortions of the $\mathbf{D}_p$ matrix near the boundaries.

While the inclusion of additional information in the assimilation interval does not appear greatly to affect the convergence rate, it does produce solutions that fit the p.e. fields better. Figure 3 displays the rms differences between the p.e. fields and the quasi-geostrophic forecasts after applying the assimilation technique to the analyses. In all cases, the adjusted quasi-geostrophic forecasts fit the p.e. forecast much better than the unadjusted forecast. As could be expected, the best fit occurred for shorter intervals, where the effects of modeling differences are smaller, and for larger quantities of inserted data.

In Fig. 3 it is also apparent that when there are three or more insertion times, a minimum in the rms values occurs near the center of the analysis interval. This feature can be explained by considering a simple linear forecast model with one dimensionless forecast variable. The forecast model will be assumed to be incorrect, overestimating the variable by one unit every hour. When equally spaced observations over a twelve hour interval are inserted into the four-dimensional analysis technique, the best fit solution will have an error that grows linearly from -6 at the initial time to 6 at the end of the 12 hour interval. This results in an rms curve which decreases from maxima of six at the initial and final times to a minimum of 0 at 6 hours, the center of the interval. In the comparison of the fit of the quasi-geostrophic model solution to the p.e. forecast, the situation is more complicated since the dynamics are nonlinear and the differences are dependent on the solution itself. For these reasons, the minimum does not reach zero at the center of the interval and the evolution of the errors cannot be represented by simple linear or quadratic curves.

From the simple one dimensional example just presented, one may also expect the spatial difference fields between the p.e. height fields and the optimal quasi-geostrophic solution to show smaller errors in the center

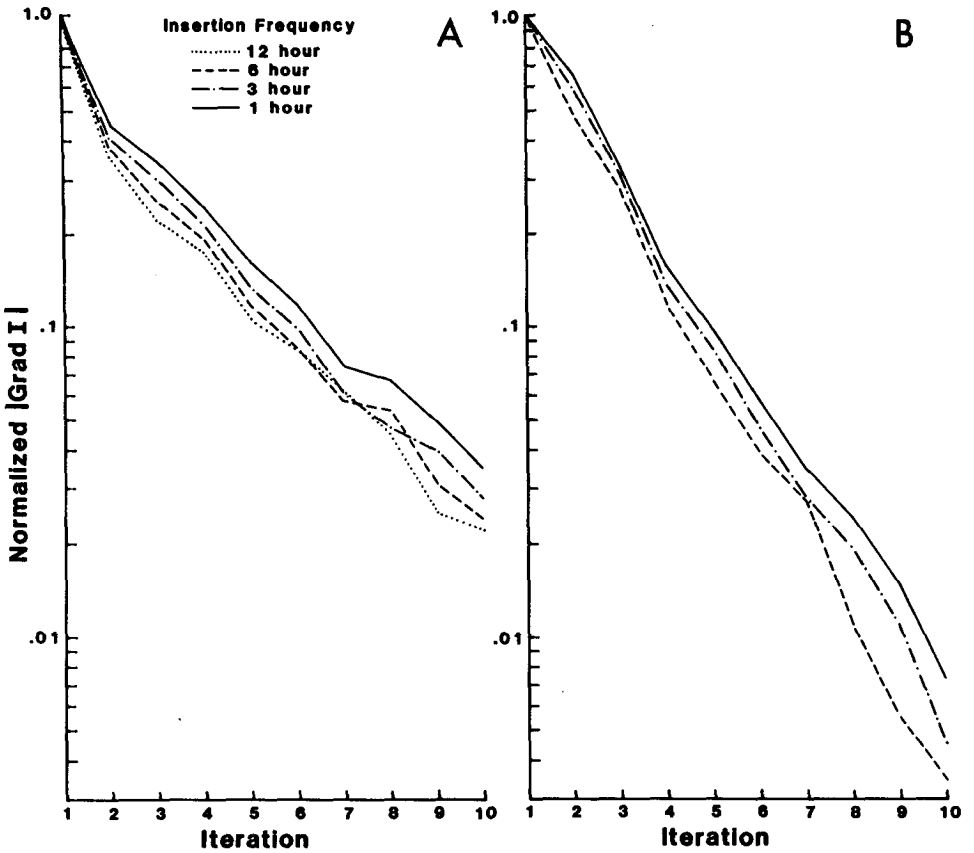

FIG. 2. Normalized magnitude of the gradient of the functional (Grad I) for the first ten iterations using a twelve hour analysis interval (panel A) and a six hour analysis interval (B). The magnitudes of the gradient are normalized by the initial magnitude of the gradient.

