# A Scheme for Continuous Data Assimilation

W. F. STERN AND J. J. PLOSHAY

*Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, New Jersey*

(Manuscript received 15 May 1991, in final form 21 October 1991)

## ABSTRACT

Major revisions to the Geophysical Fluid Dynamics Laboratory's (GFDL) continuous data-assimilation system have been implemented and tested. Shortcomings noted during the original processing of data from FGGE [First GARP (Global Atmospheric Research Program) Global Experiment] served as the basis for these improvements. This new system has been used to reanalyze the two FGGE special observing periods. The main focus here will be on assessing the changes to the assimilation system using comparisons of rerun test results with results from the original FGGE processing.

The key new features in the current system include: a reduction in the assimilation cycle from 12 to 6 h; the use of a 6-h forecast first guess for the OI (optimum-interpolation analysis) as opposed to the previous use of persistence as a first guess; an extension of the OI search range from 250 to 500 km with an increase in the maximum number of observations used per analysis point from 8 to 12; the introduction of incremental linear normal-mode initialization, eliminating the periodic nonlinear normal-mode initialization; and an increase in the horizontal resolution of the assimilating model from 30 waves to 42 waves, rhomboidally truncated.

Tests of the new system show a significant reduction in the level of noise, improved consistency between mass and momentum analyses, and a better fit of the analyses to observations. In addition, the new system has demonstrated a greater ability to resolve rapidly moving and deepening transient features, with an indication of less rejection of surface pressure data.

In addition to the quantities archived during the original FGGE data processing, components of diabatic heating from the assimilating model have also been archived. They should be used with caution to the extent that they reflect model bias and spinup in addition to real features of the general circulation.

## 1. Introduction

Techniques for dynamically producing analyses via the assimilation of observed data may be grouped into two general categories: *intermittent* (forecast–analysis system) and *continuous* data assimilation. Both of these methods continue to be used successfully for operational and research purposes.

One of the earlier demonstrations of the continuous data-assimilation technique was presented by Charney et al. (1969). Although the "observations" inserted were actually model generated, the concept of using an atmospheric model as a "dynamic" interpolation tool for data-void regions was introduced. Furthermore, their results, and later those of Stern (1974), suggested the possibility that enough observations inserted at a sufficiently frequent interval might actually reverse the error-growth tendency in those areas without any observed data.

The particular (forward) continuous data assimilation scheme to be described here was developed at the Geophysical Fluid Dynamics Laboratory (GFDL). The feasibility of the scheme was first demonstrated with the processing of observations from the GARP (Global Atmospheric Research Program) Atlantic Tropical Experiment (GATE) (Miyakoda et al. 1976). The system was further modified in preparation for the processing of data from FGGE (First GARP Global Experiment) (Stern et al. 1985, hereafter referred to as S), which included changing to the spectral technique for the assimilation model (Gordon and Stern 1982).

Now that the original FGGE IIIb datasets have been widely used and analyzed, ways to improve the assimilation systems that produced these analyses are becoming evident. With regard to GFDL's continuous data-insertion scheme, a number of shortcomings and potential improvements have already been noted in S and Puri and Stern (1985)—hereafter referred to as PS. The major deficiencies include noisy analyses containing significant small-scale structure, the tendency to reject unbalanced mass data (especially surface pressure), and the use of an OI (optimum interpolation) first guess that could be as much as 12 h old. The current assimilation system addresses these shortcomings and, in addition, has introduced some better quality-control criteria, uses an improved spectral model with higher resolution, and provides an ex-

---

*Corresponding author address:* Dr. W. F. Stern, GFDL Princeton University, Forrestal Campus, USDC/NOAA/ERL, Route 1, P.O. Box 308, Princeton, NJ 08542.

panded archive file including a number of diabatic heating quantities.

The assimilation system is an important component in providing initial analyses for prediction studies. Hence, there is a requirement to continually improve the system, even though GFDL does not produce routine operational forecasts. In addition, GFDL has reanalyzed both FGGE special observing periods using this system and observations from the final FGGE IIb dataset.

Section 2 will look at some measures for evaluating a data-assimilation system. Section 3 will give a description of the continuous data-assimilation system currently being used, with particular emphasis on the changes and new features that have been introduced since the original FGGE data processing. Section 4 presents some test results using this system, and concluding remarks are made in section 5. The Appendix contains a discussion with regard to the archiving of diagnostic quantities, and a complete list of fields included in the diagnostic data file may be found in Table A1.

## 2. Guidelines

In developing measures of the data-assimilation system performance, it is important to consider how the analyses are to be used. Some of the primary uses of these analyses include:

1) Provide initial conditions for numerical weather predictions.
2) Provide synoptic-scale weather analyses and statistics of atmospheric circulation.
3) Provide data for evaluating the validity of atmospheric general circulation models.

Lorenc (1985) presents three principles for data-assimilation analyses that serve as a good starting point for assessing this assimilation system with regard to its output functionality. His first principle, that "the analysis must fit the observations to within their estimated observational errors," is clearly a goal that is quite compatible with all of the aforementioned uses. The second principle, that "the analyzed fields must be internally consistent, matching the structure and balance of the atmosphere," is most important for the purpose of forecast initial conditions and secondly for synoptic analysis. Difficulties in establishing this principle arise when trying to produce a full nonlinear diabatic balance in the analyses. In an intermittent assimilation system, consistency among analysis fields is generally accomplished by imposing multivariate relationships on data increments during the observation to analysis interpolation phase (usually a form of optimal interpolation or objective analysis, Gandin 1963) and by periodically applying a nonlinear normal-mode initialization (NLNMI) before beginning the next forecast first guess. This works rather well in the middle and high latitudes, but in the tropics, where the observed structure of the atmosphere is largely governed by diabatic effects, the analyses provide a poorer match.

At some centers, where intermittent data assimilation is used, diabatic effects have been included in the NLNMI. This has improved the retention of divergence in the tropical analyses, but the inability to include the actual model condensation processes in the initialization and to the extent that the model diabatic heating is imperfect remains a weakness. The continuous system attempts to achieve a high-order consistency among the meteorological fields by iteratively involving the model and analysis components continuously (usually every time step) in the assimilation process. In this case the analyses will reflect more of the model's spunup circulation features. How closely the analysis structure matches the real atmosphere may depend significantly on the model validity, especially in the relatively data-poor tropics. Furthermore, since the generation of a spuriously large amount of gravity-wave activity is a problem associated with the repeated insertion of data, the extent to which this noise is controlled (typically via normal-mode initialization or a form of divergence damping) represents additional compromising constraints on the system.

The need for the model to be able to reliably define the state of the atmosphere, especially in regions void of data, leads to Lorenc's third principle for assimilation: "The analysis must be near the forecast based on earlier observations, unless current observations indicate otherwise." Hollingsworth et al. (1986) have further quantified this principle by showing that short-range forecast errors are comparable to observational errors in well-observed areas, which implies that most of the change to the state of the atmosphere in a well-behaved assimilation system should be accounted for in the forecast first guess. Furthermore, it would be expected that the atmospheric states produced by the assimilation remain reasonably close to a slow manifold. Representing the magnitude of incremental changes for the three major components of an assimilation system by $F$, $A$, and $I$ (forecast, analysis, initialization), it follows that $F > A > I$ in a well-behaved, internally consistent assimilation system.

These guidelines, along with concerns for correcting the shortcomings indicated in GFDL's original FGGE assimilation system, have served as considerations in the assimilation system design and performance.

