THE SSI ANALYSIS SYSTEM AND EXTENSIONS TO 4D

John C. Derber, David F. Parrish and Joseph G. Sela
National Meteorological Center
Washington D.C., USA

The Spectral Statistical Interpolation (SSI) analysis system is the only 3D variational analysis currently used in operational numerical weather prediction. The basic structure of this system will be described in this paper. Recent improvements in the analysis and more details of the analysis system are described in the companion paper by Parrish and Derber(1992b). The 3D analysis system can be extended to 4D by the use of a prediction model and the adjoint of the prediction model. A few simple experiments have been performed with the full 4D variational system over a 6 hour assimilation interval. Results will be presented from the 3D system and the preliminary 4D system.

1. INTRODUCTION

The Spectral Statistical Interpolation (SSI, note that this type of system is also referred to as 3D var) analysis system has been in operational use at the National Meteorological Center (NMC) for over a year (since 25 June 1991). Over this period, no major problems have developed with the system. Further enhancements from the original analysis system have been implemented and are planned for the future (see Parrish and Derber, 1992b). In this paper, the basic structure of the analysis system will be discussed and our first attempts to extend the system to the time dimension (4D SSI or 4D var) will be presented. The development of this system parallels the development of a similar system at ECMWF (Rabier and Courtier, 1992 and Thepaut and Courtier, 1991).

The analysis system used at NMC prior to the SSI system was an optimal interpolation (OI) based system (Bergman 1979, Dey and Morone 1985, DiMego 1988 and Kanamitsu 1989). With this system, it was necessary to follow the analysis procedure with a transformation of the analysis variables into the model variables before applying an initialization procedure (Ballish et al., 1992). The SSI analysis system is attempting to solve the same basic equations as the OI analysis procedure. However, the SSI equations are solved in an entirely different manner than in the OI system. Because of this, it is possible to solve the equations with fewer approximations, better constraints and in a form more compatible with the prediction model. Thus, the SSI system produces analyses directly in the model coordinates, uses a linear balance equation (with an empirical friction effect) and uses all of the data at a single time to produce a global analysis. This results in an analysis which is sufficiently balanced to eliminate the need for the application of the initialization procedure.

With the successful creation of the SSI analysis system and the creation of a tangent linear model and associated adjoint model, it is possible to extend the SSI system to 4D. A few experiments have been performed with such a system over a 6 hour assimilation period. Data is only available at the end of the assimilation period, while the guess is at the initial time. All other aspects (except a few minor differences) are kept the same as the 3D SSI system. While this experiment does not incorporate all of the advantages of a full 4D system, the experiment is useful in evaluating the correctness of the system (since the 3D and 4D analyses should be similar at the end of the assimilation period) and determining some of the difficulties of directly incorporating observations into a 4D system.

2. The SSI analysis system

The SSI analysis system is described in papers by Parrish and Derber (1992a) and Derber et al. (1991). A brief review of the system will be given in the following section. The SSI analysis system minimizes the objective function given by,

$$J = 1/2 \; [\; x^t \; B^{-1} \; x + (L \; x - y)^t \; (F + O)^{-1} \; (L \; x - y)] \qquad (1)$$

where

x = N-component vector of analysis increments;

B = NxN forecast error covariance matrix;

O = MxM observational error covariance matrix;

F = MxM representativeness error covariance matrix;

L = Linear transformation operator that converts the analysis variables to the same type and location;

y = M component vector of observational residuals i.e., $y = y_{obs} - L\, x_{guess}$;

N = number of degrees of freedom in the analysis; and

M = number of observations.

The linearity in L is not essential but does impact the choice of minimization algorithms. The equation for the minimum of J can be found by differentiating with respect to x and setting the result to zero. Multiplying this equation by B and rearranging terms results in

$$[\, I + B\, L^t\, (F + O)^{-1}\, L]\, x = B\, L^t\, (F + O)^{-1}\, y \qquad (2)$$

If B can be written in the form $B = C\, C^t$ then by defining a new variable $w = C^{-1}\, x$, the equation can be rewritten in a form that ensures symmetry in the matrix to be inverted and nicely preconditions the problem.

$$A\, w = f \qquad (3)$$

where

$$A = I + C^t\, L^t\, (F + O)^{-1}\, L\, C \qquad (4)$$

and

$$f = C^t\, L^t\, (F + O)^{-1}\, y \qquad (5)$$

This is the primary analysis equation for the SSI analysis system.

Note that in (1), it is not necessary that the analysis variables be the same as the observation variables. This is one of the advantages of this form of the problem over conventional implementations of the OI analysis system. The observations can be used in a more primitive form and are not contaminated by ill-posed inversion techniques. This is the subject of several other papers in this workshop. Another advantage of differing analysis and observational variables is that the analysis variables can be chosen based on other criteria. We have chosen to define the variables such that the balanced part of the mass field is included in the wind variables. Thus the L operator implies the balanced part of the mass field from the vorticity and divergence through a linear balance equation with empirical friction (see

Parrish and Derber, 1992a and 1992b for details). This eliminates some of the large off diagonal components in the B matrix and concentrates most of the analysis signal in a reduced set of variables.

In the vertical, the variables are decomposed in terms of Empirical Orthogonal Functions (EOFs) calculated from estimates of the vertical forecast error covariance matrices. All vertical modes are used for all variables. This differs from what is stated in Parrish and Derber (1992a) and results from the reformulation of the balance equation (Parrish and Derber, 1992b).

The L operator transforms the analysis variables to the same form as the observations. This involves reconstructing the variables in the vertical; calculating a balanced part of the mass field; converting to temperature, surface pressure, u and v components of the wind, and specific humidity; transforming from spectral to grid point values and finally interpolating to the observation locations. In the current operational system, the last step is not performed because "superobservations" are created after the initial calculation of residuals (except for satellite temperature profiles). For satellite temperature profiles, the initial residuals are only interpolated in the vertical to the sigma levels to allow the application of the vertical error covariance matrix. Experiments in increasing the order of interpolation and the elimination of the "superobservations" are given in Parrish and Derber (1992b).