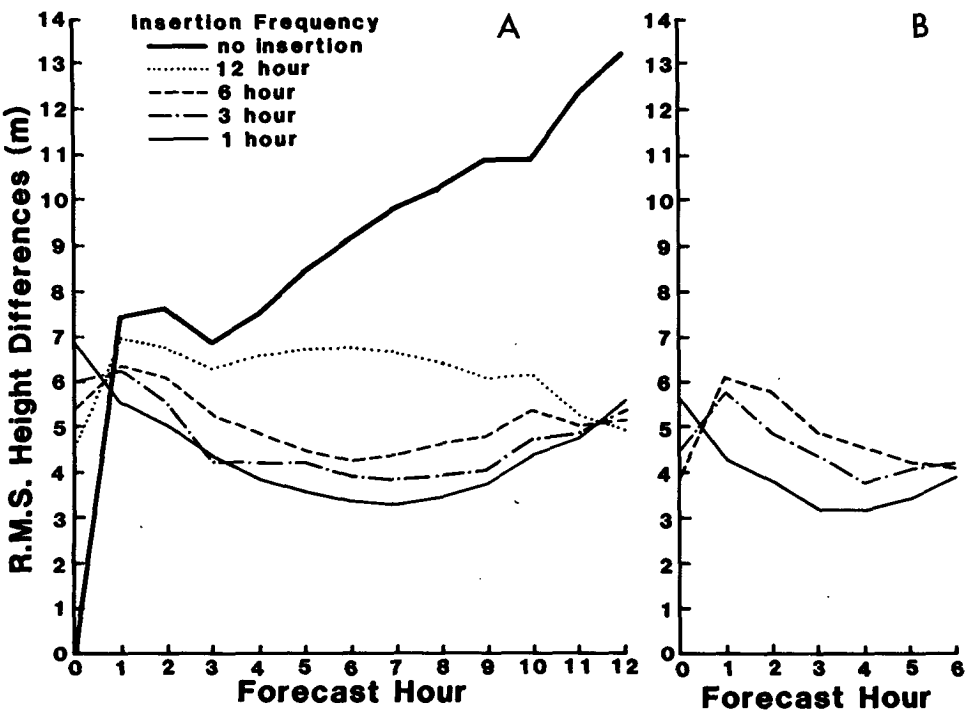

FIG. 3. Rms height differences between analysed height fields and p.e. height fields for a 12 hour analysis interval (panel A) and a 6 hour analysis interval (B). Also shown in panel A are the rms differences between the p.e. height fields and the quasi-geostrophic forecast produced from the unadjusted initial conditions (thick solid line).

of the analysis interval and a tendency to have errors of opposite signs at 0 and 12 hours of a 12 hour analysis. Figure 4 displays the 500 and 900 mb difference fields between the p.e. height fields and the adjusted quasi-geostrophic forecast height fields at 0, 6 and 12 hours for the 12 hour analysis using data every hour. As expected, while the differences are relatively small at all times, they are significantly larger at 0 and 12 hours than in the 6 hour forecast especially at the 900 mb level. There is also some tendency for the relatively large negative values in the western half of the 0 hour 900 mb difference field to balance the positive values in the 12 hour difference fields. However, the relationship is certainly not simple in this case. Finally, note that the differences are much larger at 900 mb than at 500 mb. This is probably due to the lack of inclusion of surface effects in the quasi-geostrophic model and errors from extrapolating down to the 900 mb level in mountainous regions.