## 3. System design

As indicated earlier, the current data-assimilation system configuration is the result of a continuing evolutionary process. Despite some short comings, both the original system established for the processing of data from GATE (Miyakoda et al. 1976) and the revised system used to process the original FGGE data (S) produced reasonable analyses. Hence, the focus

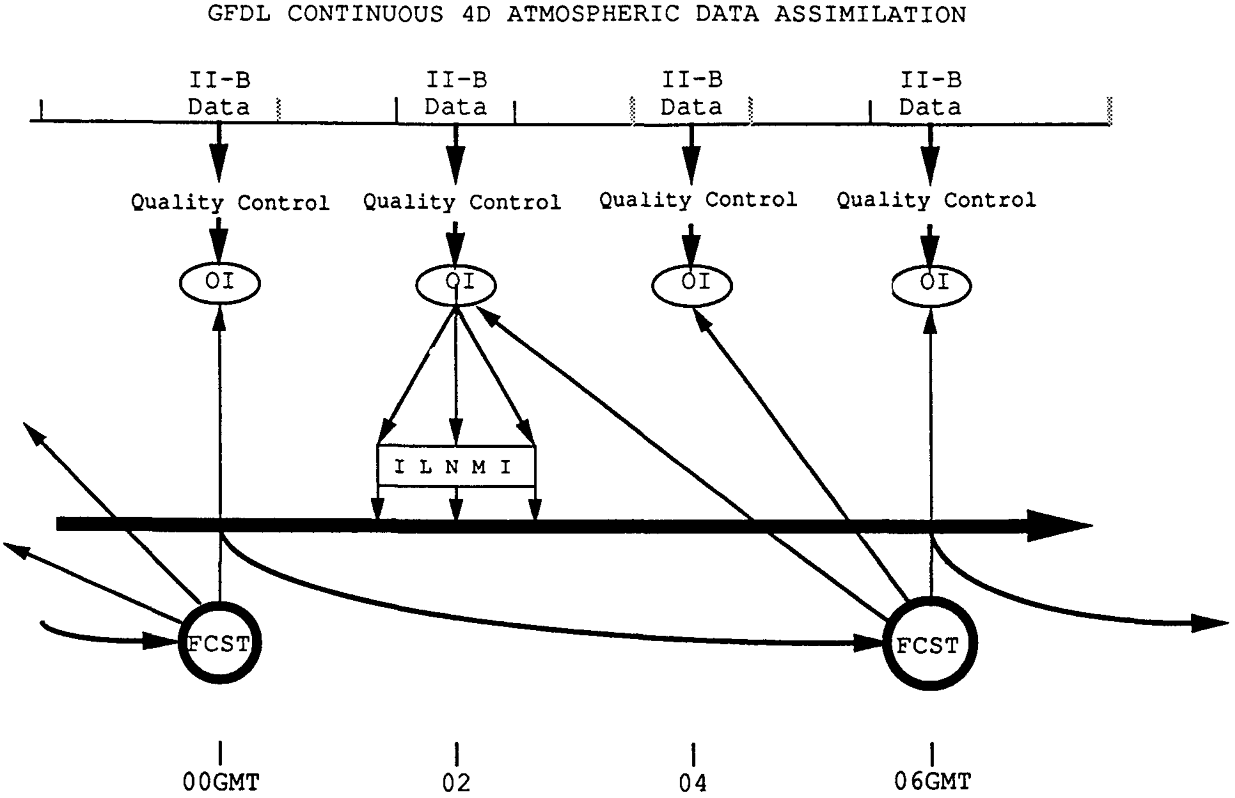


FIG. 1. Schematic overview of GFDL's four-dimensional atmospheric data assimilation. The major components are illustrated as follows: data preprocessing and quality control, optimal interpolation (OI), initialization (ILNMI), dynamic assimilation (bold line with arrow from left to right), and the 6-h forecast to determine a first guess (thin line with arrow from left to right). (Note that ILNMI is illustrated at only one time level but is actually applied at each assimilation time step.)

here is on justification and motivation of those new features in the current system.

The GFDL data-assimilation system is comprised of three main phases, as in the original FGGE system: data preprocessing, dynamic assimilation, and initialization. However, a number of significant modifications have been introduced, including the reduction of the assimilation cycle from 12 to 6 h and using a 6-h forecast to serve as a first guess during the preprocessing, increased horizontal resolution of the assimilation model, and the inclusion of an incremental linear normal-mode initialization (ILNMI) [Daley and Puri (1980) and PS], eliminating the need for the intermittent nonlinear normal-mode initialization. Figure 1 schematically shows the system overview.

### *a. Preprocessing*

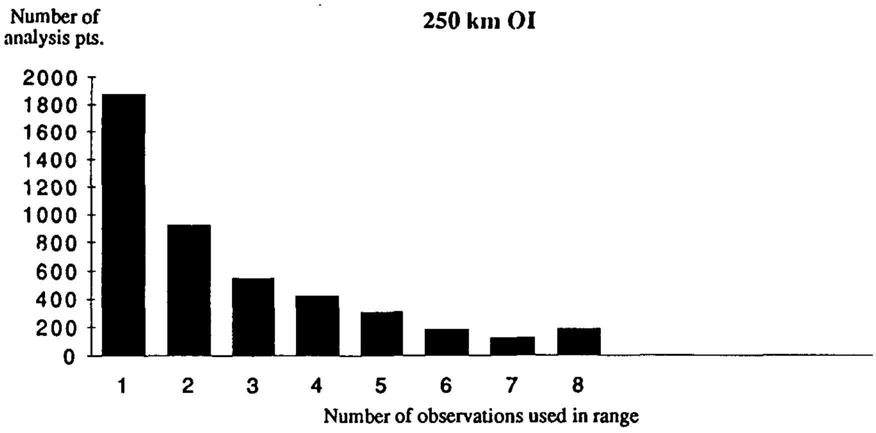
The sorting and quality-control procedures involved with preprocessing the level IIb data are only slightly changed from the original FGGE system (see S). The main difference is an increase in the data window to ±1.5 h from ±1.0 h, with the analysis times still centered about the even hours. There have, however, been two significant improvements to the optimum interpolation:

1) The extension of the observational collection range from 250 to 500 km with an increase of the maximum number of observations per analysis point used from 8 to 12.

2) Using a 6-h forecast instead of a 12-h persistence as the first guess.

These changes should at some stage be followed up by modifications to the forecast-error statistics and OI structure functions. However, at this time efforts are focused on improved consistency and data acceptance as well as reduction of noise in the analyses; therefore, a critical examination of these components of the OI will be a future consideration.

The extended OI range has been found to produce analyses that are more appropriate for the assimilating model, as well as being more consistent with observational networks (PS). By doubling the collection radius, it should be expected that the number of observations available to be used to determine an analysis value will increase on average by a factor of 4. Figure 2 shows a significant increase in the number of observations per analysis value. It can be seen that, with the 250-km range, less than 5% of the analysis values were based on the maximum of 8 observations, while now nearly 25% are based on the maximum of 12 obser-



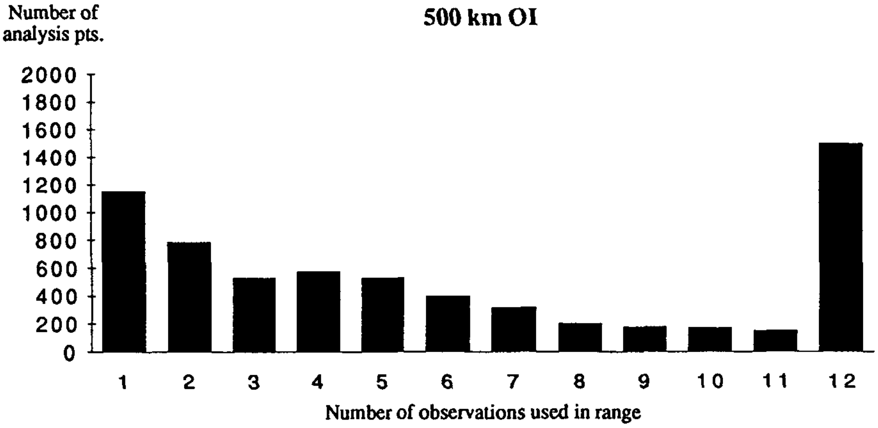


FIG. 2. Columns group analysis points by the number of observations used to determine them in the OI scheme. The top plot is for an OI with a 250-km search radius and a maximum of 8 observations used. The bottom plot is for an OI with a 500-km search radius and a maximum of 12 observations used.

vations. Furthermore, the total number of analysis points where insertion data is available has increased by more than 40%.[1] This tendency should contribute to improved quality, more consistency, and greater representativeness of the analyses.

Using the more current first guess significantly improves the background field for the OI and is consistent with one of Hollingsworth et al.'s (1986) primary criteria for a well-behaved assimilation system. This should especially help to correct the tendency to under analyze fast moving and rapidly deepening transient features, which was observed in the original FGGE analyses (S). This aspect will be examined in some detail in section 4. The use of the 6-h first guess also should allow for improved analysis quality. Since the forecast will generally account for most of the analysis change, this has permitted the use of gross rejection criteria that are significantly tighter than those used during the original FGGE processing for tossing out erroneous observations. There may be some question with regard to the use of the 6-h forecast as the first guess in the OIs performed at hours 2 and 4. Admittedly, it would probably have been most appropriate to carry three separate first-guess fields (i.e., forecast for 2, 4, and 6 h). Although this is conceptually quite feasible, it does present practical complexities, such as the need to save additional fields. Furthermore, since the resulting analyses are always at the 6-h forecast time, it was felt that by using the 6-h forecast as the first guess throughout the 6-h cycle it could help to "tug" the assimilation model solution in the proper

---

[1] The values plotted are for a single synoptic time for temperature but are indicative of the increased number of observations per analysis for all fields and analysis times.

direction. The OI vertical range of three mandatory pressure levels and the univariate structure function formulations remain the same as those used during the original FGGE processing.