The solution to (3) is found using a linear conjugate gradient algorithm. Currently, 100 iterations are performed. This is probably more iterations than necessary.

The y vector in (3) contains the initial differences between the guess and observations. It is important to represent this difference as accurately as possible so that there will not be an incorrect indication that the guess is wrong. In order to improve the representation of this difference without increasing the computational cost too much, several improvements have been added to the system which are only in the initial residual calculation, not the iterative part of the code.

First, the ten meter winds predicted by the model are used in the calculation of the residuals. The lowest sigma level of the model is above 10m and thus in order to calculate near surface winds it was necessary to extrapolate. In the initial version of the analysis, the winds were held constant below the bottom sigma level. By using the predicted 10m winds, the initial near surface residuals were reduced, producing somewhat improved near surface analyses.

The second major change in the initial calculation of the residuals was to include a time interpolation to the observation time. At this time, only the analysis from six hours before the analysis time and the guess at the analysis time are available. Thus, the time interpolation is only performed when the observation time is between -6 and 0 hours. In the future, it is hoped that the forecast will be saved more frequently in time and the forecast extended to 9 hours to allow accurate interpolation to the observation time. After the initial computation of the residuals all observations are assumed at the analysis time. The difference in observation times can be properly accounted for in the iterative part of the procedure only in a 4D system.

Results from the SSI analysis system are presented in Derber et al. (1991) and Parrish and Derber (1992a,1992b). In the next section of the paper, the strategy for extending this system to 4D will be discussed and in the following section a few preliminary results will be presented.

3. Extension of SSI to 4D

The extension of the SSI analysis to 4D will allow the use of data over a longer time interval and will account for the time distribution of the observations more correctly. The 4D version of SSI is being developed under the assumption that it will not be possible to perform the assimilation with the operational prediction model with full physics. The reasons for this assumption will be given later.

To extend the SSI system to 4D, it is only necessary to incorporate the forecast model (and a simple time interpolation) into the definition of the L operator in (1). Thus the model will be integrated through the period in which the observation occurred and then a time interpolation will be performed to the actual observation time from the nearest saved solutions from the model forecast. In the ideal case, the solution would be saved every time step of the model and the time interpolation would only involve a very short interpolation between time steps. However, because of the very large storage requirements of saving the solution every time step, we are anticipating that the solution will not be saved every time step. Since the L operator will contain the forward model and the time interpolation operator, the L^t operator will contain the adjoint of the time interpolation operator and the adjoint of the forward model. The development of the adjoint of the forecast model and some simple tests of the models on analyses are discussed in Navon et al. (1992).

For the 4D problem, the x vector in (1) is the control vector. In conventional 4D assimilation this is the initial state. However, the x vector could also be model parameters or a bias correction to the model equations (Derber, 1989; Wergen, 1992; Zupański, 1992), or any combination of these variables as long as the problem stays well-posed. For the rest of this paper, we will only consider the case of the control vector being the initial conditions. However, we anticipate that the final system will probably include both the initial state and a bias correction vector.

If the minimum to (1) were to be found for the complete nonlinear forecast model, then the L operator would no longer be linear. This does not present any theoretical problem (except, perhaps, the discontinuities in the parameterizations), but does present some practical problems. First, it would be necessary to have the exact adjoint of the complete nonlinear model, including all parameterizations. While this is possible, considerable effort must be expended. Secondly, for very nonlinear parameterization there is an additional problem of the adjoint needing intermediate solutions from the nonlinear solution in reverse order. Thus, it is either necessary to store these intermediate solutions or recalculate them. This can greatly increase the memory or computational expense. Third, the minimization algorithm is more complicated for a nonlinear forecast model. For a linear forecast model, each iteration only requires a single forward integration of the forecast model and a backwards integration of the adjoint model. For a nonlinear forecast model it is necessary to estimate a stepsize. This stepsize estimation can require additional integrations of the forecast model and introduces an uncertainty in the minimization procedure because it is not possible to exactly calculate the stepsize. Finally, the operational forecast models tend to require the limits of the available computational resources and have discontinuous parameterizations. Under these conditions, the operational use of the complete operational prediction model in a 4D variational assimilation is not practical. Of course, better results should be expected when one uses a system which is closer to the operational model and all effort should be made to make them as close to the operational model as feasible.

Thus, the tentative configuration for the assimilation system was based on the assumption that the complete forecast model with physics would not be used. The proposed algorithm has two levels of iteration. In the internal iteration, a linear version of the model is used with it's exact adjoint to solve (3) using linear conjugate gradients. Only minor approximations have been made in the solution of (3) in the linear model. In fact, fewer assumptions have been made than for the 3D SSI with the elimination of the "superobing". The minor approximations which are still made in the internal iteration include interpolation of the satellite profiles to the sigma surfaces and the winds below the lowest sigma level are assumed to be at the bottom sigma level. In the future, these minor approximations can be removed.

The external iteration defines the right hand side of (3). This is done periodically and the internal iteration restarted in order to partially account for the nonlinearities and the missing physics. The external iteration consists of adjusting the initial state for the nonlinear model based on the adjustment found in the internal iteration, an integration of the model over the assimilation period, recalculating the model - observation differences and recalculating the right hand side. Note that no assumptions are made and the right hand side of (1) is calculated exactly as in the 3D SSI.

4. Preliminary results from the 4D SSI

A simple experiment was designed to determine the feasibility of the system described in the previous section and to determine the correctness of the system. For 0000UTC 28 October 1992, all of the data and fields used in the 3D SSI analysis were saved from a T62, 18 level version of the NMC operational model being run in parallel. Included in the fields was the 6 hour forecast from 1800UTC 29 October, the analysis from 1800UTC 28 October and the analysis at 0000UTC 28 October. From these data, a series of experiments were run. In all of these experiments, the same statistics, data, balance equation, analysis variables, etc. were used as in the 3D SSI analysis. Two of these experiments will be described below. All experiments are performed at the full resolution of T62 with 18 vertical levels.