4. The assimilation of analyses created from satellite retrieved temperature profiles

Height fields created from temperature profiles retrieved from satellite measured radiances also have been inserted in the ten-level version of the four-dimensional analysis routine. These fields were inserted both with equal weighting for each point and with weighting dependent on the vertical level and observational data density. Satellite data were used since these are the only data which can be produced with high enough temporal resolution at the proper spatial scales. However, as will be shown below, the satellite data may contain spatially correlated errors which could result in errors temporally correlated to the model solution. This type of error cannot be removed by this four-dimensional analysis technique or any other purely dynamical four-dimensional analysis method.

The data assimilation method was applied to three

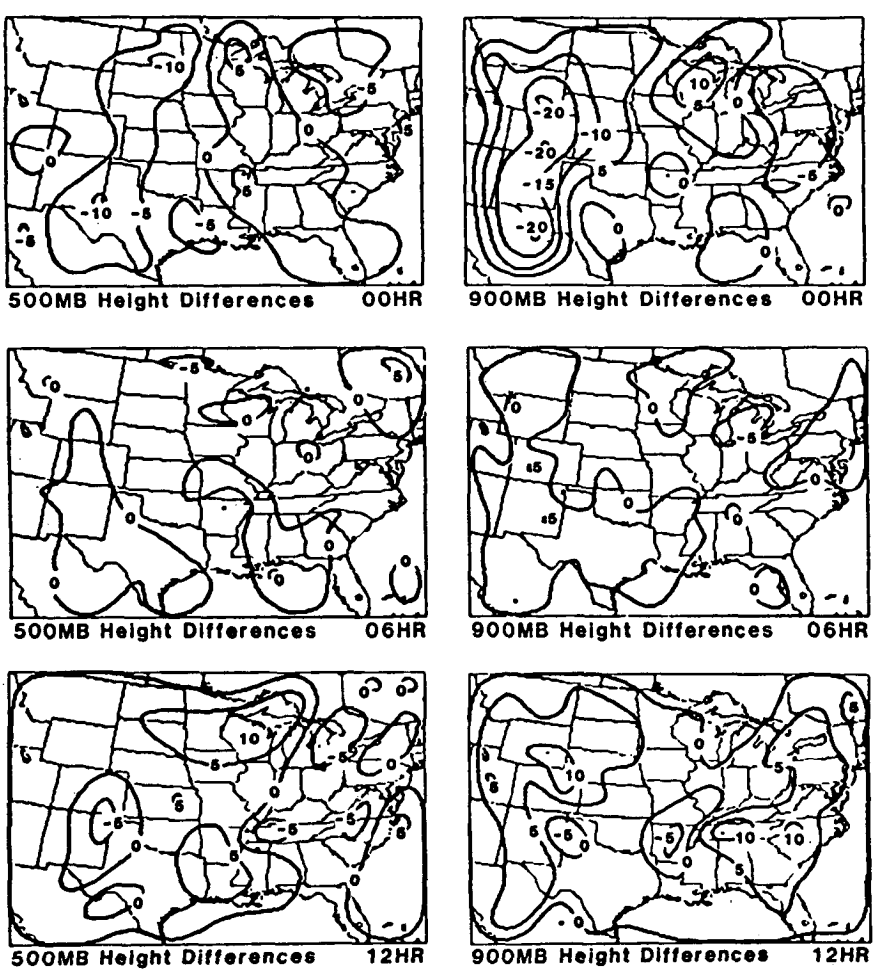

FIG. 4. 500 and 900 mb height difference fields (m) between converged four-dimensional analysis and p.e. height fields at 0, 6 and 12 hours into interval. The analysis results were created by inserting data every hour over the 12 hour interval.

satellite derived height fields at 1100, 1300 and 1400 UTC 27 June 1984. The analyses were produced from temperature profiles retrieved using an iterative technique developed by Smith (1983) from radiances measured by the VISSR atmospheric sounder (VAS) onboard a geostationary satellite. These retrieved profiles were screened manually and checked for consistency between the soundings to eliminate those soundings which are severely contaminated by cloud elements. Then the soundings along with surface hourly observations were inserted in the static objective analysis routine described in Mills and Hayden (1983). Since no wind observations were included in creating the satellite height fields, this static objective analysis technique essentially becomes a successive corrections technique similar to that developed by Cressman (1959) applied to the thicknesses implied by the temperature profiles. The first guess for the static objective analysis was a 1200 UTC 27 June radiosonde analysis. Thus all differences between the input height fields result from the retrieved temperature profiles.

