CONTINUOUS DATA ASSIMILATION AT GFDL DURING FGGE[1]

W. F. Stern, J. J. Ploshay, K. Miyakoda

Geophysical Fluid Dynamics Laboratory/NOAA
Princeton University
Princeton, New Jersey 08542 USA

1. INTRODUCTION

During the year of 1979, FGGE (First GARP Global Experiment) provided an unprecedented amount of meteorological data from a variety of sources. One aspect of FGGE was to construct analyses based on these observations. Two centers were involved in producing gridded analyses on a non-operational basis (so-called "Level III-B"). They are the European Centre for Medium-Range Weather Forecasts (ECMWF)(Bengtsson et al., 1982), and the Geophysical Fluid Dynamics Laboratory (GFDL). Four-dimensional analysis techniques were devised for this purpose.

The history of four-dimensional analysis at GFDL dates to 1968 when development began. It was first used in a near real-time mode for the processing of observations collected during the GARP Atlantic Tropical Experiment (GATE)(Miyakoda et al., 1976), at which time the viability of continuous dynamical data assimilation was demonstrated. The GFDL 4-D assimilation system continued to evolve into the configuration as described in this report for use in the processing of FGGE data. Throughout the development, some problems have been noticed mostly associated with the excessive generation of gravity waves during insertion of data into the model. However, the basic philosophies motivating this continuous data assimilation system have remained

---

[1]Some portion of this paper was presented at the Sixth Conference on Numerical Weather Prediction at Omaha, Nebraska, USA in June, 1983.

essentially intact and may be summarized as follows:

1. A "state of the art" general circulation model is used to fill in data gaps, given an incomplete data history, and to maintain an internal consistency among the variables.
2. Optimum interpolation is employed so as to provide a means of obtaining a best fit of observations onto the model's grid that remains faithful to the observations and minimizes constraints that might conflict with the model's balance.

With this commitment, the temptation to switch to the more conventional forecast-analysis cycle,so as to bypass some of the difficulties of continuous data insertion, has been avoided. In addition, it is felt that the dynamic assimilation technique has advantages including: (1) more properly incorporating asynoptic data and (2) reducing error growth in the model solution.

2. DESCRIPTION

The GFDL 4-D assimilation system consists of three main phases which may be designated as: data pre-processing, dynamic assimilation and initialization. They are shown schematically in Figure 1.

2.1 Preprocessing

The first part of the pre-processing involves sorting the level II-B data by variable type, blocking it in time and performing various quality control checks. At this stage a data format check and a hydrostatic check are performed. Next, data for insertion into the assimilation model are prepared by optimum interpolation (referred to as OI hereafter) of observations to 19 mandatory pressure levels (see Figure 2) and to the model grid points for winds u and v, temperature T, mixing ratio q and also surface pressure P_s. Before the vertical interpolation is

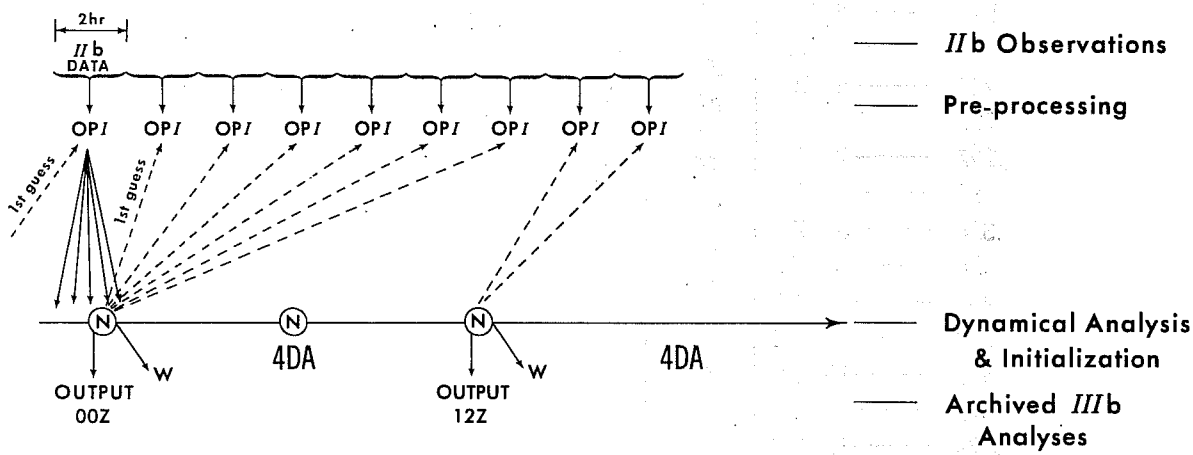

Fig. 1 -- Schematic overview of the GFDL data processing system during FGGE.

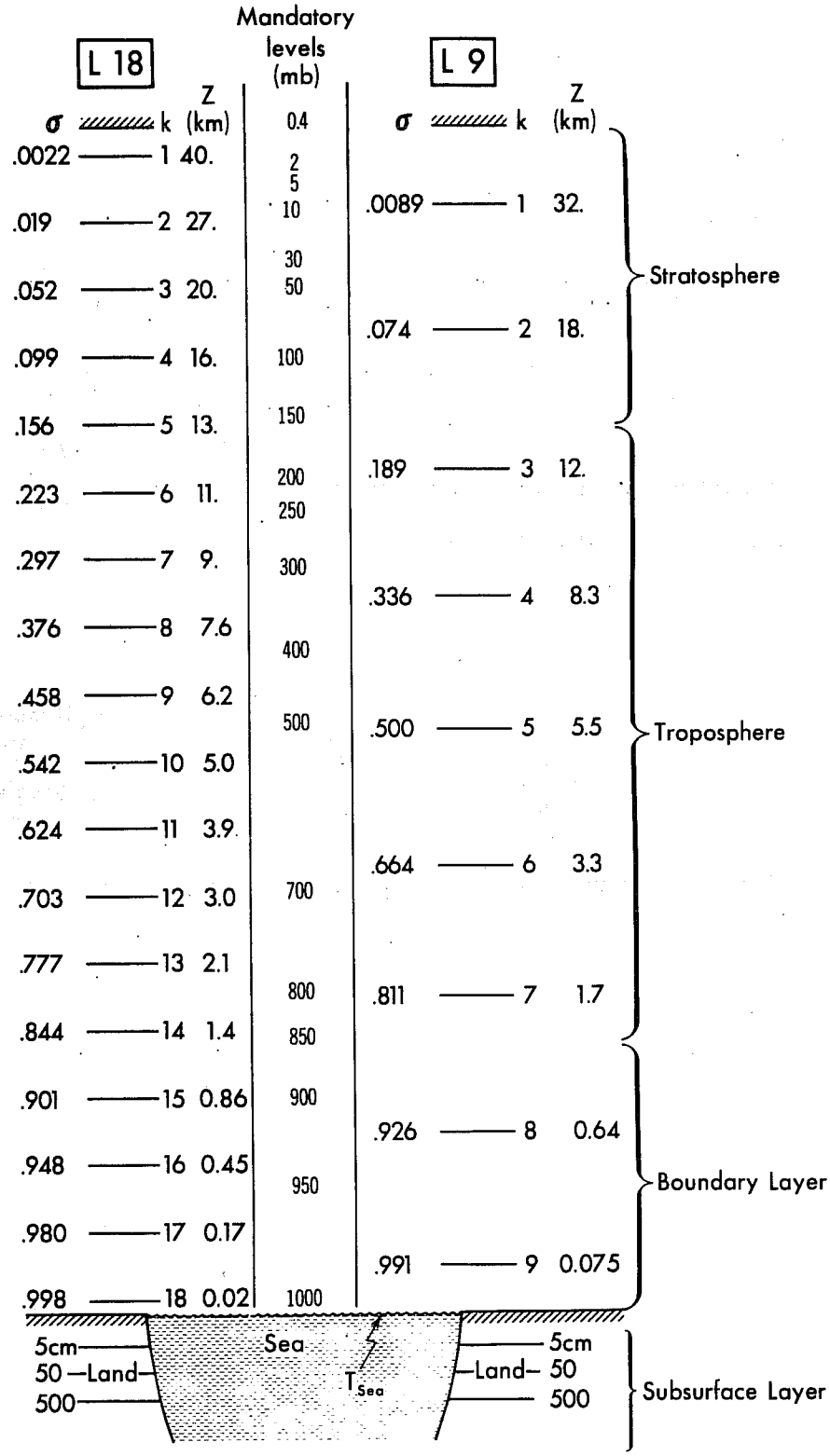

Fig. 2 -- Vertical levels for 18 level (L18), and 9 level (L9) models and 19 mandatory pressure levels. Sigma coordinate values, level indicies, and approximate level altitudes are indicated in columns headed by σ, k, and Z respectively.

performed, a third quality control check takes place by comparing observed data with the most recent synoptic time analysis; if differences exceed gross rejection criteria, then the observations will not be used. After the vertical interpolation, a comparative test with neighboring observations ("buddy" check) is made on the mandatory levels, similar to what is done by Bergman(1978).