### *b. Assimilation*

After a time interpolation, which uses available data at synoptic times to fill in temporal gaps at particular OI analysis points at off-synoptic times (i.e., hours 2, 4, 8, 10, 14, 16, 20, and 22),[2] the data is inserted into the assimilation model. Where insertion data exists and passes a final quality control, the model solution is updated by a weighted combination of the model solution and the OI insertion data. (It should be noted that the selection criteria for the acceptance of moisture insertion data at this stage have been significantly tightened. The moisture acceptance criteria used in the original FGGE processing were unrealistically large; the current values are based on realistic extremes.) The computation of the weights is the same as that described in S, but the model solution is allowed to feed back during the weighted insertion process instead of only at the beginning of each new data-insertion time (every 2 h). Hence, the updated model solution $\phi(\tau, u)$ for time step $\tau$ may be expressed as

$$\phi(\tau, u) = (1 - \omega)\phi(\tau, m) + \omega\phi(\tau, i)$$

where $\phi(\tau, m)$ is the model solution at time step $\tau$, $\phi(\tau, i)$ is the insertion value for a 2-h data block, and $\omega$ is the insertion weight for the 2-h data block. (Although new insertion data is inserted every 2 h, the data window has been increased to 3 h. This overlap allows for increased temporal consistency.)

This repeated insertion technique coupled with ILNMI at each time step distinguishes this system from others. There is some philosophical similarity to the United Kingdom Meteorological Office (UKMO) continuous data-assimilation scheme described by Lorenc et al. (1991). Both systems use repeated insertion techniques to iteratively converge toward a compromise between current observations and model forecast. In the UKMO system a simple analysis is performed iteratively with model forecast steps, while in the GFDL system the OI analysis is not involved in the iterative process. However, the ILNMI provides a method of adjusting the analyses consistent with the slowly varying modes of the model in an iterative manner. The process of moving the model solution toward the insertion data values is shown graphically in Fig. 3 for the 12-h period 0000–1200 UTC 5 May 1979. The rms (root-mean-square) differences between the model solution and the insertion values have been computed for temperature and surface pressure at all analysis points that involved observed data. Plots for the original FGGE system are compared to the current system. The spikes reflect the introduction of new insertion data. The overall reduction of the minimum difference levels (~0.4° to 0.6°C vs ~0.7° to 1.0°C previously for temperature and ~1.8 vs ~3.5 hPa previously for surface pressure) indicate a greater ability to assimilate data and more data acceptance in the current system.[3]

The original FGGE system did not allow data insertion in the top two model levels, questioning the accuracy and availability of observations at 20 hPa and above. As it turned out, however, the model's top-level climate drift toward radiative equilibrium resulted in upper-level analyses with abnormally strong polar night jets and extremely cold winter poles. This also forced the use of a very short time step during the model assimilation runs in order to avoid CFL (Courant–Friedrich–Lewy) criterion violations. Tests allowing insertion throughout the entire model vertical structure have shown considerable improvement in upper-level analyses. Presumably, available satellite temperatures are well accepted by the model, resulting in a winter polar temperature that is much more reasonable, reducing both the north–south temperature gradient and the overly intense polar night jet (see Fig. 4). It seems that the longer time and space scales associated with the stratospheric circulation features enhance the value of mass information (i.e., satellite temperatures), as would be anticipated by geostrophic adjustment theory.

The model used for data assimilation has also been improved since the original FGGE processing. The horizontal resolution has been increased from 30 to 42 waves (rhomboidal truncation), while still retaining the 18-vertical-level structure. The GFDL E4 physical parameterization package has been added (see Gordon and Stern 1982). The major feature is the Mellor–Yamada level 2.5 turbulence closure scheme for vertical transport of momentum, heat, and moisture. In addition, a parameterization for mountain gravity-wave drag has been included (Pierrehumbert 1987). This parameterization has significantly reduced the model's systematic zonal-wind error throughout the troposphere and lower stratosphere in the midlatitudes of the Northern Hemisphere, while there is very little change to the Southern Hemisphere's zonal-wind bias.

### *c. Initialization*

An undesirable effect of continuous data insertion is the excitation of spurious fast modes that can lead to noisy analyses. The NLNMI has proven quite effec-

---

[2] If an OI insertion value cannot be provided at an off-synoptic time but is available at both the earlier and later synoptic hours, a linear interpolation in time is performed to obtain a value.

[3] Note: Root-mean-square statistics do not include differences in the temperatures at the top two levels.

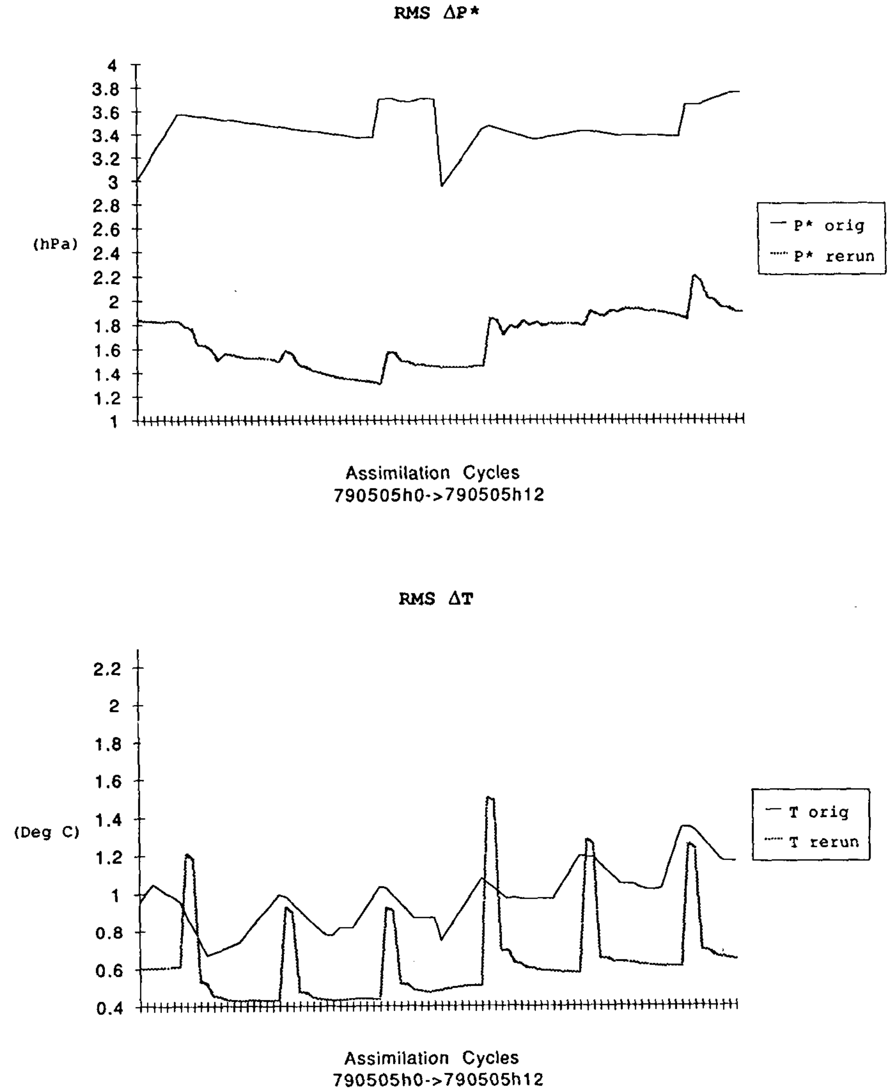


FIG. 3. Global rms differences between the model solution and insertion values computed for all points that have observed insertion data, shown for surface pressure ($\Delta P^*$) and temperature ($\Delta T$). Black curves are for the original FGGE system and the gray curves are for the current system used in the rerun of FGGE.

tive in controlling fast-mode growth (Machenhauer 1977). However, distinguishing those fast modes that are spurious from those that are real circulation features is a difficult but important challenge for any data-assimilation system. In this regard, while in the domain of the model's normal modes, NMI provides the capability for specific selectivity with respect to frequency—that is, the ability to choose to initialize particular gravity modes while not affecting the Rossby modes.