The 500 mb height fields produced from the satellite data are shown in Fig. 5. The evolution of the 500 mb height field subjectively appears reasonably consistent in time. This consistency is confirmed by the small adjustments which occurred in the fields when the four-dimensional analysis technique was applied assuming equal weighting at each grid point [i.e., $\mathbf{W}(t_p) = \mathbf{I}$, the identity matrix, for all t_p]. The rms adjustments for each level and time are shown in Table 1.

Since the analysis technique creates a four-dimensional analysis, the output 1200 UTC fields can be compared directly to a conventional analysis based on rawinsonde observations. The 500 mb height fields for the conventional and the four-dimensional analysis, assuming equal weighting, are shown in Fig. 6. The magnitude of the differences between these two fields is indicative of the differences at other levels. The largest of these differences is the presence of a short-wave trough over the southwestern United States in the satellite analysis which is not contained in the conventional analysis. This difference is believed to have resulted from the contamination of the satellite retrievals by cirrus clouds, but the source of the difference is unimportant for this paper. However, since the feature is persistent in the input satellite height fields, the four-dimensional analysis is unable to recognize that the feature is incorrect (assuming the conventional analysis is correct) and thus unable to reduce its magnitude. Thus, the presence of temporally correlated errors in the input data substantially reduced the effectiveness of the four-dimensional analysis technique. Since much of the error in the satellite height fields was temporally correlated, the adjustments shown in Table 1 are much smaller than actual magnitude of the errors in the data.

The four-dimensional analysis technique was also performed with an unequal weighting based on the expected accuracy of each gridpoint value. The weighting

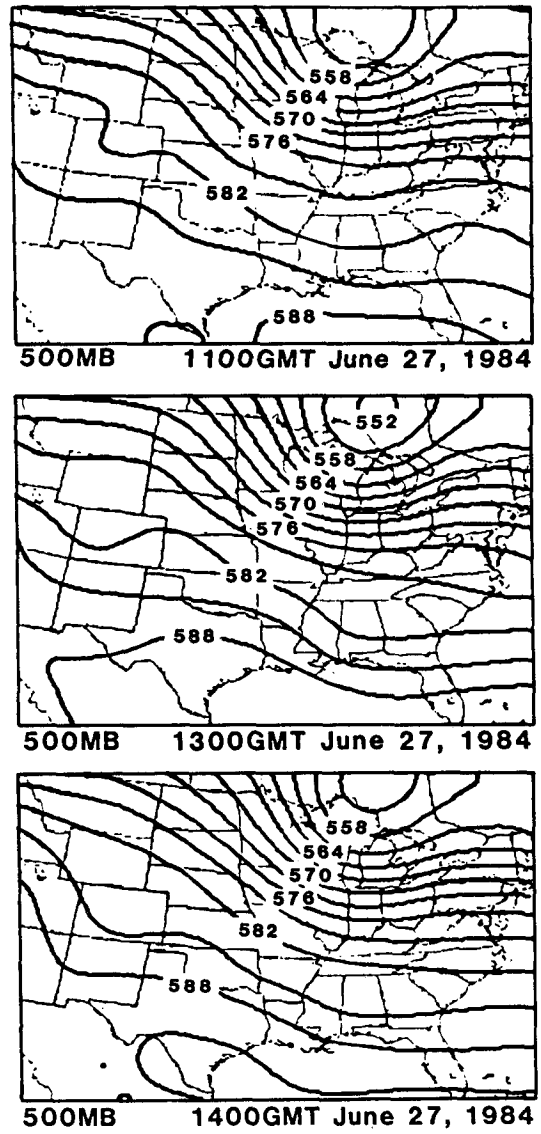

FIG. 5. 500 mb satellite height fields (dam) for 1100, 1300 and 1400 UTC 27 June 1984.

was designed to make the functional approximately equal to the integral

$$\sum_{r=1}^{R} \int_x \int_y \int_p w(\Psi(t_r) - \tilde{\Psi}(t_r))^2 dp dy dx, \qquad (5)$$

where w is a weight dependent on the expected error variance estimated using the product of the inverse of a vertically varying error variance and a horizontal data density. The horizontal data density fields are shown in Fig. 7 and the vertical error variances given in Table 2.