OI is a statistical analysis which has been extensively discussed by Gandin (1963) and later by many others. The OI employed by the GFDL FGGE system is univariate for all fields with a maximum vertical range of 3 mandatory pressure levels and a maximum horizontal influence region of 250 km.

To help discuss the OI, it may be convenient to briefly describe the interpolation equations. Following Bergman's notation (1978; 1979), where two types of observations of an arbitrary variable are considered, let T^R and T^S represent radiosonde and satellite observations respectively. Then the linear equation to find the analysis value, $\tilde{T}_o$, at a spatial point o may be written as

$$\tilde{T}_o = T_o^* + \sum_{i=1}^{m} a_i \hat{t}_i^R + \sum_{j=1}^{n} b_j \hat{t}_j^S \qquad (1)$$

where $\hat{t}_i$ is the deviation of the observed data, $\hat{T}_i$, at a point i from the initial guess, T_i^*. In this example there are m radiosonde and n satellite data values within the collection range (For the GFDL OI used for processing FGGE data, the maximum number of observations is limited to 8). Weighting factors, a_i and b_j, are to be determined so that the expected analysis error variance is minimized. This requirement will lead to a set of linear algebraic equations:

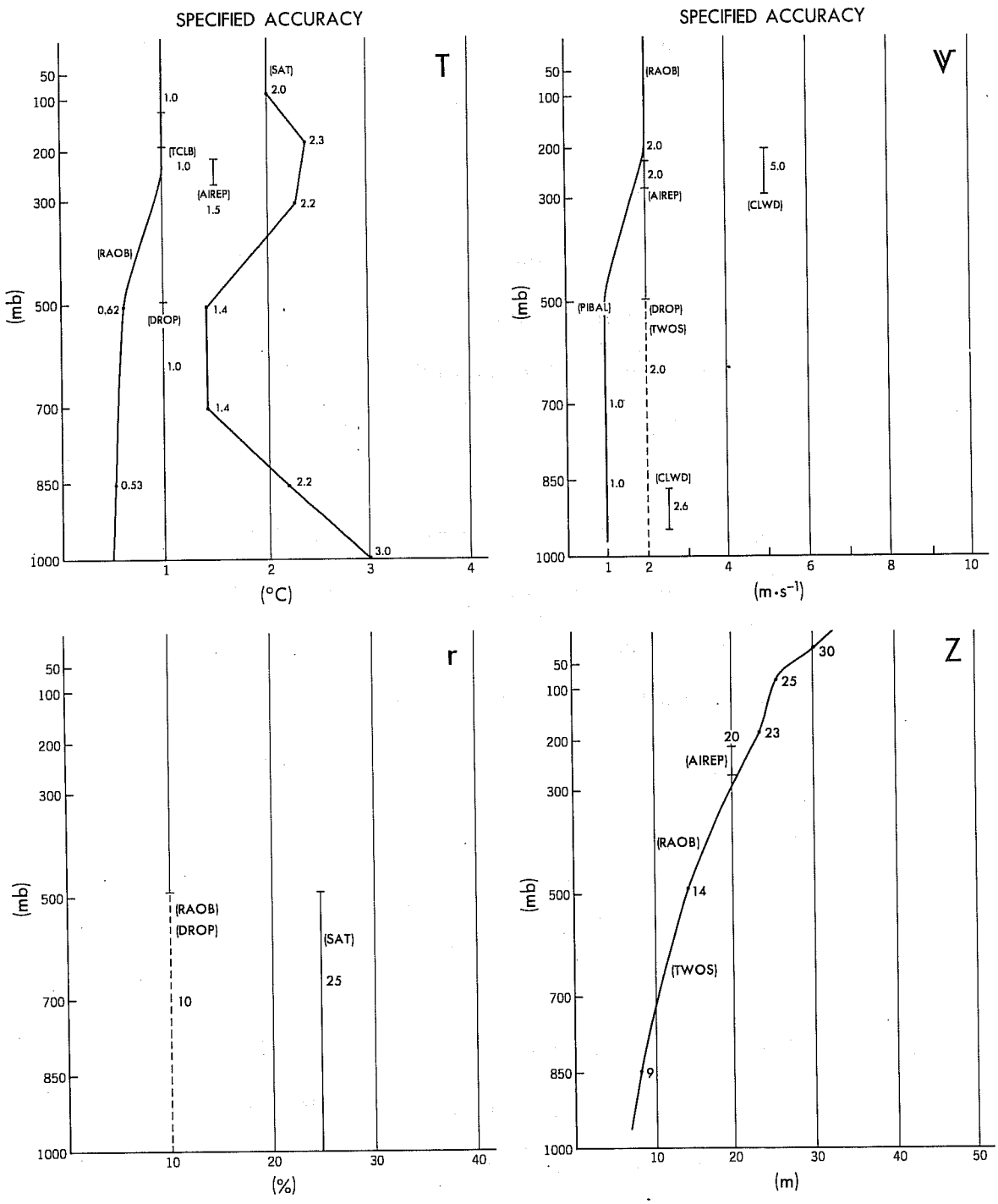

Fig. 3 -- Specified observational errors for temperature (T), wind vector (V), humidity (r), and geopotential height (Z). Types of observations are, radiosonde (RAOB), pilot baloon (PIBAL), aircraft reports (AIREP), Tropical Wind Observing Ships (TWOS), dropsonde (DROP), Tropical Constant-Level Balloons (TCLB), satellite (SAT), and cloud-tracer wind (CLWD).

$$\sum_{k=1}^{m} a_k \left[\gamma_{ik} + \rho_{ik}^{R}\ \sigma_i^{R}\right] + \sum_{j=1}^{n} b_j \gamma_{ij} = \gamma_{oi} \qquad (2)$$

$$i = 1,2, \ldots\ldots m$$

$$\sum_{\ell=1}^{n} b_\ell \left[\gamma_{j\ell} + \rho_{j\ell}^{S}\ \sigma_j^{S}\right] + \sum_{i=1}^{m} a_i \gamma_{ij} = \gamma_{oj} \qquad (3)$$

$$j = 1,2, \ldots\ldots n$$

where γ_{ij} is the correlation coefficient of background field errors between points i and j, σ_i is the normalized observational error variance and ρ_{ij} is the autocorrelation coefficient of observed errors at points i and j (generally = 0 for $i \neq j$).

By knowing γ_{ij}, σ_i and ρ_{ij}, the set of $m + n$ equations (2) and (3) can be solved for a_k and b_ℓ. Then the estimate, $\tilde{T}_o$, is obtained from equation (1), and the expected analysis error variance at a point o, ε_o, is calculated by

$$\varepsilon_o = \overline{E_o^2} \,/\, \overline{t^2} = 1 - \sum_{i=1}^{m} a_i \gamma_{oi} - \sum_{j=1}^{n} b_j \gamma_{oj} \qquad (4)$$

where $\overline{t^2}$ is the atmospheric variance. The form of the structure functions γ_{ij} is quite similar to that described by Bergman (1978). The observational errors used in the GFDL system are summarized in Figure 3.

The FGGE system involves a 12 hour cycle with data processed in two hour blocks via six separate OI analyses (see Figure 1). The first guess used for all OI is the most recent synoptic time analysis available (00Z or 12Z) after initialization is applied. Before the preparation of insertion data is complete, there are two more steps. First, weights are computed, corresponding to each analysis value and based on the expected analysis error variance; this is a reflection of the number and

reliability of the observations used to form the analysis estimate. Finally, an interpolation in time is performed in order to fill in gaps at particular analysis points. By using earlier and later values that are available, a smoother time history of insertion data is maintained.

2.2 Dynamic Assimilation

Insertion data is interpolated to the model's sigma levels (see Figure 2) using cubic splines where possible. The data is then assimilated into the forecast model in a weighted manner. At points where insertion data exists, the updated model solution for each variable, ϕ^U_{ijk}, may be written as

$$\phi^U_{ijk} = (1 - w_{ijk}) \cdot \phi^M_{ijk} + w_{ijk} \cdot \phi^I_{ijk} \tag{5}$$

where w_{ijk} is the weight value at point i, j, and level k, ϕ^I_{ijk} is the insertion value for a two hour data block, and ϕ^M_{ijk} is the model solution at the beginning of the two hour data block. At points where there are no insertion data values, the current model solution is unaltered. The weights have been adjusted prior to assimilation as follows:

$$w_{ijk} = 1.0 \quad \text{if} \quad \varepsilon_{ijk} < \varepsilon_{crit}$$

$$w_{ijk} = [1/(1 + \varepsilon_{ijk} - \varepsilon_{crit})] \exp\left(\frac{\varepsilon_{crit} - \varepsilon_{ijk}}{.5}\right) \quad \text{if} \quad \varepsilon_{ijk} > \varepsilon_{crit} \tag{6}$$

where ε_{ijk} is the expected analysis error variance for a particular variable at a point i, j, and level k, and ε_{crit} is empirically determined and will vary depending on the variable and level.