The first experiment, which will be referred to as LINEAR 4D, performed the 4D assimilation using only the internal iteration procedure after the initial residuals were calculated from the 6 hour nonlinear forecast. 30 iterations were performed. The corrections to the initial conditions were then inserted into the initial conditions for the full nonlinear model with physics and run over the 6 hour forecast period. Note that in this experiment, the linear model and adjoint model used a mean state which was constant in time and equal to the average of the analysis at 1800UTC and the six hour nonlinear forecast from this field. After the initial time interpolation to the observation times from the initial full nonlinear forecast, the observations were assumed to all be at 0000UTC. Thus in this case, the right hand side was exactly equal to the right hand side of the 3D SSI analysis and the same assumption that the observation times were at 0000UTC was used for both the 3D SSI analysis and the LINEAR 4D in the iterative procedure. Thus, the major difference between the 3D SSI analysis and the LINEAR 4D experiment was that the correction to the guess was made at 1800UTC for the LINEAR 4D and six hours later for the 3D SSI. The only other difference between the LINEAR 4D and 3D SSI experiments is in the "superobservation " assumption in the 3D SSI. If the 4D assimilation is working properly, the resultant fields should be similar between the 3D SSI and the LINEAR 4D solution at 0000UTC.

In fig. 1, the results from the 3D SSI analysis are shown while the results from the LINEAR 4D are shown in fig. 2. In both figures, the differences from the same initial 6 hour forecast valid at 0000UTC are shown. While significant differences between these fields exist the large scale differences are similar. Note, that with the 4D LINEAR solution the maximum differences are larger in magnitude and smaller in scale than the 3D SSI solution. Some of this difference is due to the "superobing" in the 3D SSI. Fig. 3 shows the 3D SSI solution with the removal of the "superobing" and the use of 7^{th} order polynomial interpolation discussed in Parrish and Derber (1992b). Note that the magnitude of these differences are larger than those in fig. 1, as was seen in the LINEAR 4D solution.

In fig. 4, the same fields are shown for another experiment. In this experiment, referred to in NONLINEAR 4D, after every 10 internal iterations, an external iteration is performed, i.e., the full nonlinear model with physics is rerun, the right hand side recalculated and the internal iteration restarted. In this case, 3 external iterations are performed to give about the same total number of iterations as for the LINEAR 4D. Note that two other differences from the LINEAR 4D were included. First, the nonlinear solution around which the linear and adjoint model are linearized varies with each time step. However, the time varying solution is not determined from the full nonlinear model but rather is a simple linear interpolation between the initial and