As can be seen in Fig. 6, the inclusion of the weighting did not change the 500 mb height analysis significantly. At other levels, the results were similar, with little difference between the analyses when unequal

TABLE 1. Rms height differences (m) between four-dimensional analysis and input height fields for 1200 UTC 27 June 1984. Analysis was performed assuming equal weighting.

	Pressure level (mb)									
UTC	1000	850	700	500	400	300	250	200	150	100
1100	4.15	4.06	4.52	6.28	6.91	8.09	8.92	8.59	8.97	8.27
1300	4.78	2.10	2.63	4.72	5.68	6.57	7.11	7.35	7.59	7.20
1400	7.15	3.86	3.20	4.89	5.28	6.74	7.75	7.74	7.74	8.02

weighting was or was not used. The one major difference between the equal and unequally weighted cases was that for the unequally weighted cases the analyses technique required over five times as many iterations to achieve convergence. However, since the convergence rate is a function of the solution algorithm, the

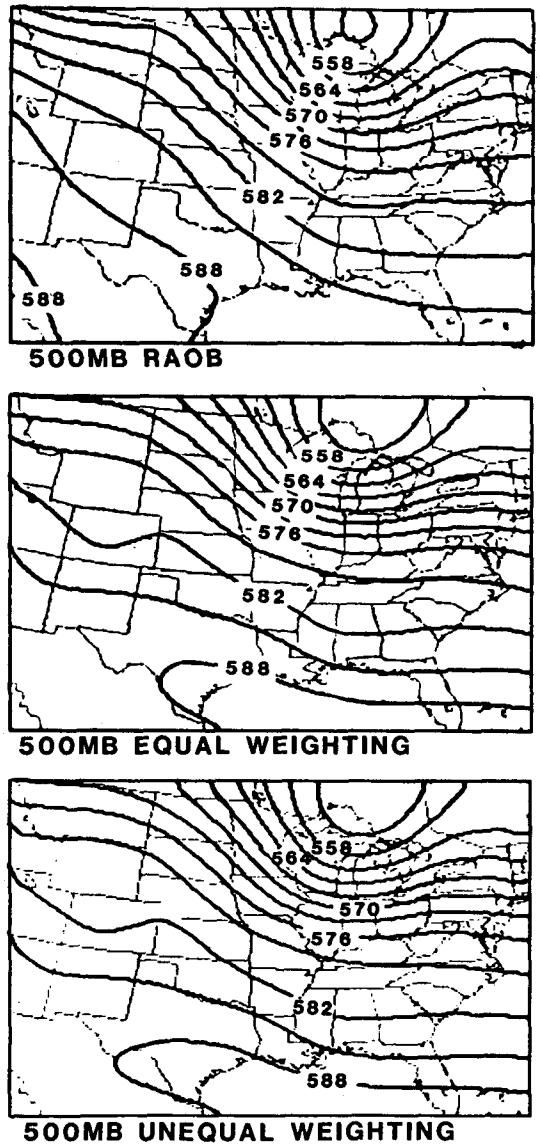

FIG. 6. 500 mb height fields (dam) created from rawinsonde data (a), four-dimensional analysis with equal weights (b), and four-dimensional analysis with unequal weights (c).