The FGGE data assimilation model is a global spectral model, rhomboidally truncated at 30 waves and with 18 levels in the vertical. A

Monin-Obukhov surface boundary layer formulation and diurnally varying solar radiation are used. More details concerning the dynamics and physical processes may be found in Gordon and Stern (1982).

The continuous insertion of data into the model is a key feature of the GFDL data assimilation system. As indicated in (5), the same data is repeatedly inserted at each time step over a 2 hour interval. In earlier studies (i.e., Charney et al., 1969), where continuous updating was simulated by partially replacing one model solution with another model solution representing "observations", the results were overly optimistic in indicating the convergence of the updated model solution. In the real data case, the approach of the updated forecast solution toward the observations is not only opposed by error growth (as in the simulation studies) but also by the inability of the model to accept some of the observed data. A sample of the detailed behavior of the model solution during a 12 hour FGGE assimilation cycle may be seen in Figure 4. The rms differences between the model solution and insertion data (F-I), for temperature (T) and surface pressure (p*), are plotted as a function of time (denoted by the solid lines). This rms difference is computed only for those points where the OI has produced insertion values. Sharp increases in this difference occur at time steps when new insertion data is introduced (indicated by arrows) while sharp drops are the result of the model accepting some of the data. Note that most of the acceptance takes place very quickly, after which time there is very little additional acceptance. A general upward trend seen in the rms differences over each 12 hour cycle reflects an evolution of the model solution away from the first guess for the insertion data (the previous synoptic time analysis) and also possibly indicates poorer quality observations after the synoptic time. Also indicated in Figure 4 (by

letters 'N') are those time steps at which nonlinear normal mode initialization is performed. The control of spurious fast mode amplitudes by the initialization results in a temporary additional reduction in the rms difference. The dashed curves show the rms differences at the same insertion data points but during a 12 hour forecast without data insertion. When compared to the solid curve, this dashed curve shows that, at least for temperatures, the rate of error growth is being slowed by the continuous assimilation. (Although it is not shown here, wind differences behave quite like temperatures.) However, in the case of surface pressure, the rms differences during assimilation and the forecast errors are comparable, thereby indicating the model's inability to accept most of the observed surface pressure data.

It is felt that this insertion technique produces analyses that are reasonably faithful to the observations and also consistent with the model's dynamics and physical processes. However, there is evidence of significant excitation of fast modes, and there is an indication of a need to reduce data rejection, especially in the surface pressure fields.

2.3 Initialization

A nonlinear normal mode initialization using Machenhauer's method (1977) is performed every 6 hours, as indicated by "N" in Figure 1. Most of the details of the scheme used in the FGGE data processing system are described by Ballish (1980). The intent of the initialization in the GFDL system is to help control the growth of spurious gravity modes, but not to alter the model balance, especially in the tropics, where adiabatic initialization has been known to weaken circulation features. For this reason, in addition to initializing only the first 7 of 18

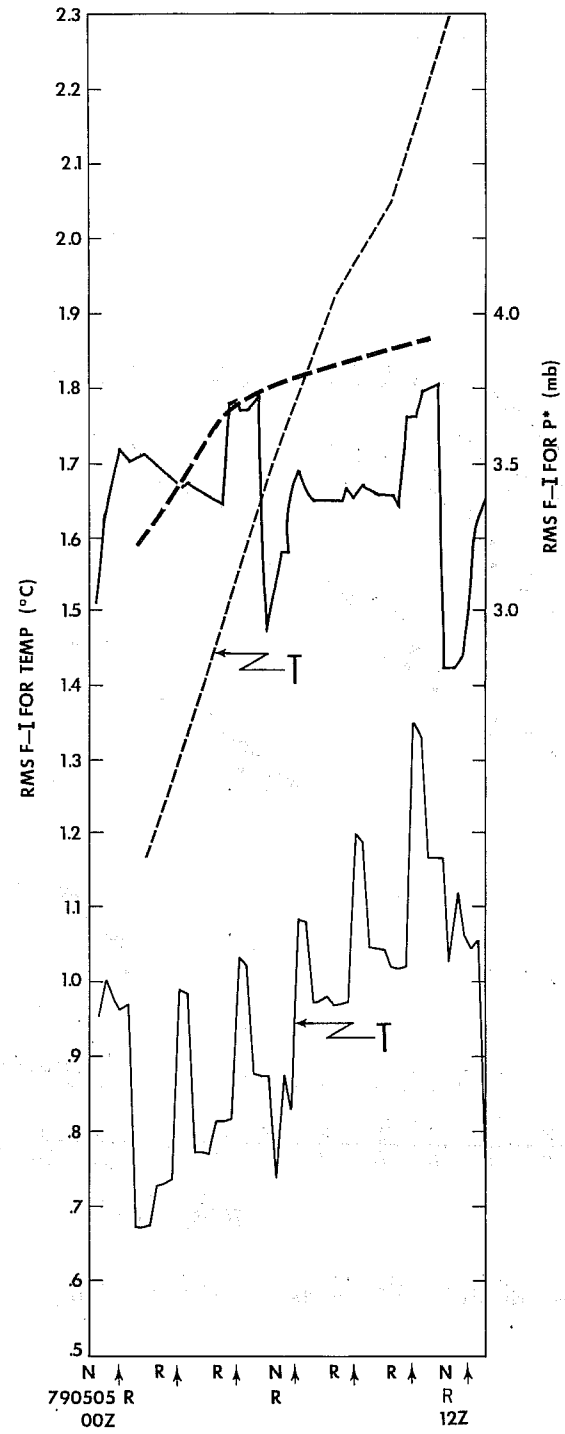

Fig. 4 -- Global rms differences between the model solution and insertion values (F-I) computed for all points that have insertion data - solid curves. Forecast error - dashed curves. Arrows indicate new data insertion times; N and R denote nonlinear normal mode initialization and radiation time steps, respectively. Temperature curves are labeled (T) and surface pressure curves are unlabled.

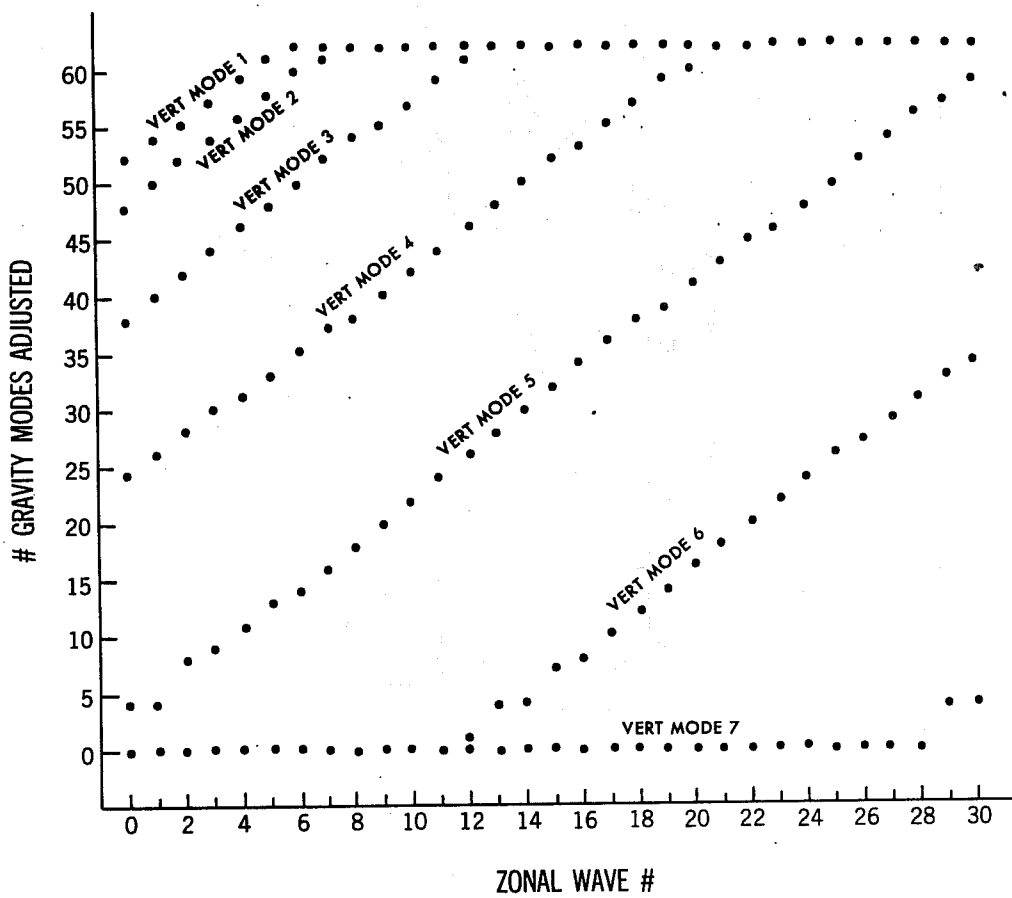

Fig. 5 -- Number of modes initialized for the FGGE data assimilation model.