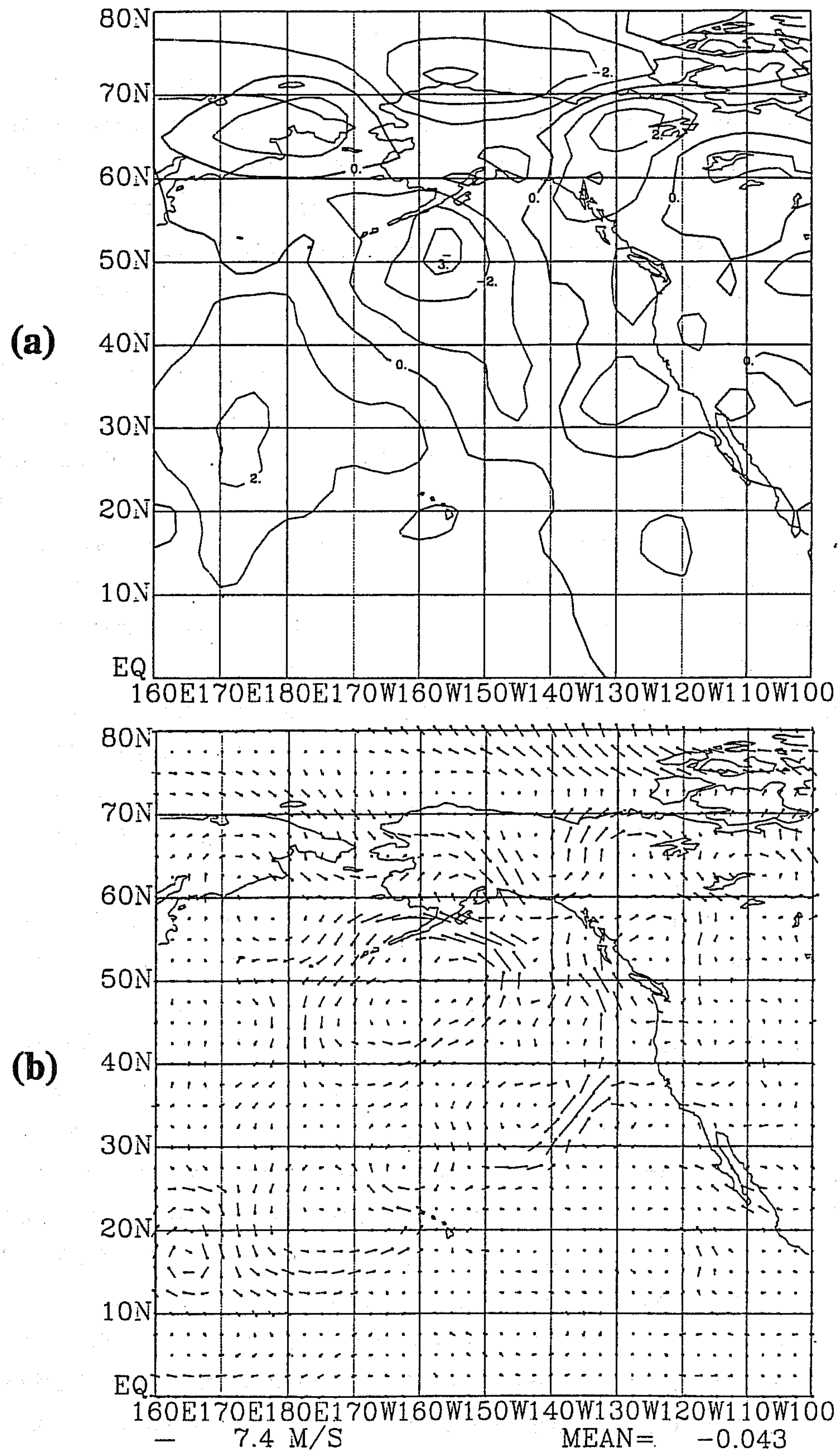


Fig. 1: Difference between the 3D SSI analysis at 0000UTC 28 October 1992 and 6 hour forecast from analysis at 1800UTC 27 October 1992. Results are shown for temperatures (a) and for winds (b) from sigma level 12 (about 325mb).

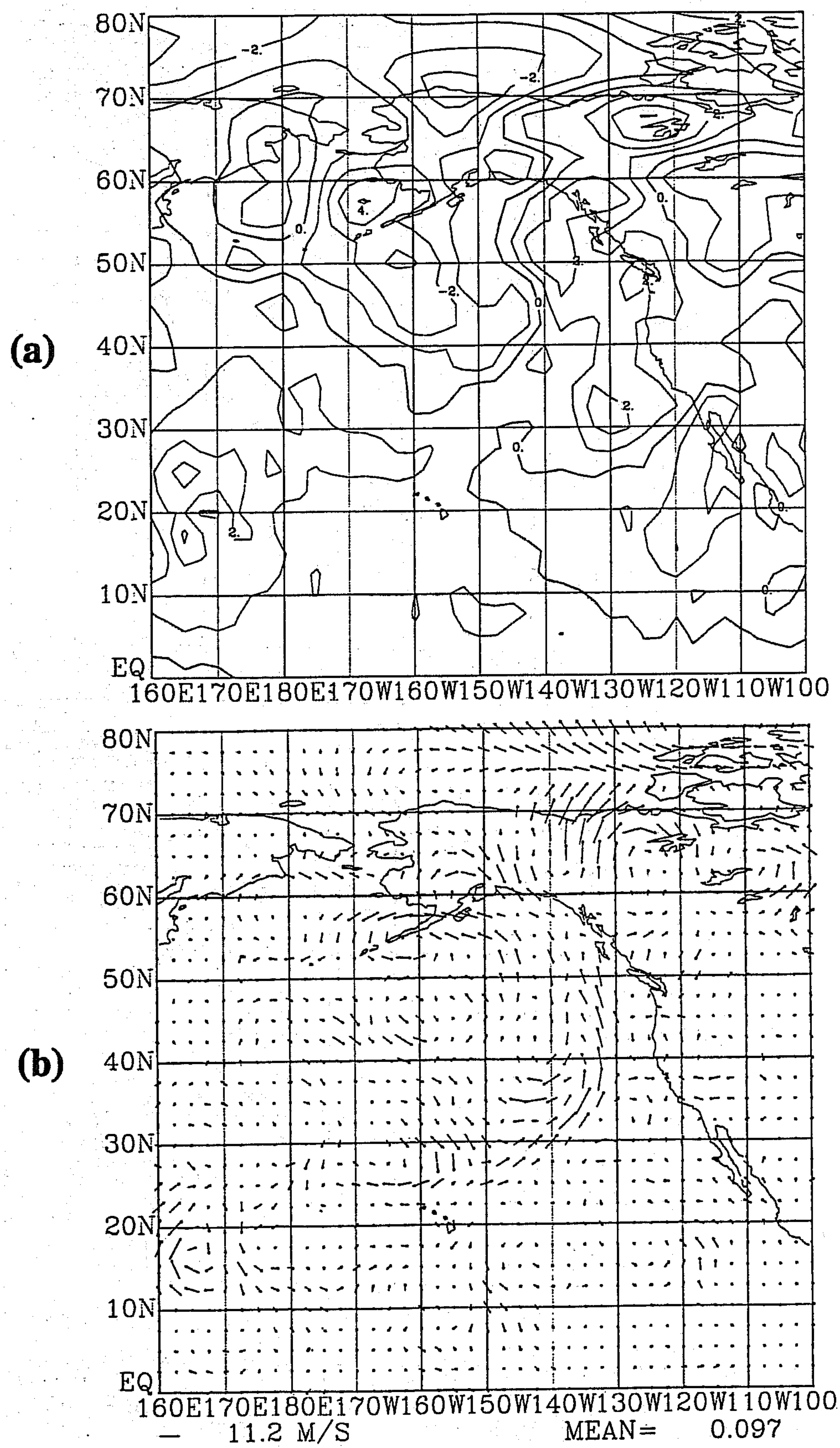


Fig. 2: Same as Fig. 1 except LINEAR 4D analysis.