In the earlier applications of both linear and nonlinear NMI, the usual choice was to essentially eliminate the fast modes by initializing all of the gravity waves without any modification to the Rossby modes. This was found to work reasonably well in the extratropics but greatly diminished the tropical divergent circulations. In the original FGGE system a selective NLNMI was applied periodically every 6 h, using a frequency and vertical mode cutoff (see S and PS). Although this intermittent NLNMI does keep the noise from dominating the analyses, there are some uncomfortable aspects to this initialization technique within the context of continuous data assimilation. First, the changes due to the initialization are significant when compared to the changes due to new insertion data; this is inconsistent with a previously noted measure of a well-behaved assimilation system (Hollingsworth et al. 1986). This may be seen in the top of Fig. 3 where

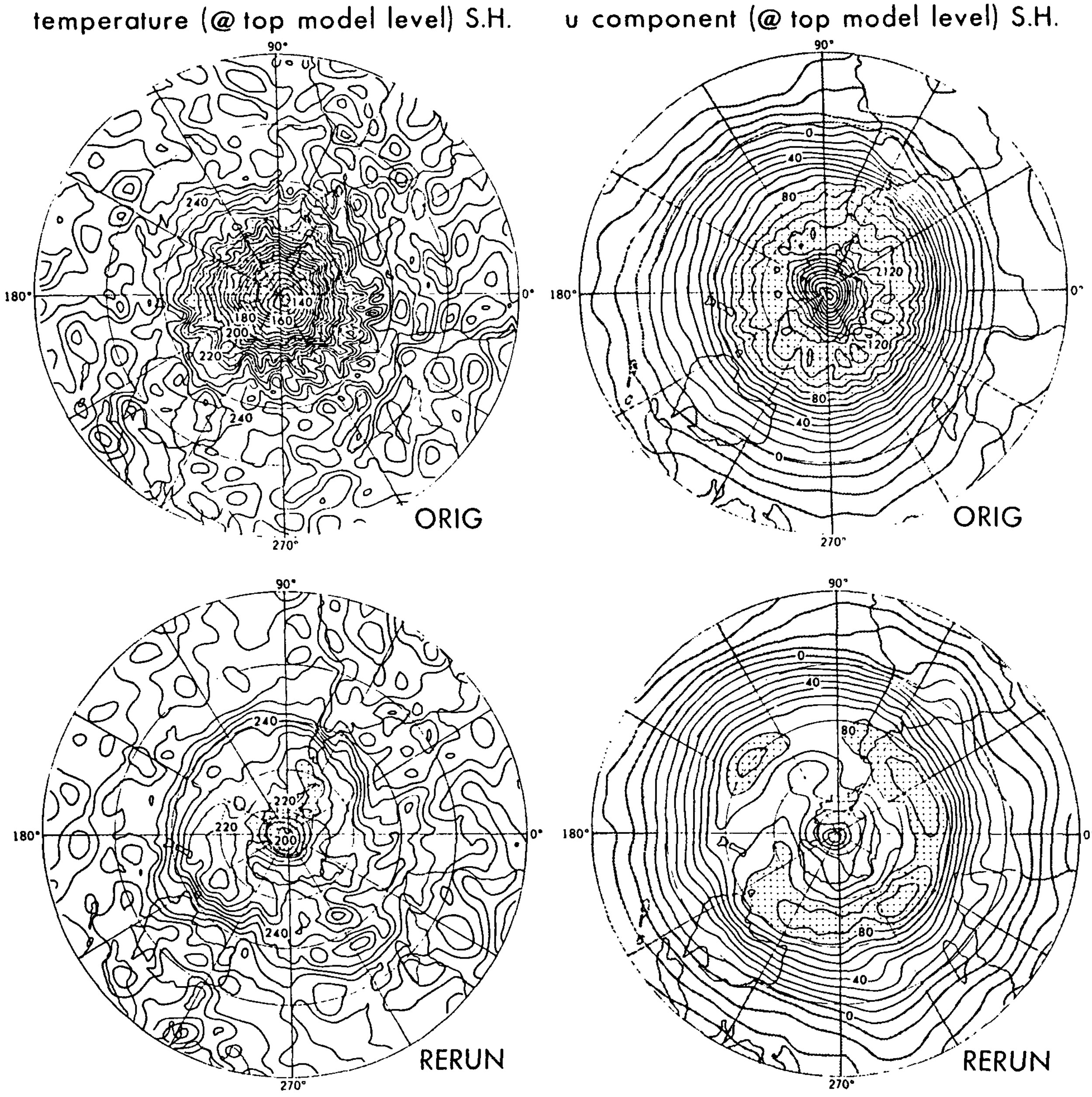


FIG. 4. Southern Hemisphere maps of temperature (left) and zonal-wind (right) analyses at the top model level (2.2 hPa) for the original FGGE system (top) versus the system used for the FGGE rerun (bottom).

large changes in the rms differences between the model solution and the insertion values occur every 6 h, corresponding to the NLNMI steps. Furthermore, some noise will be generated each time step as data increments are inserted into the model.

In order to keep the analyses closer to a diabatically balanced state throughout the assimilation cycle, an ILNMI has been adopted as part of this data-assimilation scheme, and the periodic NLNMI has been eliminated. It keeps only that part of the data increments that projects onto the Rossby modes and slower gravity modes; that is, those gravity waves whose structures are determined by vertical modes 1–4 and whose periods are less than 6 h are adjusted, while retaining the full model-produced circulations. The motivation for choosing which gravity modes to adjust was based on results of tests with the FGGE assimilation system (S and PS) as well as the findings of Puri and Bourke (1982). The ILNMI is applied every time step before the analyses are archived. Although the initialization

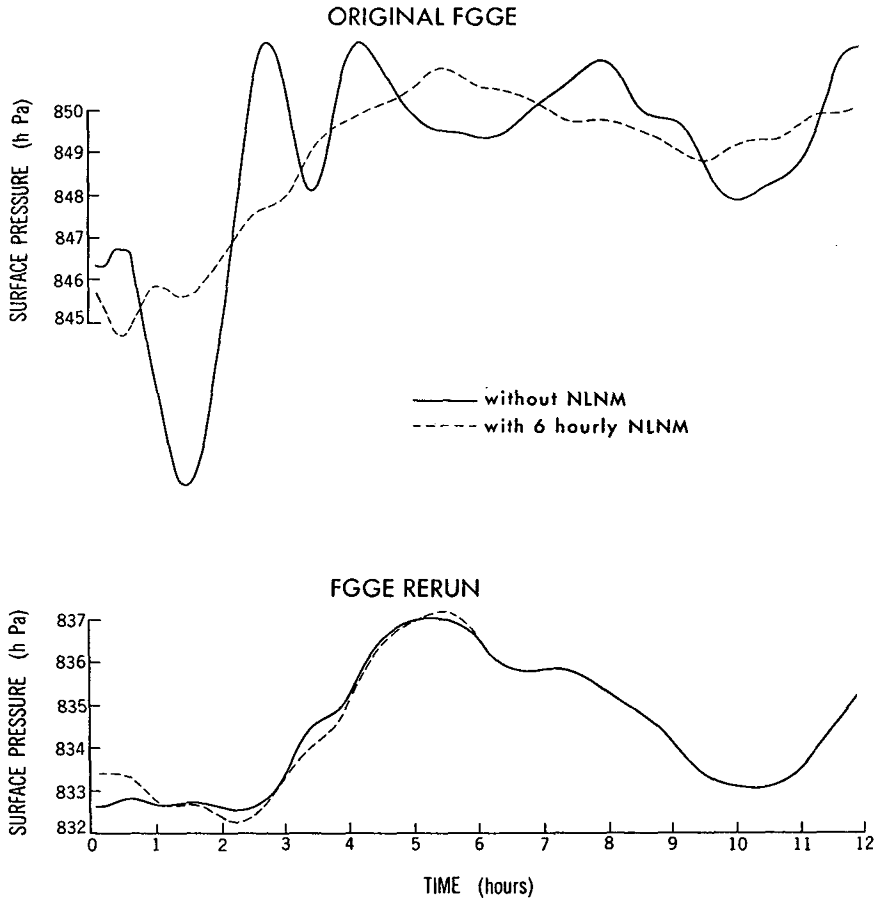


FIG. 5a. Surface pressure traces at an extratropical point during 12-h forecasts run from an original FGGE analysis (top) and from a comparable FGGE rerun analysis (bottom). The solid curves show forecasts from the assimilation system analyses, and the dashed curves show the impact of applying nonlinear normal-mode initialization to the analyses before running the forecasts.

is applied only to data increments and is linear, spurious fast-mode activity appears to be controlled quite well. Since the model solution should remain close to a slow manifold and the data increments are generally small, the linear initialization is sufficient to control the high-frequency gravity modes. This is shown in Fig. 5a, which presents 6- and 12-h surface pressure traces at an extratropical location, and in Fig. 5b, for a tropical location, for four forecasts from the same analysis time but from four different assimilation-system configurations. The comparison is for the original FGGE system, the original system with the NLNMI performed just after analysis time, the present system, and the present system plus a 6-h periodic NLNMI. It can be seen that the current system with or without NLNMI does quite well at suppressing high-frequency oscillations in the subsequent forecast, but this was not the case for the original FGGE analyses. The ILNMI appears to be a well-suited, effective technique of controlling spurious fast modes generated during continuous assimilation, while allowing the model to achieve a spunup diabatic solution.