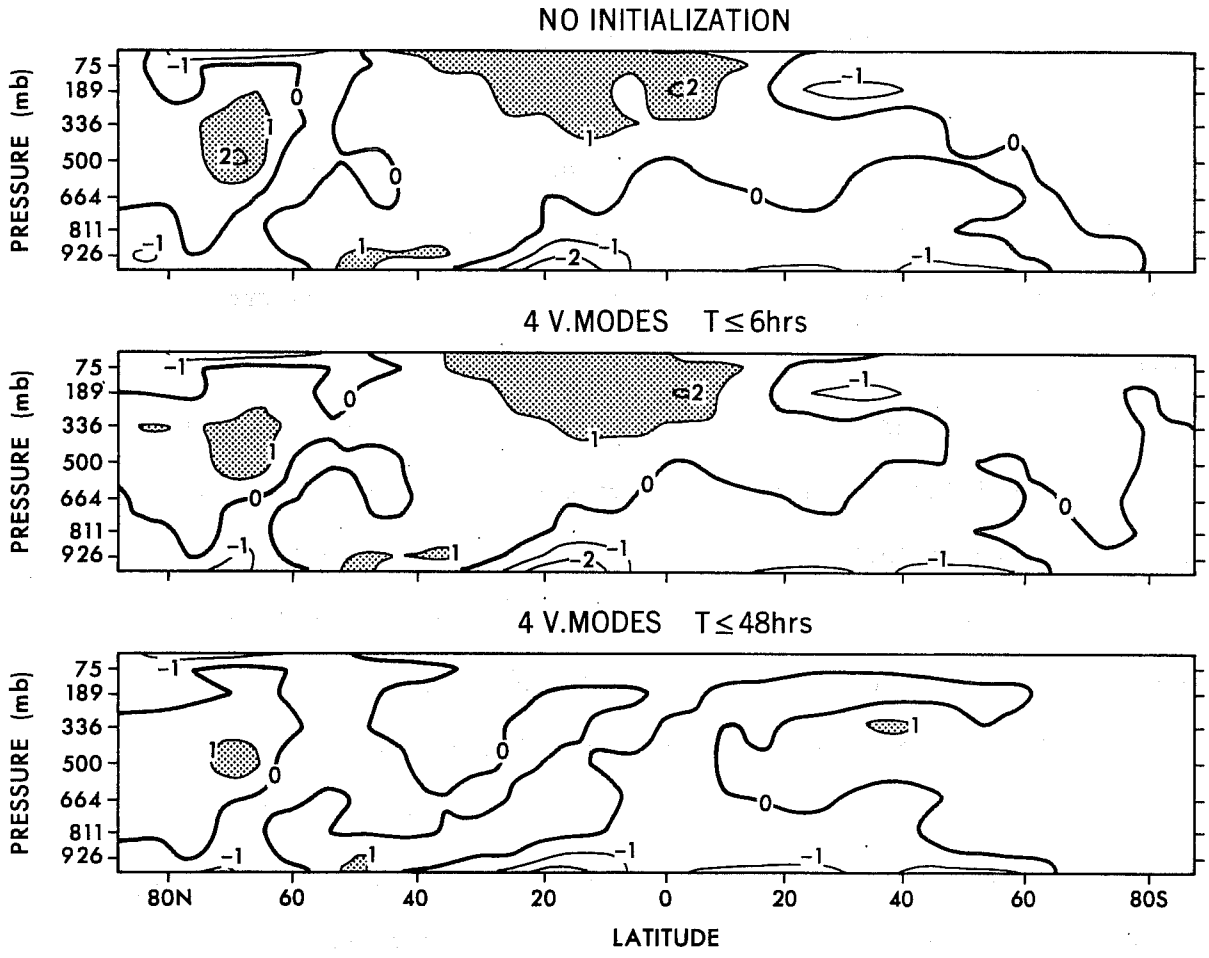

Fig. 6 -- Zonal mean meridional wind for 4-D analyses from December 7, 1978 00Z with no initialization (top), modes with periods ≤6 hours from first 4 vertical modes initialized (middle), and modes with periods ≤48 hours from first 4 vertical modes initialized (bottom). Contour interval is 1 m sec^{-1}, >1 m sec^{-1} is shaded.

vertical modes, a frequency cutoff has been imposed which limits the initialization to those modes with periods of 6 hours or shorter. Figure 5 shows the number of modes initialized as a function of vertical mode and zonal wave number. The choice of a specific cutoff frequency was motivated by observing that most of the amplitude in tropical surface pressure traces from uninitialized forecasts is in periods greater than 6 hours. It turns out that the Hadley circulation projects mostly onto modes that are not adjusted by this initialization scheme and, therefore, will be essentially unaffected. Figure 6 compares the effects of this frequency filtered initialization with that of a more general initialization (modes with periods up to 48 hours being initialized) using a 9 level version of the FGGE assimilation model. Latitude - height profiles of the zonal mean meridional wind are shown for a 4-D analysis before any initialization and after the two initialization schemes mentioned. It is clear that the 6-hour cutoff scheme maintains the strength of the return Hadley flow to be quite close to the uninitialized case, while this circulation has been suppressed by the 48-hour cutoff initialization.

3. ASSESSING III-B ANALYSES

Determining the usefulness and quality of the FGGE III-B analyses is quite an involved process that will probably continue for many years. Studies involving a range of appliations are needed to thoroughly assess the strengths and weaknesses of the analyses. GFDL is involved in at least three areas of study regarding the III-B analyses: Using the analyses to compute various general circulation features, monitoring of daily FGGE analyses, and using several FGGE analyses as initial conditions in a long range forecast study.

General circulation statistics computed from GFDL analyses are examined in a study by Rosen et al. (1984), which compares circulation features derived from the III-B analyses with those computed from station-based analyses during the FGGE special observing periods. One of the major aims of their study is to see whether the FGGE analyses can add to our understanding of the atmospheric general circulation which was based on earlier less extensive observational networks. A comparison of January 79 monthly mean 200 mb zonal wind, u, may be seen in Figures 7 and 8, taken from Rosen et al. (1984). In Figure 7, the top two diagrams are analyses based on the upper-air sounding stations only[2] (the tropical winds observing ships are included) and the lower two are the GFDL and ECMWF III-B analyses which used the complete FGGE observational network. The main point here is that the III-B analyses show some consistent differences relative to the station analyses, specifically in extending the mid-pacific jet and indicating a break in the jet across the Atlantic. Since these are in regions generally lacking conventional station data, it seems reasonable to conclude that the III-B analyses represent an improvement. However, it should be noted that differences between the GFDL and ECMWF analyses are not insignificant. This may be seen in Figure 8 which shows the two level III-B analyses and their difference field in more detail for a section of a latitude belt from 20°N to 40°N and ranging from 300°W to 150°W longitude. Also included in this diagram are the monthly means of the observations used in the

[2]In areas that have fairly sparse station coverage, the dependence on the initial guess is significant; hence, station analyses using two types of first guess fields are provided. The analysis referred to as "station" uses a climatological first guess, while the other uses the zonal mean at each latitude as the guess for that latitude; see Rosen et al. (1984) for more details.