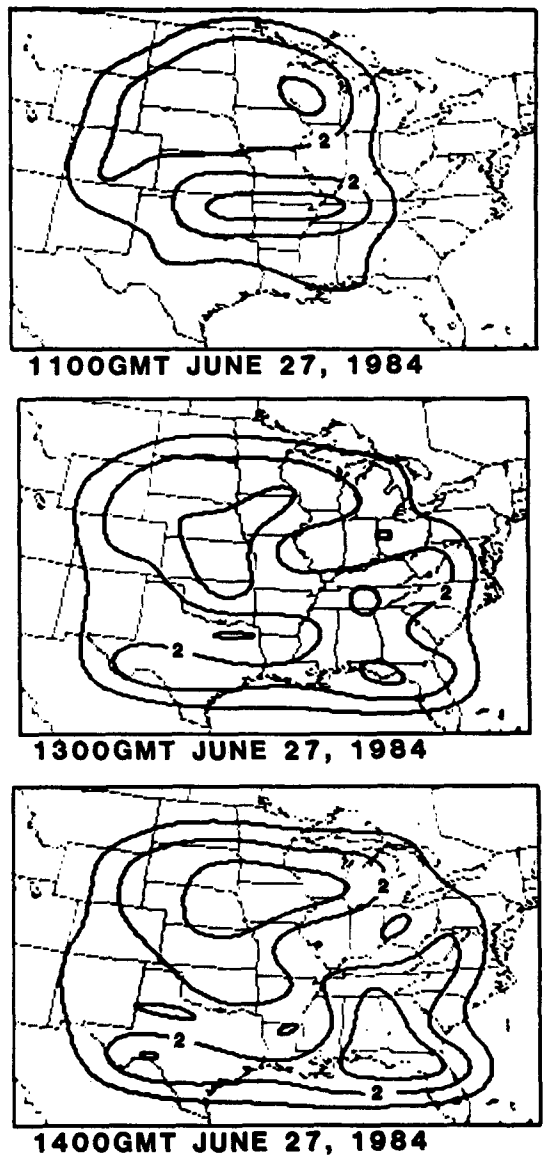

FIG. 7. Horizontal data density weights for 1100, 1300 and 1400 UTC 27 June 1984.

TABLE 2. Assumed vertical error variances (m^2) used in unequal weighting of input height fields.

	1000	850	700	500	400	300	250	200	150	100
variance (m^2)	400	625	900	1600	2500	2500	3600	3600	4900	3600

slow convergence rate with unequal weighting may be improved by a more advanced solution algorithm.

5. Summary and conclusions

In this paper, a variational four-dimensional data analysis technique for use on data distributed over some time interval has been presented. The four-dimensional analysis technique constrains the solution to be a quasi-geostrophic model solution while fitting the input fields as closely as possible. This best fit trajectory is found by use of a standard nonlinear minimization algorithm (conjugate gradients). The minimization algorithm uses the gradient of the functional with respect to the initial conditions which is calculated using an adjoint model as in Lewis and Derber (1985). The four-dimensional analysis technique was tested using a five-level quasi-geostrophic model on a series of fields created from a primitive equations model forecast. The four-dimensional analysis technique was well behaved and converged smoothly to the solution, with most of the adjustment occurring in the first few iterations. The convergence rate appears to be relatively independent of the amount of data in the analysis interval but does depend on the length of the analysis interval and the presence of horizontal boundaries. As could be expected, the quasi-geostrophic model solution was able to fit the p.e. solution better for shorter time intervals and with the inclusion of a greater amount of data. When three or more evenly distributed insertion times were used, the fit to the p.e. height fields was best at the center of the analysis interval. These results indicate that the four-dimensional analysis technique will work well irrespective of the amount of data included in the analysis period.

The assimilation technique also worked well on height fields created from satellite retrievals. In this case, the four-dimensional technique was applied using a ten-level quasi-geostrophic model as a constraint. Small but reasonable adjustments were made to the fields when each grid point value was equally weighted or a weighting dependent on an expected error variance was used. The adjustments with unequal weights produced very similar fields to the equally weighted results. However, the convergence rate decreased substantially, when unequal weights were included. Comparison to the 1200 UTC rawinsonde analysis indicates that the satellite analyses contained temporally correlated errors which cannot be removed by this or any other dynamical assimilation technique.