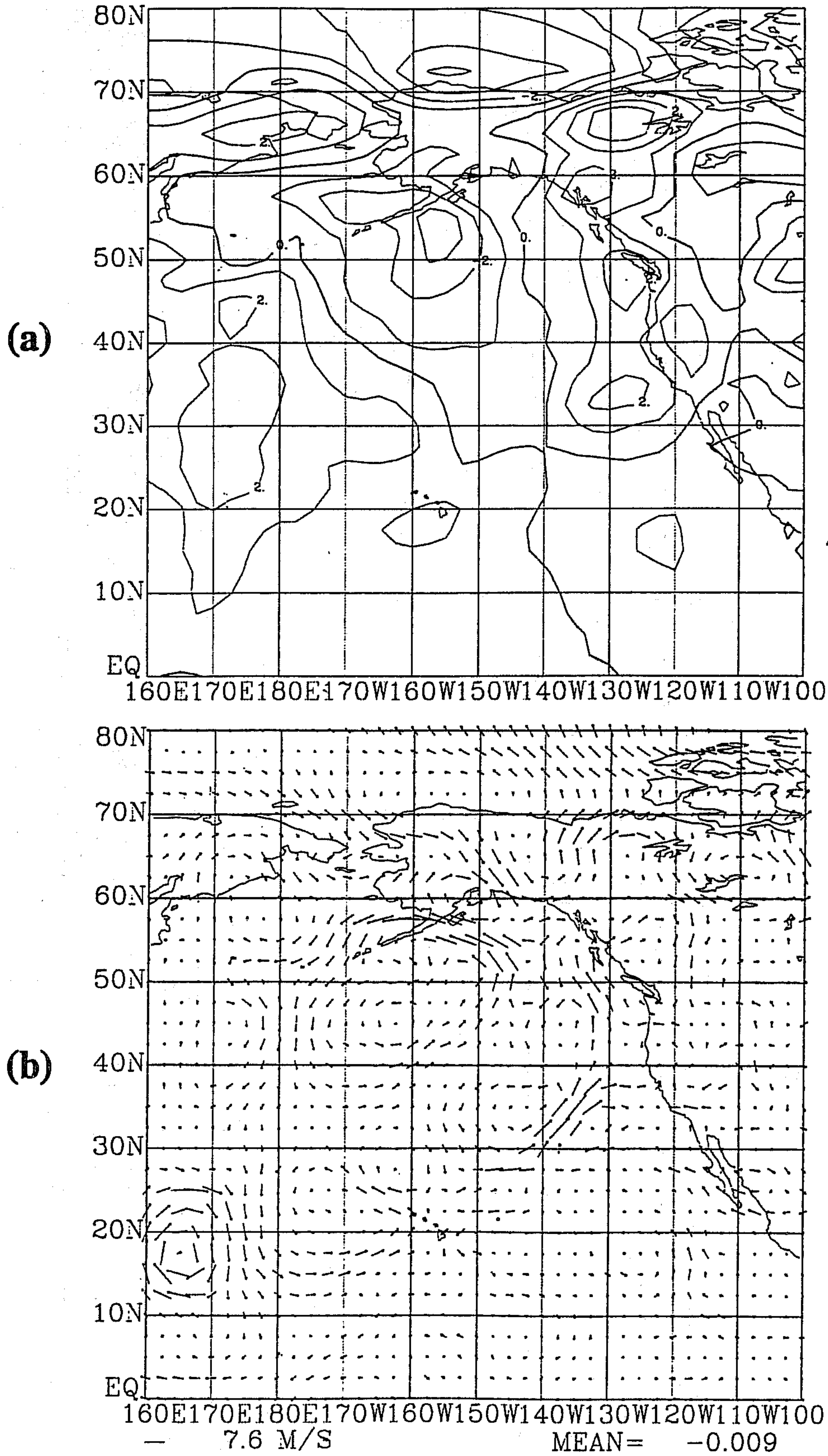
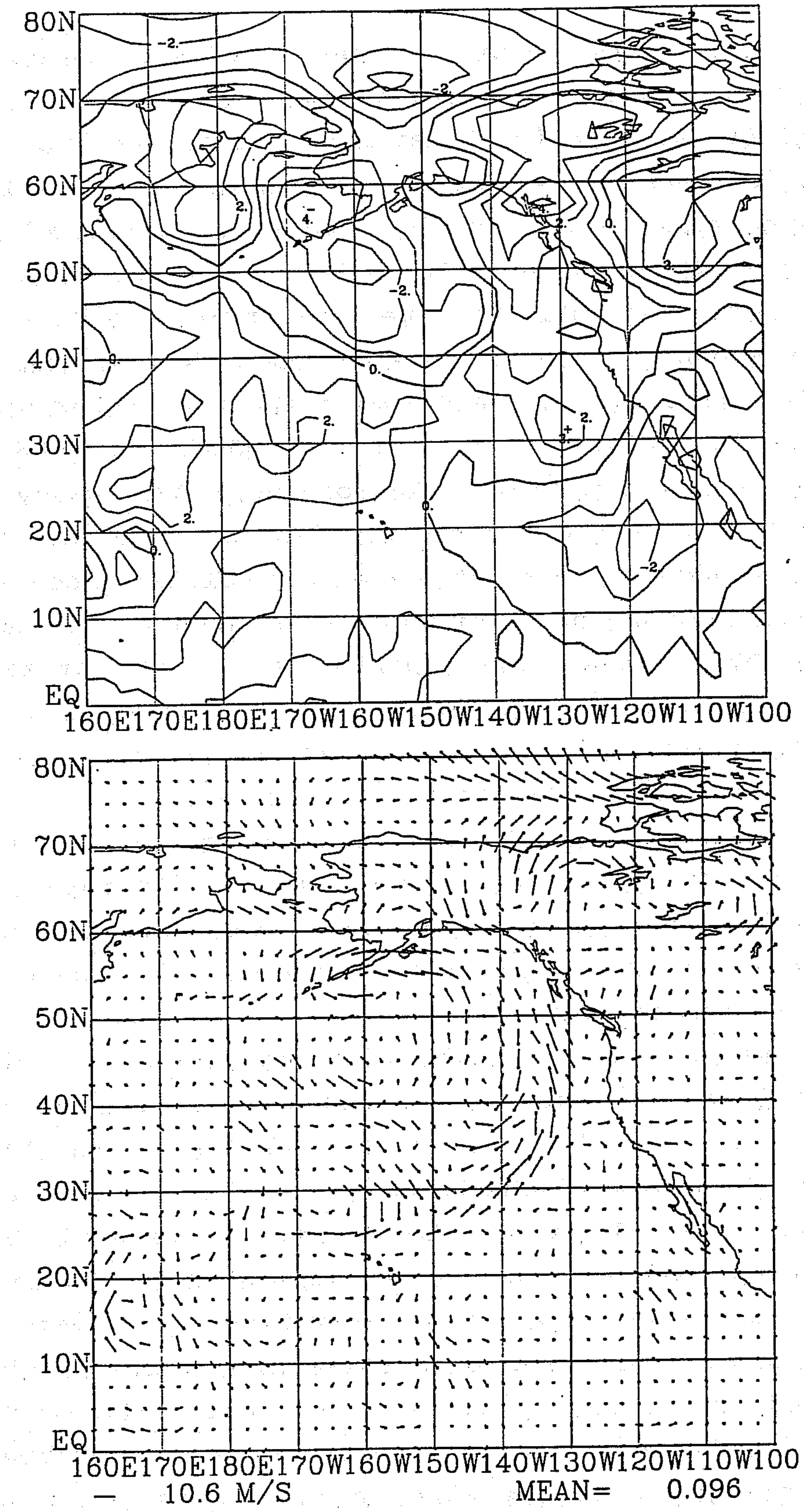