## 4. System performance

In this section some test results produced by this continuous four-dimensional data-assimilation system will be presented. The primary focus will be to assess the extent to which the current system has corrected some of the shortcomings noted in the original FGGE processing system. Specifically, how system changes have improved the representation of rapidly changing atmospheric features and the extent to which mass data acceptance and noise reduction is evident. In addition, some discussion with regard to how well analyses fit observations, model systematic bias in the analyses, and understanding heating tendencies during assimilation seems appropriate.

One challenging case to be properly analyzed by a global data-assimilation system is the Presidents' Day

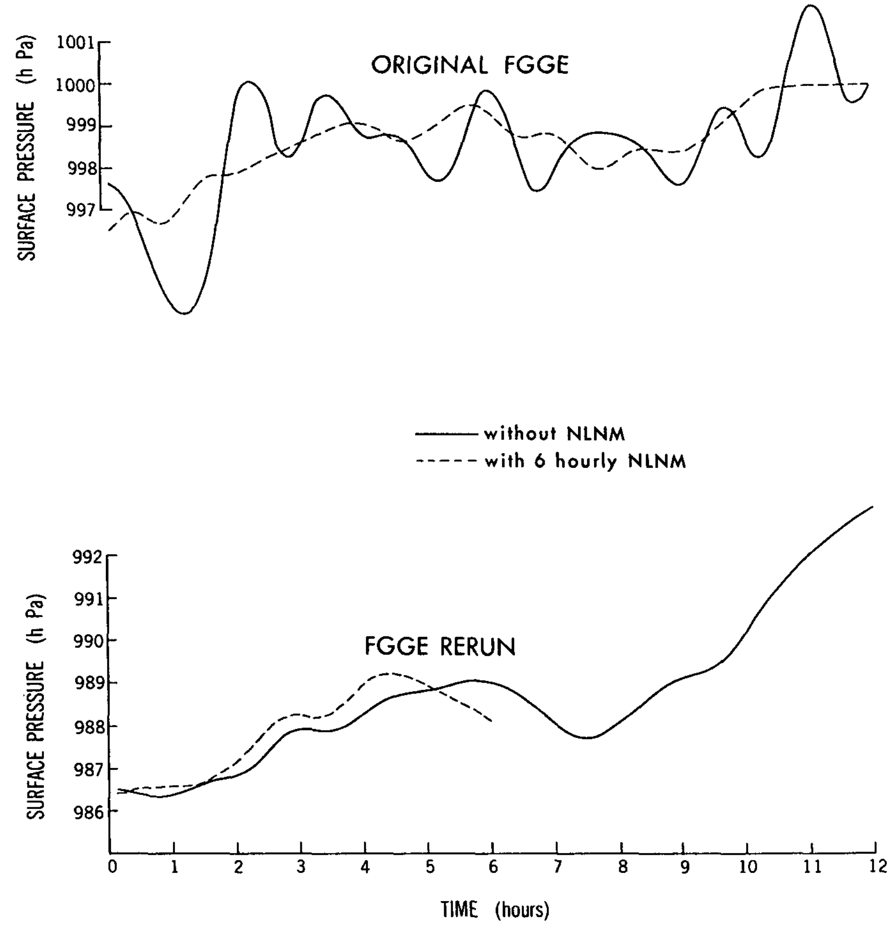


FIG. 5b. Same as Fig. 5a except at a tropical point.

storm, which developed and deepened off the United States east coast on 19–20 February 1979. This storm has been studied extensively because it produced extreme snowfall totals in the United States mid-Atlantic region and was poorly forecast (see Bosart 1981). Hollingsworth et al. (1985) studied the impact of analysis differences on the ability to accurately forecast this event. The focus here is on those aspects of the new GFDL data-assimilation system that allowed for a much improved analysis of this storm's location and circulation when it was off the United States east coast. What made this case particularly difficult was a combination of rapid cyclogenesis (originating mostly at low levels), fast movement, and the location of the storm over the ocean. In this case the original FGGE system was severely handicapped by its use of persistence as a first guess and the limited influence region for the OI (250 km). Hence, not only did this leave gaps in the OI analysis insertion data field, but it was significantly weighted by a first guess, which was nearly 12 h old. The results was a storm that was far too weak and too slow. This may be seen in Fig. 6, which shows observations (top), insertion data (middle), and analyses (bottom) for 1000-hPa wind at 0000 UTC 20 February 1979. The insertion data from the original FGGE system (middle left) defines a circulation that lacks organization and strength and is analyzed generally to the southwest of the observed center (indicated by a "+"). The subsequent analysis (bottom left) is more organized but remains retrograded and too weak. In contrast, the new system (right side) shows a stronger, more organized, and more properly positioned circulation in both the insertion data and analysis, primarily the result of the 6-h forecast first guess and the increased OI influence region (500 km). Furthermore, in the reanalysis of sea level pressure (not shown) a significantly deeper low center develops (approximately 1000 vs 1010 hPa), indicating a greater acceptance of mass data, which was anticipated based on the results plotted in Fig. 3.

The issue of noise reduction in the analysis system has been addressed in the earlier section on initialization and in PS. Evidence of the reduced level of noise in analyses produced by the current system when com-

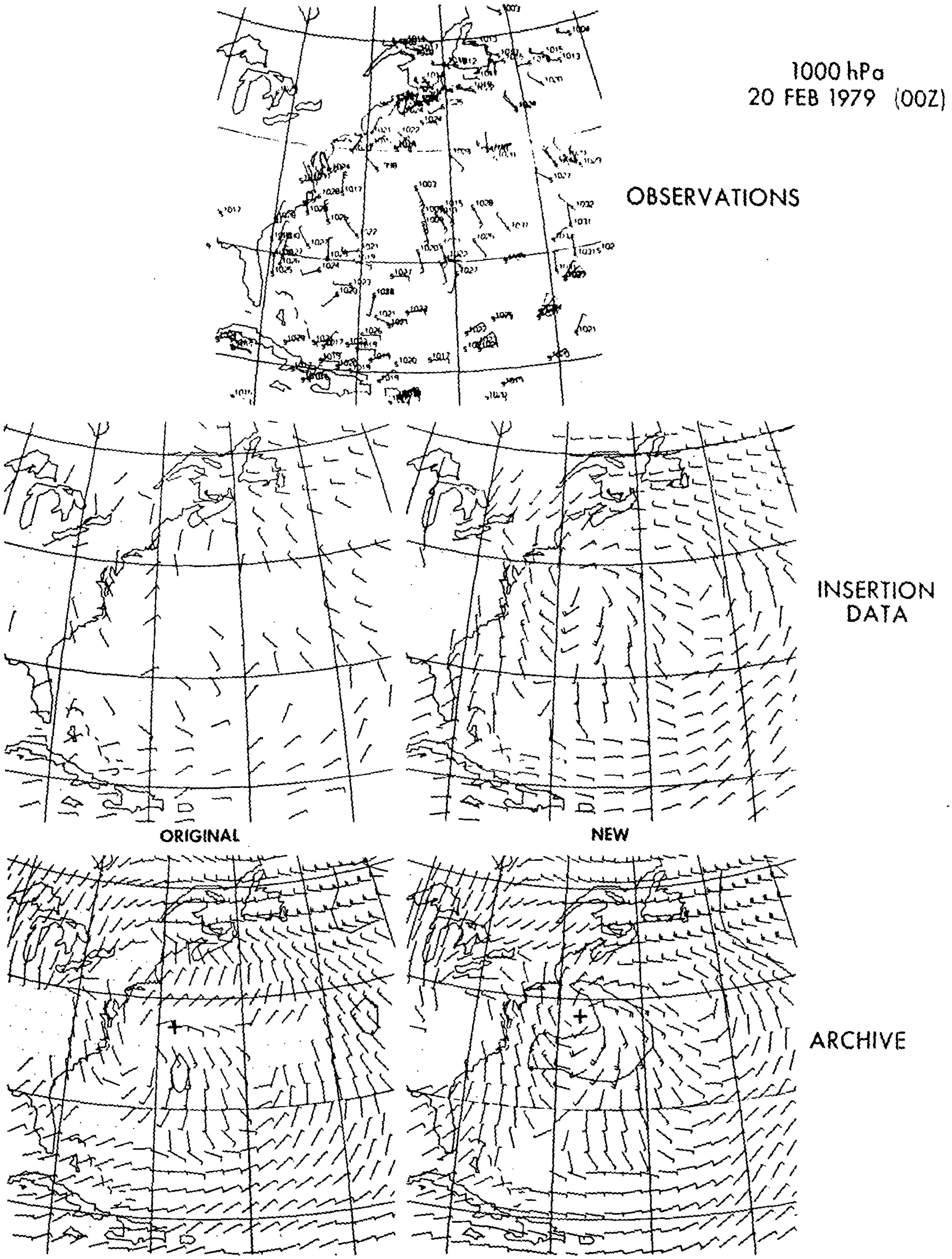


FIG. 6. Maps of the region of interest for the President's Day storm showing observations (top), OI insertion data (middle), and final analysis values (bottom) for 1000-hPa wind at 0000 UTC 20 February 1979. Insertion data and final analysis from the original FGGE system is shown on the left and the FGGE rerun system on the right.