The four-dimensional analysis technique, described in this paper, has many potential applications to meteorological problems. For example, since the technique produces the solution which fits the data best, it could be used to examine the adequacy of the forecast model for describing a particular phenomenon. Also, the variational framework of the algorithm provides the capability to combine the four-dimensional analysis technique with other variational problems. Thus, by using a functional which compares the model solution to the observations at the observation locations and penalizes the presence of high frequency modes in the solution, a four-dimensional analysis could be produced directly from the observations. The extension of the technique to more complex problems does result in some practical difficulties, owing to the extensive computational resources required in the implementation of the four-dimensional analysis technique. However, with the current growth of computational resources, these limitations may become insignificant in the near future and the four-dimensional analysis technique could be successfully implemented on a variety of meteorological problems.

Acknowledgments. The author would like to thank John M. Lewis and Francois LeDimet for introducing the author to the assimilation technique and for their many helpful suggestions and Donna MacKenzie/Jan Schuster-Waite who typed the final manuscript. Many useful comments were also provided by Graham Mills, Thomas L. Koehler, David D. Houghton, R. James Purser, William L. Smith, and Lyle H. Horn and two anonymous reviewers. The satellite analyses were provided by John LeMarshall and Geary Callan. This work was supported by the National Aeronautic and Space Administration through Contract NAS5-21965 and the National Oceanographic and Atmospheric Administration under Grant NA-84-DGC-00155.

REFERENCES

Bengtsson, L., 1975: Four-dimensional assimilation of meteorological observations. WMO, Int. Council of Scientist Unions, Joint Organizing Committee, GARP Publ. Ser. No. 15, 75 pp.

——, M. Ghil and E. Kallen, Eds., 1981: *Dynamic Meteorology: Assimilation Methods.* Springer-Verlag, 330 pp.

Bloom, S. C., 1983: The use of dynamical constraints in the analysis of mesoscale rawinsonde data. *Tellus,* **35A,** 365–378.

Cressman, G. P., 1959: An operational objective analysis scheme. *Mon. Wea. Rev.,* **87,** 367–374.

Derber, J. C., 1985: The variational four-dimensional assimilation of analyses using filtered models as constraints. Ph.D. thesis, University of Wisconsin-Madison, 142 pp.

Gill, P. E., W. Murrey and M. H. Wright, 1981: *Practical Optimization.* Academic Press, 401 pp.

Hoffman, R. N., 1986: A four-dimensional analysis exactly satisfying equations of motion. *Mon. Wea. Rev.,* **114,** 388–397.

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

Lewis, J. M., 1982: Adaptation of P. D. Thompson's scheme to the constraint of potential vorticity conservation. *Mon. Wea. Rev.,* **110,** 1618–1653.

——, 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.* Translated by S. K. Mitter, Springer-Verlag, 396 pp.

McPherson, R. D., 1975: Progress, problems and prospects in meteorological data assimilation. *Bull. Amer. Meteor. Soc.,* **56,** 1154–1165.

Mills, G. A., and C. M. Hayden, 1983: The use of high horizontal resolution satellite temperature and moisture profiles to initialize a mesoscale numerical weather prediction model—a severe weather event case study. *J. Climate Appl. Meteor.,* **22,** 649–663.

Polak, E., 1971: *Computational Methods in Optimization, A Unified Approach.* Academic Press, 329 pp.

Sasaki, Y., 1958: An objective analysis based on the variational method. *J. Meteor. Soc. Japan, II,* **36,** 77–88.

——, 1970: Some basic formalism in numerical variational analysis. *Mon. Wea. Rev.,* **98,** 875–883.

Shanno, D. F., 1978: Conjugate gradient methods with inexact searches. *Mathematics of Operations Research,* **3,** 244–256.

Smith, W. L., 1983: The retrieval of atmospheric profiles from VAS geostationary radiance observations. *J. Atmos. Sci.,* **40,** 2025–2035.

Thompson, P. D., 1969: Reduction of analysis error through constraints of dynamical consistency. *J. Appl. Meteor.,* **8,** 738–742.