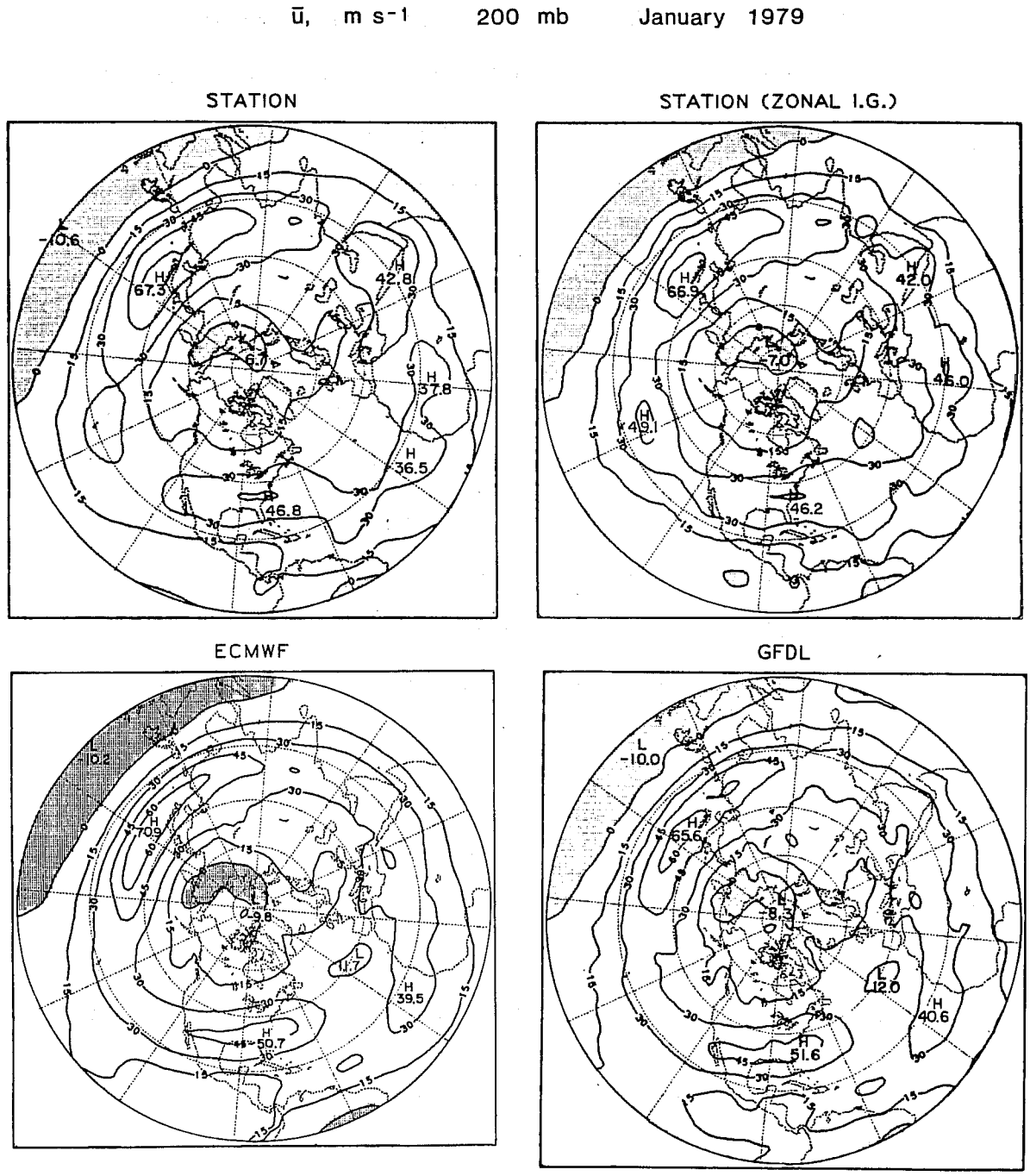

Fig. 7 -- Northern hemispheric maps of the mean zonal velocity at 200 mb for January 1979. The top two charts are derived from the two station analyses, the GFDL III-B analysis is in the lower right and the ECMWF III-B analysis is in the lower left. The contour interval is 15 m s^{-1}. Negative values, indicating easterlies, are shaded.

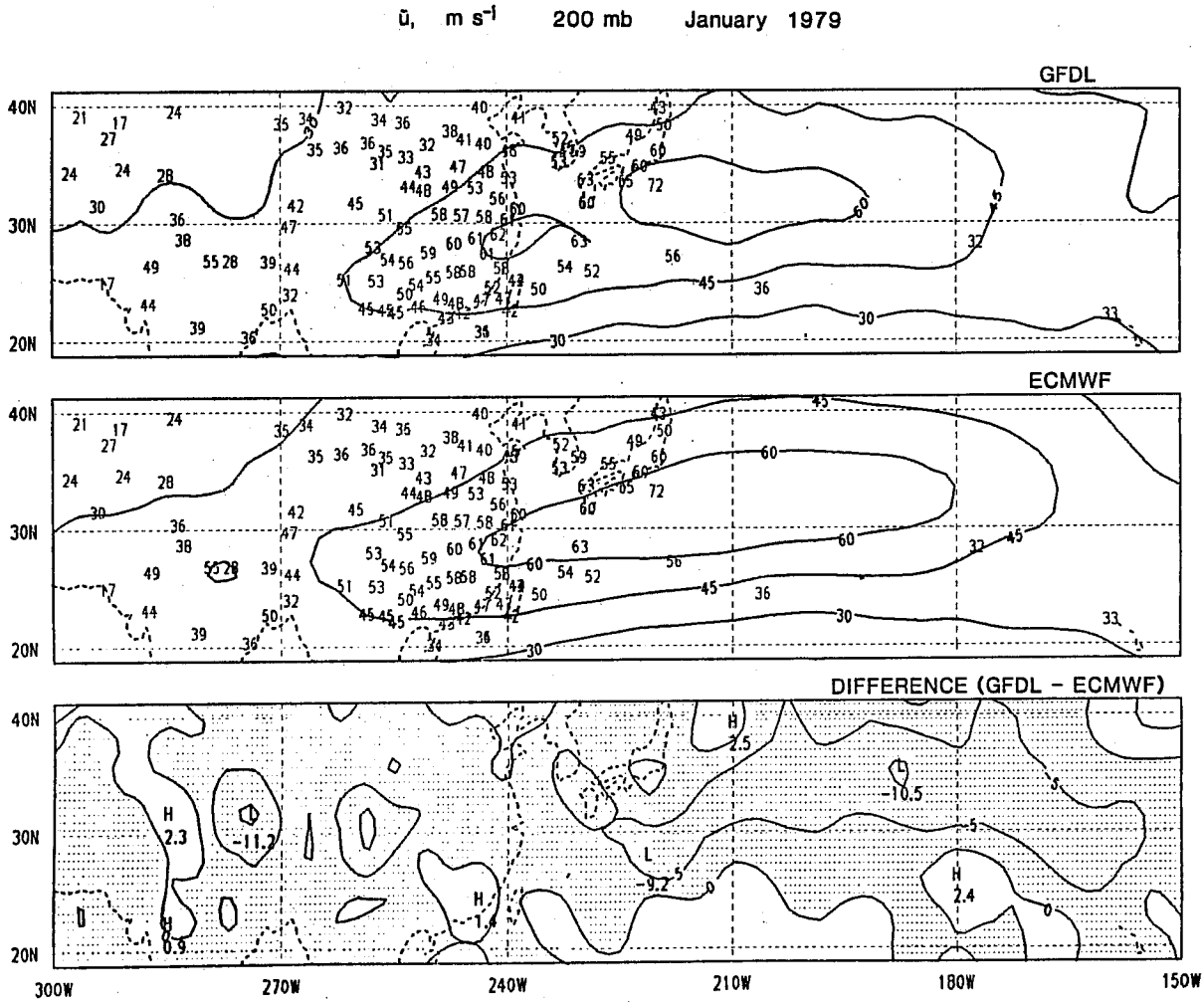

Fig. 8 — Zonal velocity at 200 mb for January 1979 mapped in the region from 20°N to 40°N and 150°W to 300°W computed from the GFDL (top) and ECMWF (middle) level III-B analyses. The contour interval is 15 m s^{-1}; monthly mean values of u at the rawindsonde stations that passed the cut-off criterion for the month are also displayed. In the bottom panel is the difference, GFDL minus ECMWF, with a contour interval of 5 m s^{-1} and negative values are shaded.

station analyses plotted at the station locations. It is apparent that in at least several places, the III-B analyses differ significantly in their fit to the observations, on the order of 10 m/sec, and it is not obvious that one analysis is providing a consistenly closer fit than the other. Another feature that should be noted here is the greater amount of small scale activity exhibited in the GFDL analysis (as well as the "zonal" station analysis) which is probably the result of a more locally structured influence region in GFDL's OI when compared to the OI scheme used by the ECMWF (Bengsston et al., 1982). In the tropics even more significant differences are apparent between the two III-B analyses, as GFDL's shows stronger divergent flow. This may be seen in Figures 9a and 9b, also taken from Rosen et al. (1984), which compare mass streamfunction for January and June. Overall, the GFDL analysis seems to be closer to the station analyses, especially when comparing to the zonal initial guess and more so for the January circulation. Although the GFDL cases are consistently stronger, the magnitude and position of the Hadley cell appears better than in the ECMWF cases. (In contrast to the 200 mb mean zonal wind comparison, where differences were probably due to different data handling techniques, it is felt here that the much more selective initialization in the GFDL assimilation system is a major factor in allowing for the retention of the Hadley circulation - see Figure 6). Other circulation statistics, including several variance quantities, have been discussed in some detail by Rosen et al. (1984), but are not included here because they are beyond the scope of this report. In summary, it appears that the III-B analyses depict realistic general circulation patterns and in some cases seem to capture features that conventional analyses have missed. However, as Rosen et al., (1984)

indicate, the significant differences that exist between the two III-B analyses present a measure of the limitations that one should be aware of when using these analyses. Therefore, it seems that further improvements in the assimilation schemes are needed before full confidence can be placed in the circulation patterns they depict.