pared to the original FGGE analyses has been presented in Fig. 5. In addition, a visual inspection of either instantaneous or time-averaged horizontal analyses reveals far less high-wavenumber structure in the FGGE reanalyses.

The importance of physical processes in determining aspects of the general circulation is obvious. In an attempt to enhance our understanding of the relative roles of various processes, the major components of diabatic heating from the assimilating model have been included in the FGGE reanalysis. As indicated previously, these heating rates represent averaged tendency contributions from the 6-h forecast to produce the first guess. Because this system assimilates data as the model is integrating, it is also possible to obtain diabatic tendency contributions during this data-assimilation cycle. Since the model atmosphere is being continuously updated with observed insertion data, the tendencies obtained during this phase of the assimilation cycle should be very much like tendencies obtained from a collection of single–time-step runs. Klinker and Sardeshmukh (1987) showed that a large number of single–time-step integrations may be useful in diagnosing systematic errors of numerical models. Hence, a comparison of these sets of diabatic tendencies may also prove useful as a model-bias diagnostic tool, although a clear separation of model bias from analysis bias is not possible using only these products.

Figure 7a shows the net diabatic heating from the 6-h data-assimilation cycle (top), the 6-h forecast first-guess integration (middle), and a monthly forecast (bottom). These tendencies have been time averaged for the period 6 January–4 February 1979. The most significant region of net diabatic cooling in the assimilation heating diagnostics is seen at all latitudes, ranging from about 800 hPa in the extratropics up to approximately 600 hPa in the tropics. This distinctive cooling band is associated with longwave cooling from the zonal mean low cloud tops. In the monthly forecast this cooling is not nearly as apparent, since the much longer integration time allows the model to adjust (but the resulting balanced model structure has a significant midtropospheric cold bias). Two regions of significant diabatic heating are seen in the top two panels of Fig. 7a. One is confined to the tropical midtroposphere and is associated with model-generated convective activity, and the other net heating region appears near the lower boundary, extending across much of the extratropics and tropics. The midtropospheric tropical heating is stronger in the forecast first guess than in the data-assimilation cycle. This is apparently the result of more convective activity and, therefore, more latent heat release, as seen in a comparison of the top two panels of Fig. 7b (which is identical to Fig. 7a except that latent heating is displayed). Furthermore, the adiabatic heating fields (not shown) are much more consistent with the diabatic heating structure in the first guess, indicating stronger and more organized vertical-motion fields (i.e., a more spunup solution). The heating near the lower boundary is significantly stronger during the assimilation cycle than in the 6-h forecast. The difference is almost entirely the result of increased low-level latent heat release (primarily large-scale condensation), as indicated in the top two panels of Fig. 7b. This implies that the model's planetary boundary layer wants to drift toward a warmer (and perhaps drier) state relative to the observed atmosphere. It is beyond the scope of this paper to explore all of the model's biases and their potential causes. The main point to be made here is that the diagnostics during the assimilation cycle may be useful for better understanding model bias and model spinup, but they are not appropriate for general circulation studies, where measures of a more balanced atmospheric state are needed. For this reason and for better comparability with forecast–analysis assimilation systems, the diagnostics obtained from the 6-h forecast for the first guess have been archived.

It seems reasonable to revisit some of those guidelines discussed in section 2 to provide some quantitative measure of system performance. The use of a 6-h forecast as a first guess for the OI (versus a persistence used in the original system), has reduced the size of the analysis increments in the current system, which is much more consistent with the previously noted criteria for a well-behaved assimilation system. Furthermore, the better first guess then allows the analysis scheme to fit the observations with a smaller incremental change relative to the first guess, which should be easier for the assimilating model to accept. The end result leads to a final analysis that retains more of the observational information. This may be seen in Figs. 8a,b, which depict a time series of rms differences between analyses and observations that are spaced 10 days apart during the first special observing period. The analysis values have been interpolated to the station locations before the differences are calculated. These diagrams are based on 518 radiosonde stations distributed between 24° and 75°N. Figure 8a shows 850-hPa temperature, and Fig. 8b shows 200-hPa vector wind. Both figures clearly indicate that the analyses from the current assimilation system (reanalysis) fit the observations better than the analyses from the original FGGE system. It is also seen that the fit to observations in the GFDL reanalysis is comparable to the fit in the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (Uppala 1986). (It should be noted that temperature is not directly analyzed in the ECMWF assimilation system, which presumably accounts for a poorer fit to observations).

A detailed presentation of analysis products from the reanalysis of FGGE is included in Ploshay et al. (1992), where comparisons include other assimilation systems as well as the GFDL assimilation system. These results indicate that the differences between reanalyses