Fig. 3: Same as Fig. 1 except 3D SSI without "superobing" and 7th order polynomial interpolation.

(a)



(b)

Fig. 4: Same as Fig. 1 except NONLINEAR 4D analysis.

final time of the latest 6 hour forecast. Other experiments appear to indicate that the average nonlinear solution used in the LINEAR 4D experiment may give better results. This result may arise because some small scale noise in the solution may be overemphasized in the interpolated solution while being smoothed in the averaging case. It is hoped that when the solution is saved more frequently in time, this problem will be reduced.

The second difference between the LINEAR 4D and NONLINEAR 4D experiments is in the handling of the time interpolation to the observation time. In the NONLINEAR 4D case, the time interpolation is accounted for correctly with observations before 0000UTC contributing to the gradient at 1800UTC. This did not appear to have a strong effect on the solution because all of the data was within 3 hours of 0000UTC and most of the data was close to 0000UTC. Thus, the contribution at the initial time was small. The proper inclusion of this term will become more important when the nonlinear solution is saved more frequently in time or longer time intervals are used.

In comparing figs. 2 and 4, several important differences can be seen. The magnitude of some of the differences can be seen to be reduced for the nonlinear case and the patterns are different in some locations. However, the basic large scale patterns remain similar.

One measure of the quality of the analysis is the final fit to the data. In table 1, the reduction of the weighted squared error (the contribution to the objective function by each variable) by each of the three analyses is given. Note that the best fit is still given by the 3D SSI. In general, the next best fit to the data is given by the NONLINEAR 4D. Some of this can be explained by the fewer iterations performed in the 4D experiments. With one more external iteration, the NONLINEAR 4D's fit to the wind improves upon that for the 3D SSI. However for the conventional temperatures, it does not appear that the 4D assimilation results will improve on the 3D SSI with further iterations. This may be a result of problems we are having in the height to temperature and surface pressure transformations.

Table 1: Percentage of original contribution to objective function for each variable type after production of analysis. Results are shown for 3D SSI, LINEAR 4D and NONLINEAR 4D analysis.

	3D SSI	LINEAR 4D	NONLINEAR 4D
Surface Pressure	48.7	55.9	52.0
Conventional Temperature	47.8	59.0	57.4
Satellite Temperature	99.1	101.7	102.3
U - Component Wind	47.9	53.1	48.8
V - Component Wind	48.0	52.7	48.4
Specific Humidity	63.4	82.7	69.0
Total Objective Function	49.0	54.6	51.1

Note the substantial improvement in the specific humidity fit to the data from the LINEAR 4D to the NONLINEAR 4D experiments. This is probably do to the inclusion of the precipitation mechanisms in the assimilation by including the external iterations.

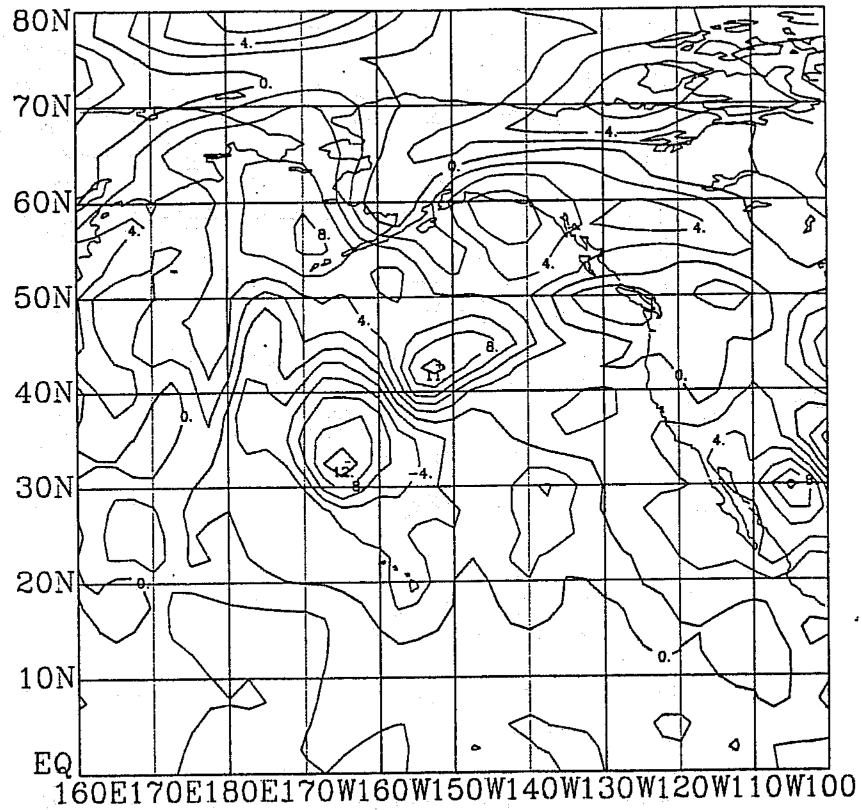
Finally, 5 day forecasts were made from the 3D SSI solution, the NONLINEAR 4D and the LINEAR 4D solutions. The forecast error (when compared to a T62 18 level parallel assimilation) for the 3D SSI solution and the NONLINEAR 4D solution are shown in figs. 5 and 6, respectively. Note that the patterns are quite similar. This is to be expected because of in the design of the experiment no additional data was added to the analysis and the handling of the time distribution of the data was only slightly improved. However, the magnitude of the maximum wind error is substantially reduced over this region and the magnitude of the temperature errors is slightly reduced. This good result is supported for the northern hemisphere in the comparison of the RMS height errors and the anomaly correlations for the experiments in table 2. However, in the southern hemisphere, the errors are slightly worse in the NONLINEAR 4D experiment. We suspect that this may result from difficulties in the specification of the satellite error covariances. Note that in table 2, the fit to the satellite data actually was slightly worse in the 4D assimilation than in the guess and only marginally better in the 3D SSI. This problem will be examined further.