At GFDL, numerous III-B extratropical height and wind analyses have been examined. Generally, they compare favorably with those produced by the ECMWF in the phase and amplitude of medium to long wave features. However, there is some tendency for the ECMWF troughs to be deeper, especially in data sparse regions. Figures 10a and 10b show this trend by plotting trough minima from GFDL analyses versus those in the corresponding ECMWF analyses for about 70 cases from December 78 - February 79 and from September 79 - October 79. Figure 10a, which compares 500 mb troughs, shows more points above the 45 degree equality line indicating values generally higher for GFDL. The surrounding dashed lines show an observational error range of 15 m. Figure 10b is identical to 10a in format, except that mean sea-level pressure (MSLP) systems are being compared and the error range is .5 mb. In this, the tendency for deeper lows in the ECMWF analyses is fairly clear. The sample size does not appear to be sufficient to show any preferences for hemisphere or time period; however, for individual cases the greatest differences appear to be in data-sparse regions, as one would expect, and also in cases of intense cyclogenesis. It seems that a combination of factors are probably responsible:

1. The multivariate OI used by the ECMWF system will allow for greater acceptance of surface pressure data than with the univariate GFDL scheme, and the larger influence region in the

Mass Streamfunction, 10^{10} kg s^{-1} January 1979

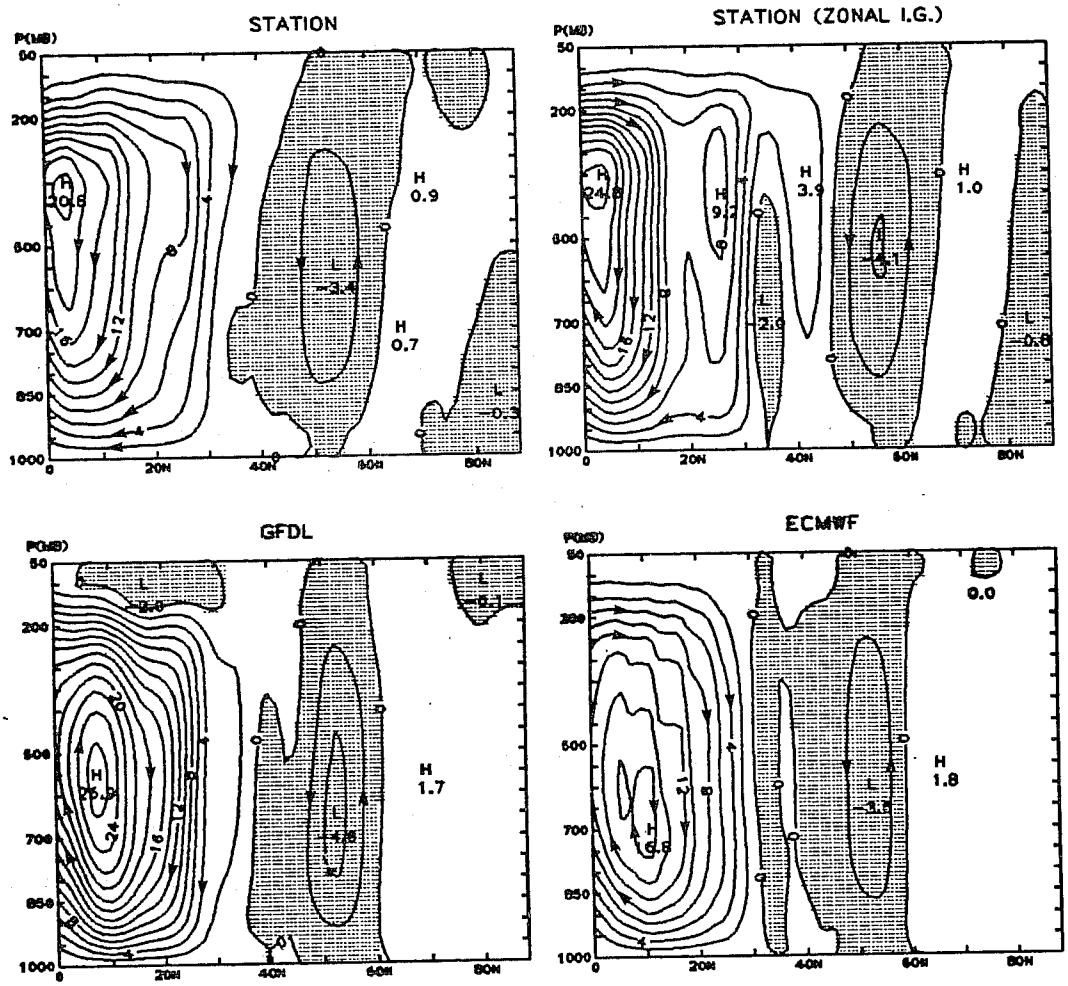

Fig. 9a -- Cross sections of the mass streamfunction for the mean meridional circulation for January 1979. It has been computed for both station analyses (top panels) and for the GFDL (lower left) and ECMWF (lower right) III-B analyses. The contour interval is 2 x 10^{-10} Kg s^{-1} and negative values are shaded.

Mass Streamfunction, 10^{10} kg s^{-1} June 1979

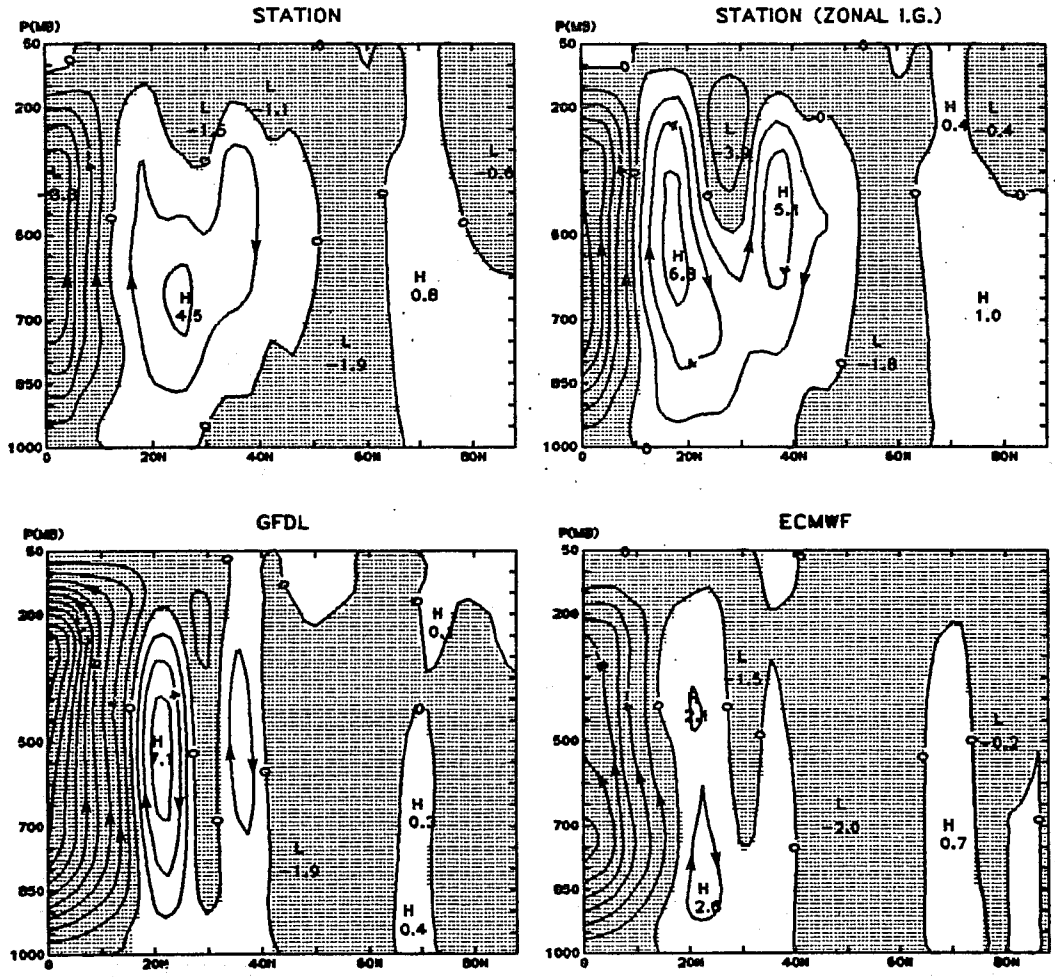

Fig. 9b -- Same as 9a, except for June 1979.