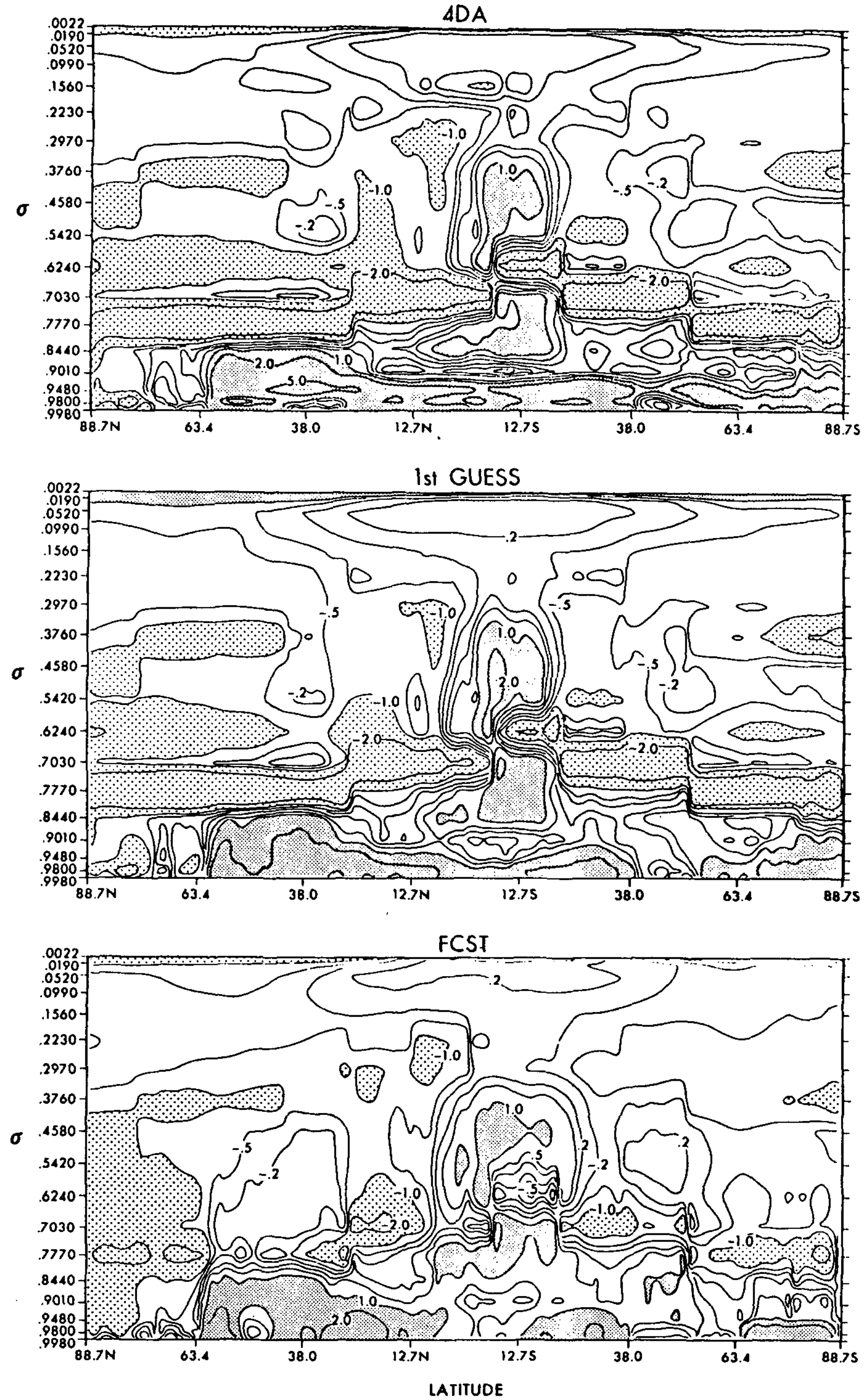


FIG. 7a. Zonal-mean latitude–height cross sections of net diabatic heating averaged for the period 6 January–4 February 1979. Values from the 6-h data-assimilation cycle (top), the 6-h forecast first-guess integration (middle), and a monthly forecast (bottom) are shown. Regions where heating rates exceed 1°C day$^{-1}$ are densely shaded, and regions where heating rates are less than −1°C day$^{-1}$ are lightly shaded.

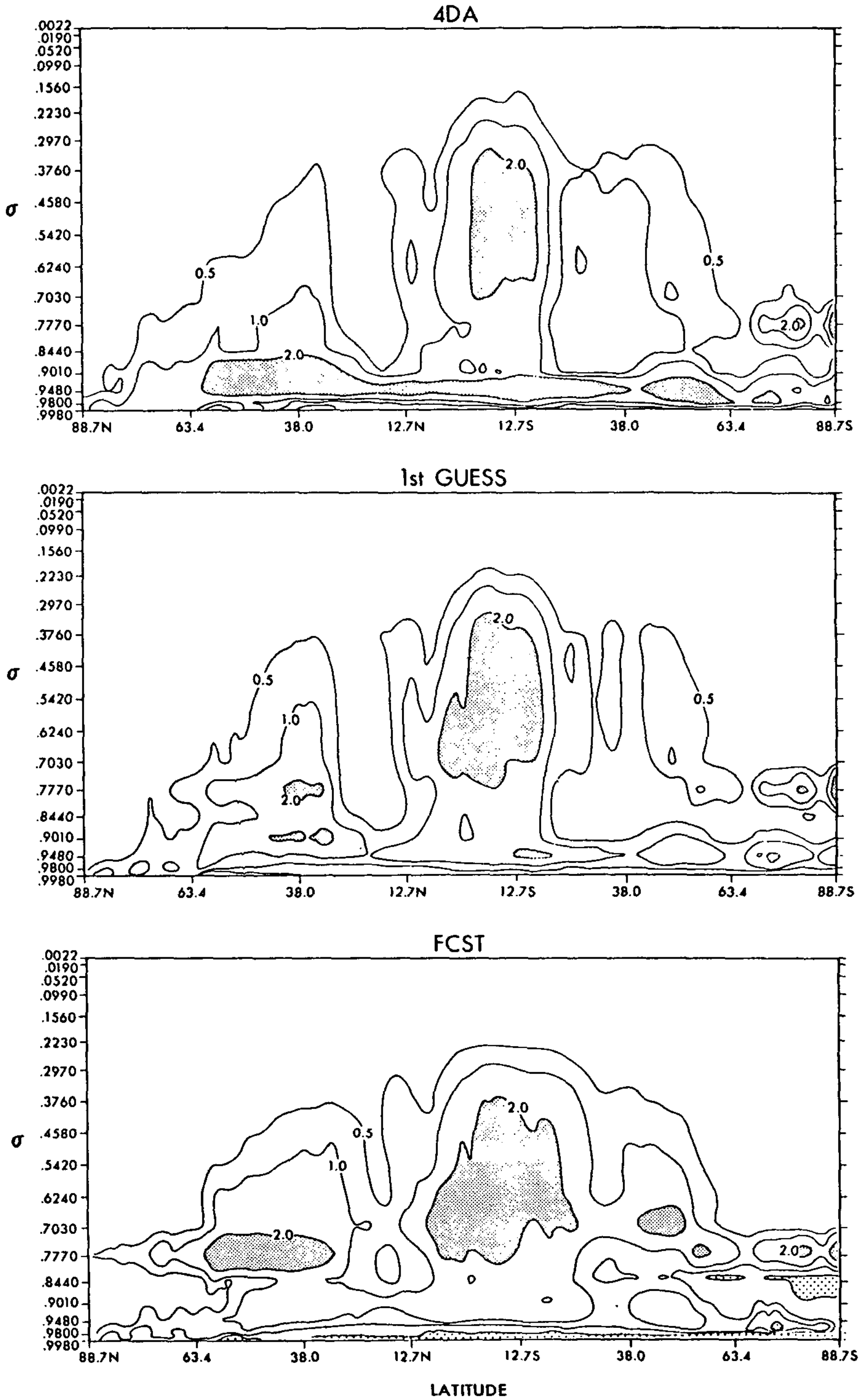


FIG. 7b. Same as Fig. 7a except for total latent heating. Regions where heating rates exceed 2°C day$^{-1}$ are densely shaded, and regions where heating rates are equal to zero are lightly shaded.

from GFDL (i.e., produced with the new GFDL data-assimilation system) and reanalyses from ECMWF are significantly smaller than corresponding differences in the original FGGE analyses.

## 5. Concluding remarks

Preliminary analyses produced with the post-FGGE system indicate improvement with regard to some of