5. Conclusions and future research

The 3D SSI analysis has proven to be a very useful and reliable analysis system at NMC. Additional improvements to the 3D SSI analysis system have been made over the last year and are continuing to be made. One of the most fruitful set of improvements to the system will be in the incorporation of new data types. Substantial effort is expected to be directed towards this problem in the future at NMC.

With the creation of the 3D SSI and the linear and it's adjoint, a prototype 4D assimilation system has been developed.

(a)



(b)

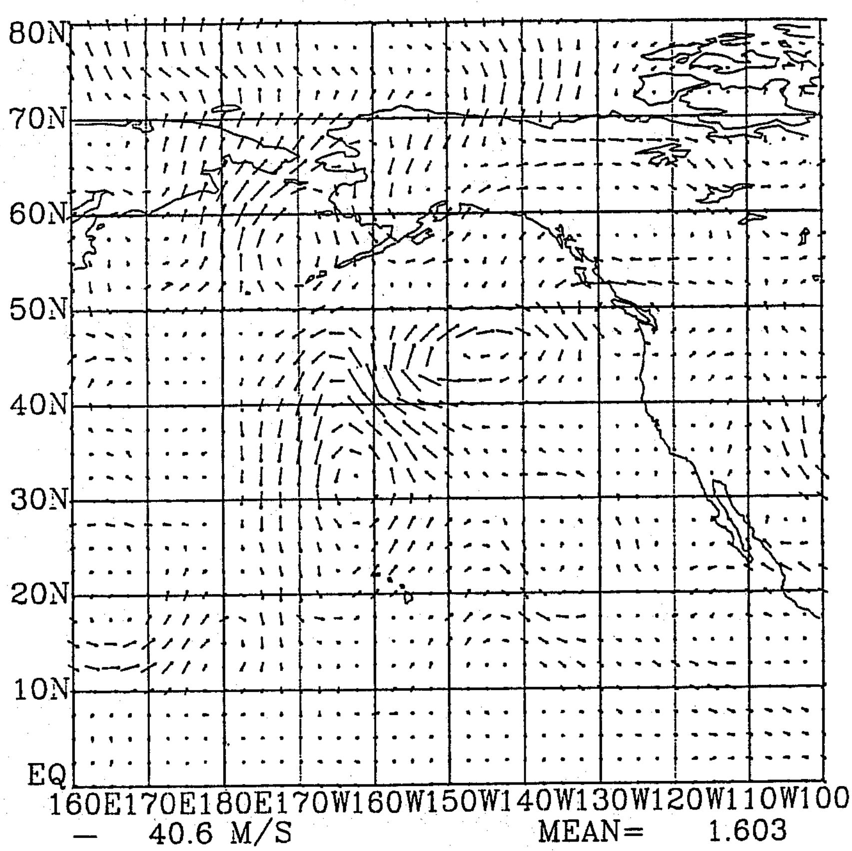


Fig. 5: Same as Fig. 1 except difference between 120 hour forecast from 3D SSI analysis and verifying analysis. Analysis is from conventional 4D assimilation using 3D SSI analysis system.