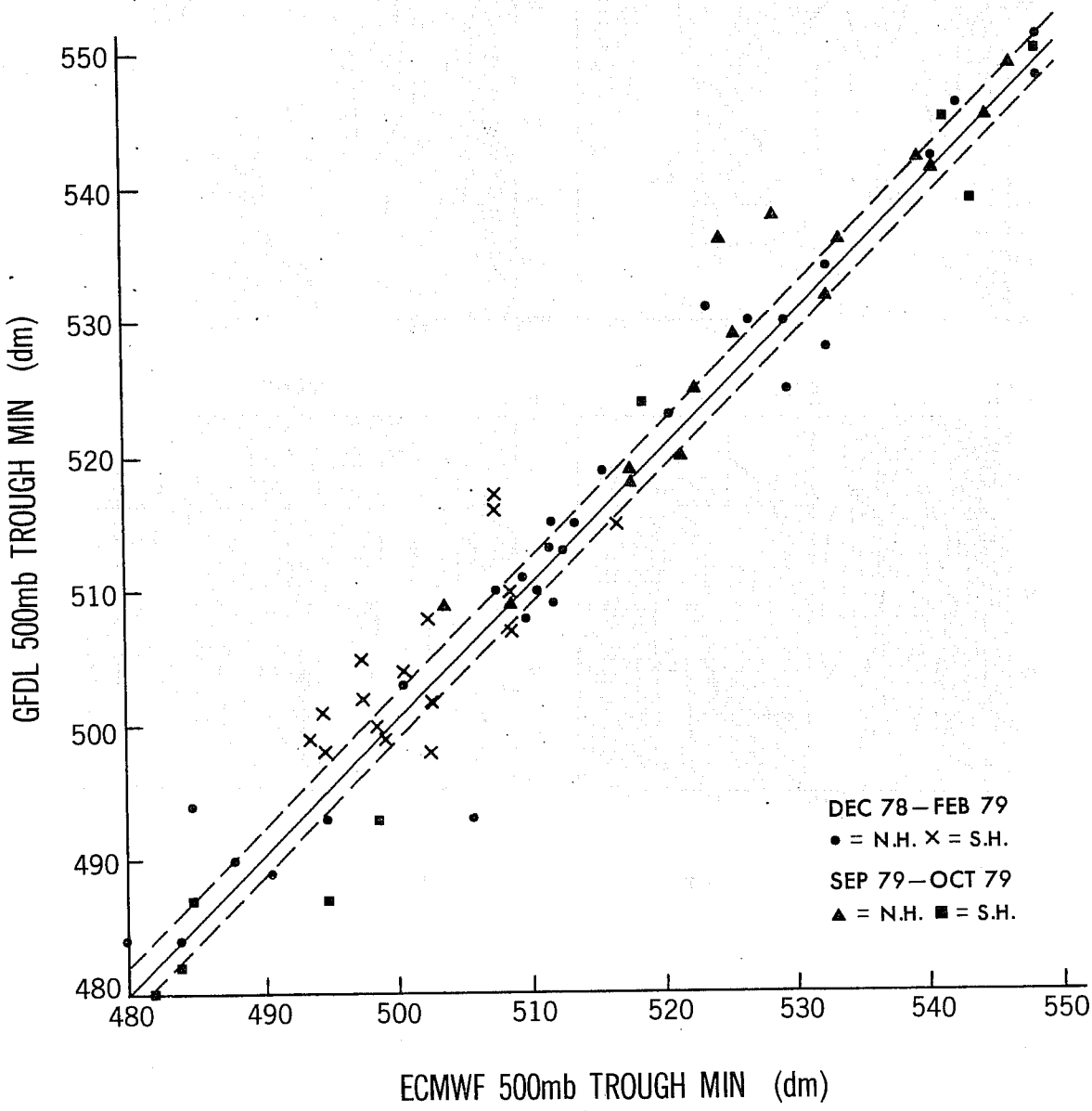

Fig. 10a -- Individual 500 mb trough minima from GFDL III-B analyses versus corresponding minima from ECMWF analyses. Those taken from the periods December 1978 through February 1979 are noted with a "•" for northern hemisphere and a "x" for southern hemisphere; for the period September and October of 1979, a northern hemisphere trough is indicated by a "Δ" and a southern hemisphere trough is shown as a " ■ ". The 45° solid line indicates equal depths for both analyses and the surrounding parallel dashed lines show an approximate observational error range.

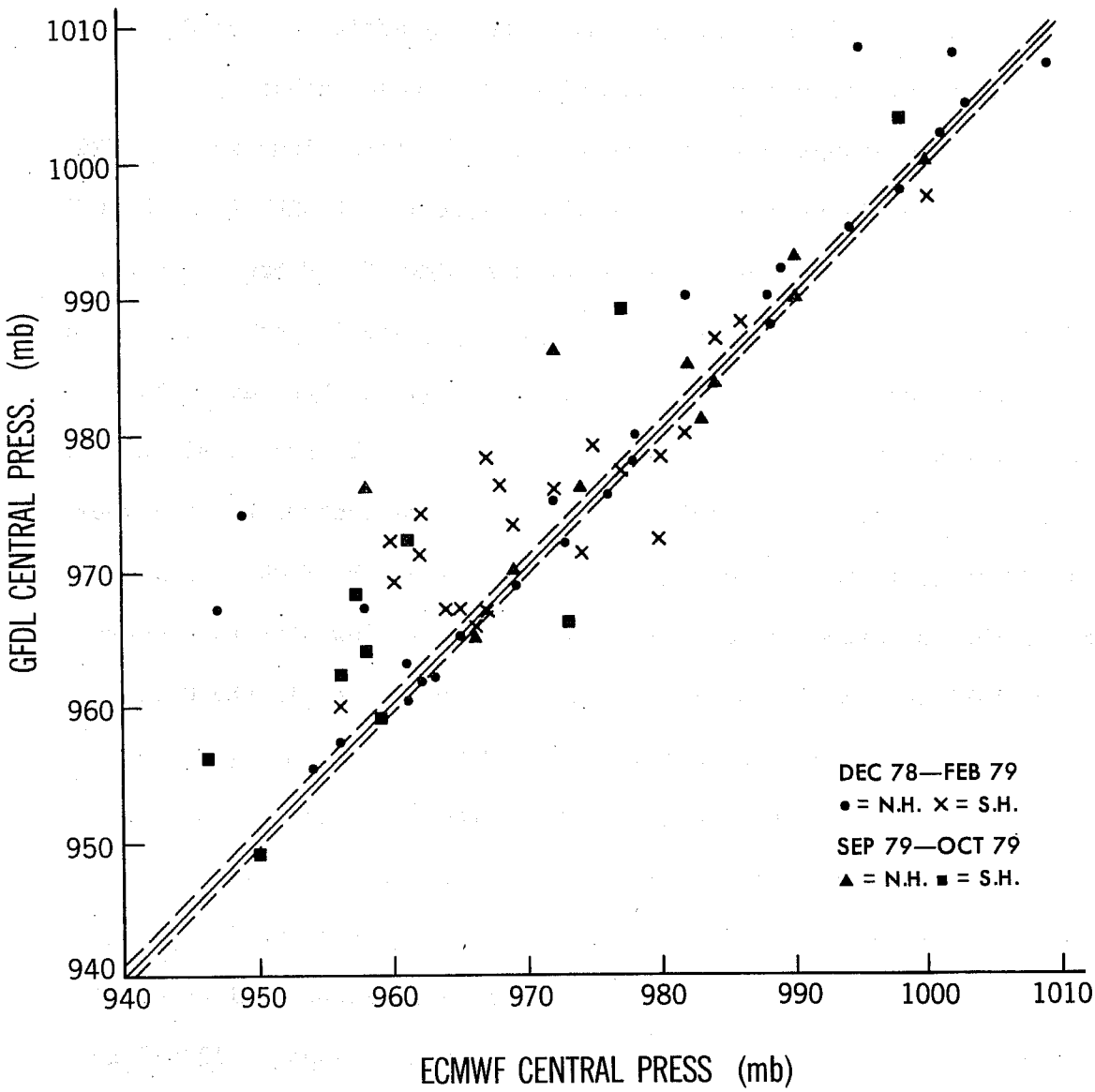

Fig. 10b — Same as in Figure 10a, except sea-level low pressure systems are compared.

ECMWF OI should be beneficial in data-sparse areas.

2. In cases of rapidly moving systems or intense cyclogenesis, the 12 hour old first guess could have an effect in reducing the magnitude of surface pressure changes.
3. The ECMWF analysis is higher resolution - 1.875° x 1.875° vs. 3.75° x 2.2° for the GFDL analysis; in addition, the GFDL analysis is truncated rhomboidally at wave number 30.

An extreme example of the tendency for shallower lows in some GFDL analyses is shown in Figures 11a and 11b. Analyses of MSLP from the GFDL system are compared to those produced at the ECMWF for February 19 00Z and February 20 00Z, 1979. This was a case of dramatic cyclogensesis off the U.S. east coast (the Presidents' Day snowstorm). Excluding those regions of high topography, the analyses in Figure 11a are fairly close, with the largest difference being the storm in the central Pacific where the ECMWF analysis is 6 mb deeper. In Figure 11b, the ECMWF analysis captures the explosive deepening of the east coast low with the central pressure falling from 1020 mb to 995 mb in 24 hours, while the GFDL analysis indicates a drop from 1021 mb to 1008 mb. An investigation of the lack of depth of this feature helps to exemplify some of the system's shortcomings that should be addressed. The storm center on the stereographic projection in Figure 11b on 20 February 1979 00GMT has a sea-level pressure value of 1008.3 mb. This value is determined mainly from two points on the Gaussian analysis grid with values of 1005.7 and 1009.5, respectively. The 1009.5 is the model solution at that point, as no observations were located within its radius of influence (250 km), while only a single observation of 998.0 mb falls within the other point's region. After incorporating a first guess value of approximately

1030 mb based on the 19 Feb 79 12GMT analysis, the OI analysis sea-level pressure value is 1002.5 mb. Hence, an insufficient amount of observational information, a single observed value 196 km from the analysis grid point, is putting too much reliance on the 12 hour old first guess, which is hardly representative of this moving, rapidly deepening cyclone a half day later. It turns out that the addition of more up to date information, i.e., using a more current first guess (even 6 hours old) or extension of the OI range to include just one additional observation would deepen the operational OI analysis value by over 4 mb. (It should be noted that a higher resolution grid might also produce a deeper MSLP analysis, with more weight probably being given to the single observation as it is likely to be closer to an analysis gridpoint.) The remaining 3.2 mb is apparently lost due to at least a partial rejection of the incremental surface pressure change by the model solution, which in this case gets over a 30% weight with the insertion data.