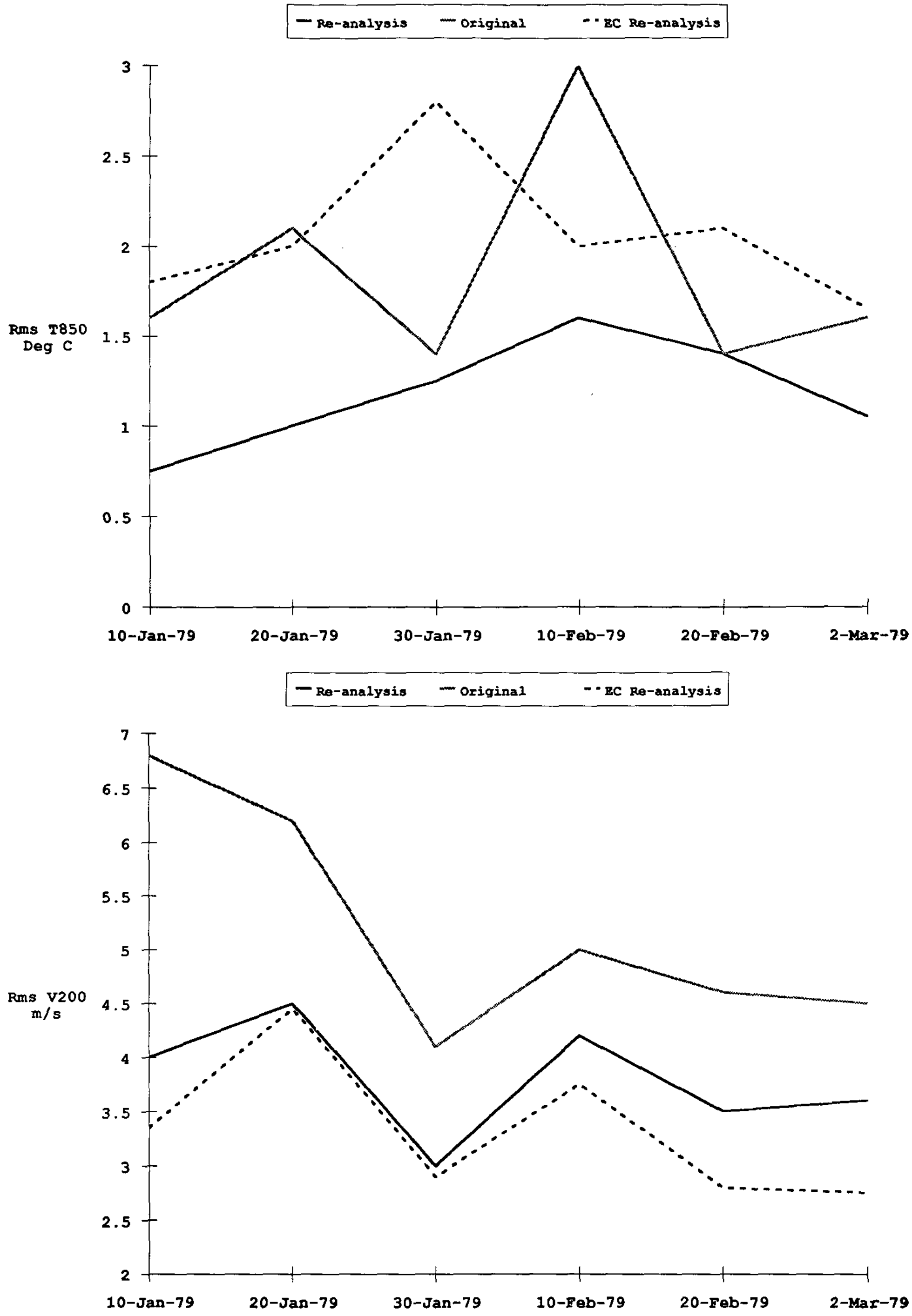


FIG 8. (a) Northern Hemisphere extratropical rms differences between analysis values and observations versus time during the first special observing period for temperature at 850 hPa. Values for the original FGGE analysis are indicated via the gray curve, and values from the reanalysis are indicated via the solid black curve, and values for the ECMWF reanalysis are indicated via the dashed black curve. (b) Same as (a) except for vector-wind magnitude at 200 hPa.

the shortcomings seen in the original GFDL FGGE analyses. Specifically, there is considerably less noise apparent, and the analyses also show a greater degree of consistency between mass and momentum, which is primarily the result of the ILNMI at each time step. In addition, the ability to resolve rapidly moving and deepening transient features has been improved. There is also some indication of less rejection of surface pressure information, although this still remains an area for further study.

Without significantly altering the current system structure, future work should include improvements to the optimum interpolation, particularly with regard to the first-guess error covariance structure functions, which should be updated to reflect the current 6-h forecast.

Investigations involving more radical system design changes may include a look at combining analysis and initialization and more appropriate temporal weighting of insertion data that might employ variational continuous assimilation and adjoint techniques (Derber 1989).

*Acknowledgments.* The authors wish to thank Kikuro Miyakoda for his helpful discussions and feedback. In addition we would like to thank Robert Smith who helped in the processing of the analysis results and Phil Tunison and his staff for drafting the figures.

## APPENDIX

### Diabatic Diagnostics

In addition to the archiving of the standard prognostic quantities, as in the original FGGE format (Ploshay et al. 1983), a file containing diabatic heating tendencies, flux, and boundary quantities is being included, as shown in Table A1. These diagnostics are accumulated and averaged during the 6-h forecasts to produce the first-guess fields. The sum of diabatic heating tendencies $Q_d$ may be defined as follows:

$$Q_d = Q_r + Q_c + Q_s, \tag{A1}$$

where $Q_r$ is the sum of short- and longwave radiative heating, $Q_c$ is the total condensational heating, and $Q_s$ is the turbulent sensible heating. Since the total heating (net temperature tendency $Q_t$) is also archived, the adiabatic heating may be defined as,

$$Q_a = Q_t - Q_d. \tag{A2}$$

TABLE A1. List of quantities included in the diagnostic data files archived for the SOPs during the FGGE year analysis.

| Archived diagnostic file |
|---|
| Radiative: |
| Shortwave heating—all levels |
| Longwave heating—all levels |
| Net heating—all levels |
| Upward longwave flux—top |
| Net longwave flux—bottom |
| Downward longwave flux—bottom |
| Upward longwave flux—bottom |
| Incoming solar flux—top |
| Outgoing solar flux—top |
| Downward solar flux—bottom |
| Upward solar flux—bottom |
| Moist condensation: |
| Convective heating—all levels |
| Total heating (moist convection and large-scale conditions)—all levels |
| $\partial q/\partial t_{\text{convective}}$—all levels |
| $\partial q/\partial t_{\text{total}}$—all levels |
| counter for $\partial q/\partial t_{\text{convective}}$—all levels |
| Turbulent sensible heating—all levels |
| Total heating—all levels |
| Accumulated precipitation: |
| Convective |
| Total |
| Other surface quantities: |
| Latent heat flux |
| Sensible heat flux |
| Wind stresses |
| Soil moisture |
| Snow cover |
| Albedo |
| Surface pressure |
| Topography |
| Subsurface: |
| Ground temperature—5 cm |
| Ground temperature—50 cm |
| Ground temperature—5 m |

These diagnostics will generally be archived at 6-h intervals (along with the prognostic quantities) and represent time averages for that period with the exception of precipitation, soil moisture, and snow cover being accumulated and albedo, topography, and ground temperatures being values at the time of archiving.

## REFERENCES

Bosart, L. F., 1981: The Presidents's Day snowstorm of 18–19 February 1979: A subsynoptic-scale event. *Mon. Wea. Rev.*, **109**, 1542–1566.

Charney, J., M. Halem, and R. Jastrow, 1969: Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, **26**, 1160–1163.

Daley, R., and K. Puri, 1980: Four-dimensional data assimilation and the slow manifold. *Mon. Wea. Rev.*, **108**, 85–99.

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

Gandin, L., 1963: *Objective Analysis of Meteorological Fields.* Gidrometeorologichesckoe Izdatel Stvo, 242 pp.

Gordon, C. T., and W. F. Stern, 1982: A description of the GFDL global spectral model. *Mon. Wea. Rev.*, **110**, 625–644.

Hollingsworth, A., M. S. Tracton, K. Arpe, G. Cats, S. Uppala, and P. Kallberg, 1985: The response of numerical weather prediction systems to FGGE II-b Part I: Analyses. *Quart. J. Roy. Meteor. Soc.*, **111**, 1–66.

——, D. B. Shaw, P. Lonnberg, L. Illari, K. Arpe, and A. J. Simmons, 1986: Monitoring of observations and analysis quality by a data assimilation system. *Mon. Wea. Rev.*, **114**, 861–879.

Klinker, E., and P. D. Sardeshmukh, 1987: The diagnosis of systematic errors in numerical weather prediction models. *Report at Seminar/Workshop on Diabatic Forcing,* ECMWF, Reading, U.K., 209–244.

Lorenc, A., 1985: Data assimilation by repeated insertion into a forecast model—principles, practice, problems and plans. *Report at Seminar/Workshop on data systems and observing system experiments with particular emphasis on FGGE,* ECMWF and WMO, Reading, U.K., 191–214.

——, R. S. Bell, and B. MacPherson, 1991: The Meteorological Office analysis correction data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **117**, 59–89.

Miyakoda, K., L. Umscheid, D. H. Lee, J. Sirutis, R. Lusen, and F. Pratte, 1976: The near real-time global four-dimensional analysis experiment during the GATE period, Part I. *J. Atmos. Sci.,* **33,** 561–591.

Pierrehumbert, R. T., 1987: An essay on the prameterization of orographic gravity wave drag. *Report at Seminar/Workshop on Orographic Effects,* ECMWF, Reading, U.K., 251–282.

Ploshay, J. J., R. White, and K. Miyakoda, 1983: FGGE level III-b daily global analyses. Part I, II, III, IV. NOAA Data Report ERL GFDL-1, 278 pp.

——, W. Stern, and K. Miyakoda, 1992: FGGE reanalysis at GFDL. *Mon. Wea. Rev.,* **120,** in press.

Puri, K., and W. Bourke, 1982: A scheme to retain the Hadley circulation during non-linear normal mode initialization. *Mon. Wea. Rev.,* **110,** 327–335.

——, and W. F. Stern, 1985: Investigations to reduce noise and improve data acceptance in the GFDL 4-dimensional analysis. *Report at Seminar/Workshop on Data Systems and Observing System Experiments with Particular Emphasis on FGGE,* ECMWF and WMO, Reading, U.K., 157–190.

Stern, W. F., 1974: Updating experiments with a simple barotropic model. *J. Atmos. Sci.,* **31,** 134–141.

——, J. J. Ploshay, and K. Miyakoda, 1985: Continuous data assimilation at GFDL during FGGE. *Report at Seminar/Workshop on Data System and Observing System Experiments with Particular Emphasis on FGGE,* ECMWF and WMO, Reading, U.K., 125–156.

Uppala, S., 1986: The assimilation of the final level II-b data set at ECMWF, Part I. Preprints, *National Conf. on the Scientific Results of the First Garp Global Experiment,* Amer. Meteor. Soc., Miami, 24–31.