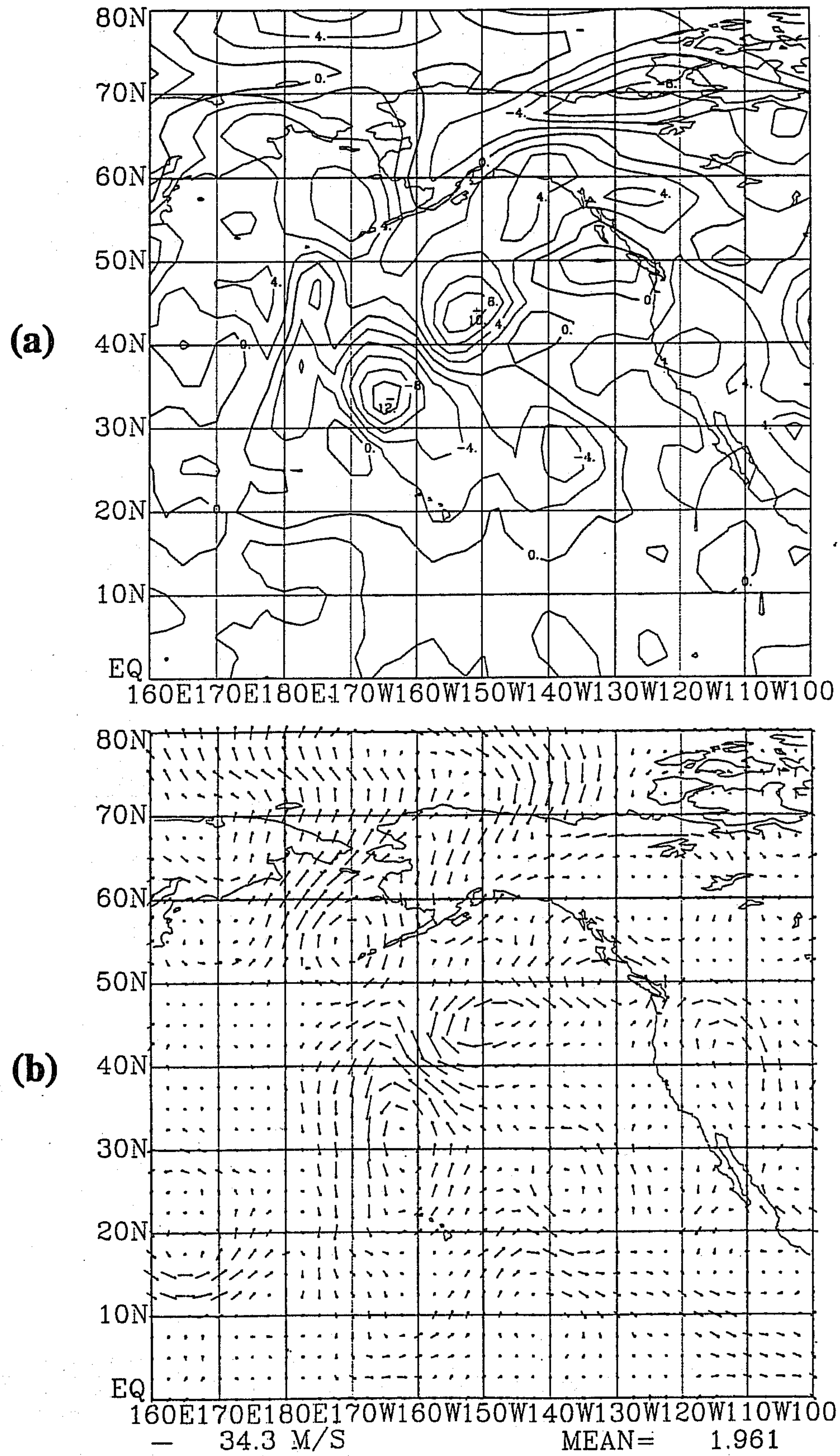


Fig. 6: Same as Fig. 5 except forecast is from NONLINEAR 4D analysis.

Table 2: 5 day anomaly correlations and RMS errors for 3D SSI, LINEAR 5D, and NONLINEAR 4D. Results are for 500mb height fields. Anomaly correlations are in percent and RMS errors are in meters.

Anomaly Correlation

DAY	0	1	2	3	4	5
3D SSI	100	97.8	93.3	88.9	85.1	72.2
LINEAR 4D	99.8	97.8	93.6	88.9	85.0	72.3
NONLINEAR 4D	99.8	97.7	93.9	90.0	86.0	72.8

RMS error

DAY	0	1	2	3	4	5
3D SSI	0.0	17.1	30.5	40.6	51.2	69.7
LINEAR 4D	5.7	17.3	29.8	40.1	49.6	69.5
NONLINEAR 4D	6.0	17.5	29.1	38.0	47.3	67.7

The preliminary experiments shown in this paper, show that the system appears to be performing properly. In the near future, further assimilation experiments will be performed. The system will be applied to longer assimilation periods, compared to the nonlinear solution more frequently in time and the optimal number of iterations examined. Also, the use of different control variables will be explored and the system will be applied to many different cases. The computational expense, the presence of model errors and the unavailability of future data may make this system less than optimal for the operational prediction problem. However, for the climate analysis problem, the 4D variational assimilation may be ideal since the analysis time will be the middle of the assimilation period, thus allowing the use of future data and the reduction of the effects of the model error (Derber, 1987). During the next few years the feasibility and practicality of the variational assimilation system should be determined. However, the initial results from this system are very encouraging.

References

Ballish, B., X. Cao, E. Kalnay, and M. Kanamitsu, 1992: Incremental nonlinear normal mode initialization, Mon. Wea. Rev., 120, 1723-1734.

Bergman, K., 1979: Multivariate analysis of temperature and winds using optimum interpolation. Mon. Wea. Rev.,107, 1423-1444.

Derber, J.C., 1987: Variational four-dimensional analysis using quasi-geostrophic constraints. Mon. Wea. Rev., 115, 998 - 1008.

Derber, J.C., 1989: A variational continuous assimilation technique. Mon. Wea. Rev., 117, 2437-2446.

Derber, J.C., D.F. Parrish, and S.J.Lord, 1991: The new global operational analysis system at the National Meteorological Center. Wea. Forecasting, 6, 538-547.

Dey, C.H., and L.L. Morone, 1985: Evolution of the National Meteorological Ceter global data assimilation system: January 1982 - December 1983. Mon. Wea. Rev., 113, 304-318.

DiMego, G.J., 1988: the National Meteorological Center Regional Analysis System. Mon. Wea. Rev., 116, 977-1000.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system Wea. Forecasting, 4, 335-342.

Navon, I.M., Zou, X., Derber, J.C. and J. Sela, 1992: Variational data assimilation with an adiabatic version of the NMC spectral model. Mon. Wea. Rev., 120, 1433-1446.

Parrish, D.F. and J.C. Derber, 1992a: The National Meteorolgical Center's spectral statistical-interpolation analysis system. Mon. Wea. Rev., 120, 1747-1763.

Parrish, D.F. and J.C. Derber, 1992b: Experimnts with the background error covariance matrix in the SSI analysis system., This volume.

Rabier, F. and P. Courtier, 1992: Four-dimensional assimilation in the presence of baroclinic instability. Q.J. Roy. Meteor. Soc., 118, 649-672.

Thepaut, J-N. and P. Courtier, 1991: Four-dimensional variational data assimilation using the adjoint of a multilevel primitive equation model. Q. J. Roy. Meteor. Soc., 117, 1225-1254.

Wergen, W., 1992: The effect of model errors in variational assimilation, Tellus, 44A, 297-313.

Zupanski, M., 1992: Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment. Submitted to Mon. Wea. Rev.