Two dates from the first special observing period (SOP1) of FGGE were chosen for initial conditions as part of a monthly forecast study using the GFDL gridpoint model (Miyakoda and Sirutis, 1977). They are 01 January 1979 00GMT and 16 January 1979 00GMT. For each of these cases, forecasts were initiated with III-B analyses from GFDL and ECMWF and the corresponding III-A analysis from NMC. Figures 12a and 12b show 500 mb height comparisons of the initial conditions, 2 day forecasts and 10 to 30 day time average forecasts from the January 01 and January 16 cases respectively. The GFDL initial conditions exhibit more noise, which most likely is the result of the relatively high level of fast modes generated during assimilation and to some extent the more local influence region of the GFDL OI. In addition, for both the initial time and the day 2

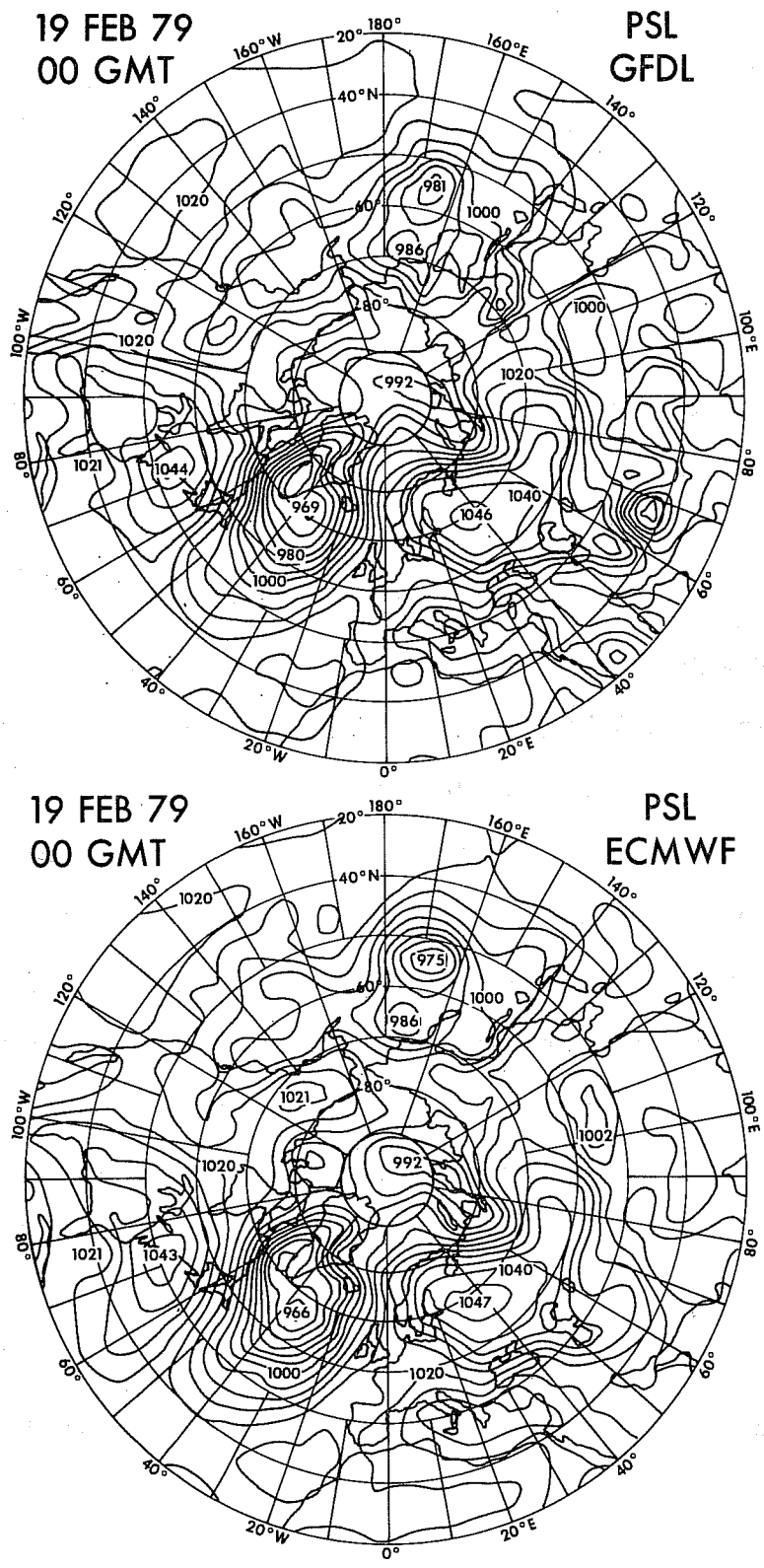

Fig. 11a - Northern hemispheric maps to 20°N of sea-level pressure for 00Z February 19, 1979 from the GFDL analysis (top) and the ECMWF analysis (bottom). Contour interval is 5 mb.

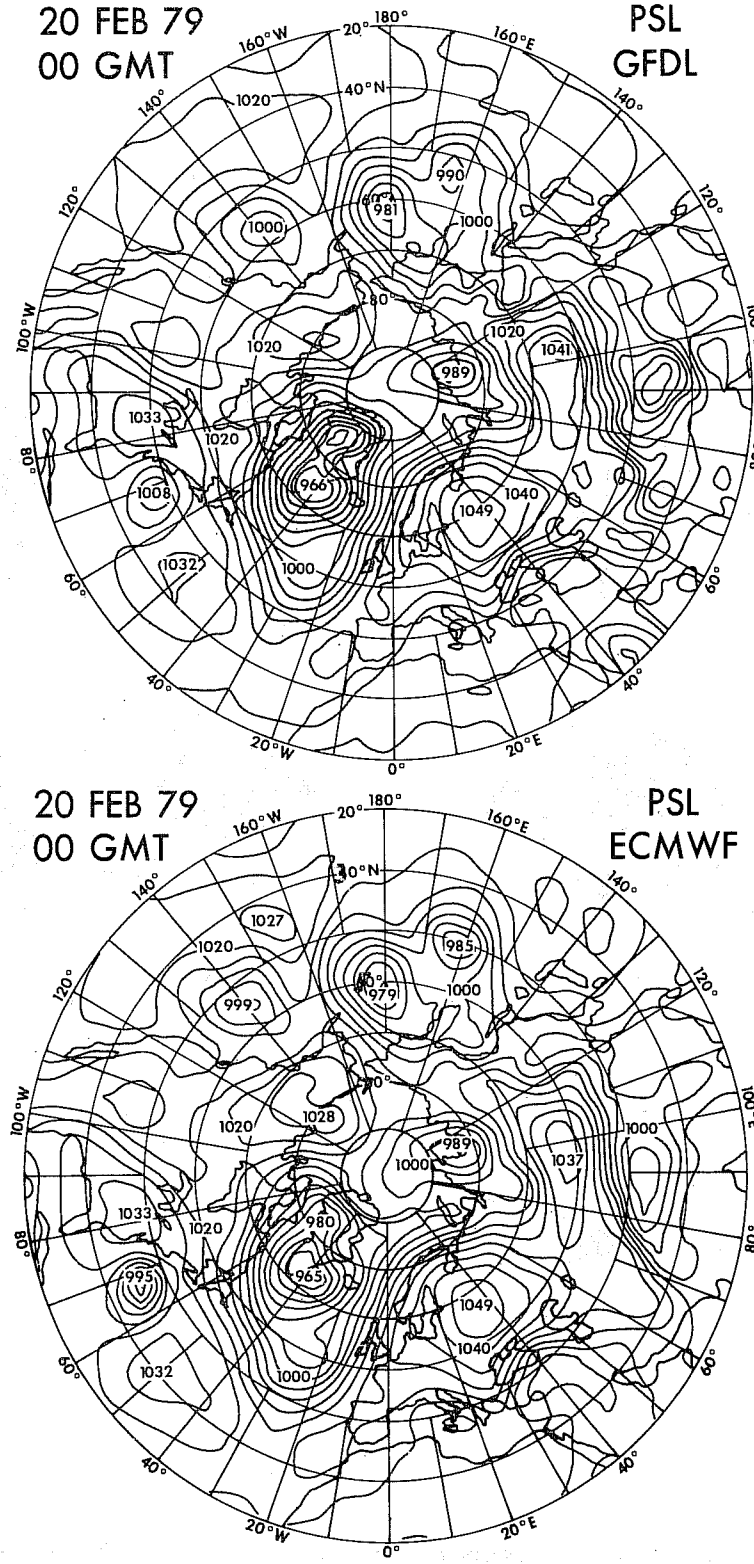

Fig. 11b -- Same as in Figure 11a, except for 00Z February 20, 1979.

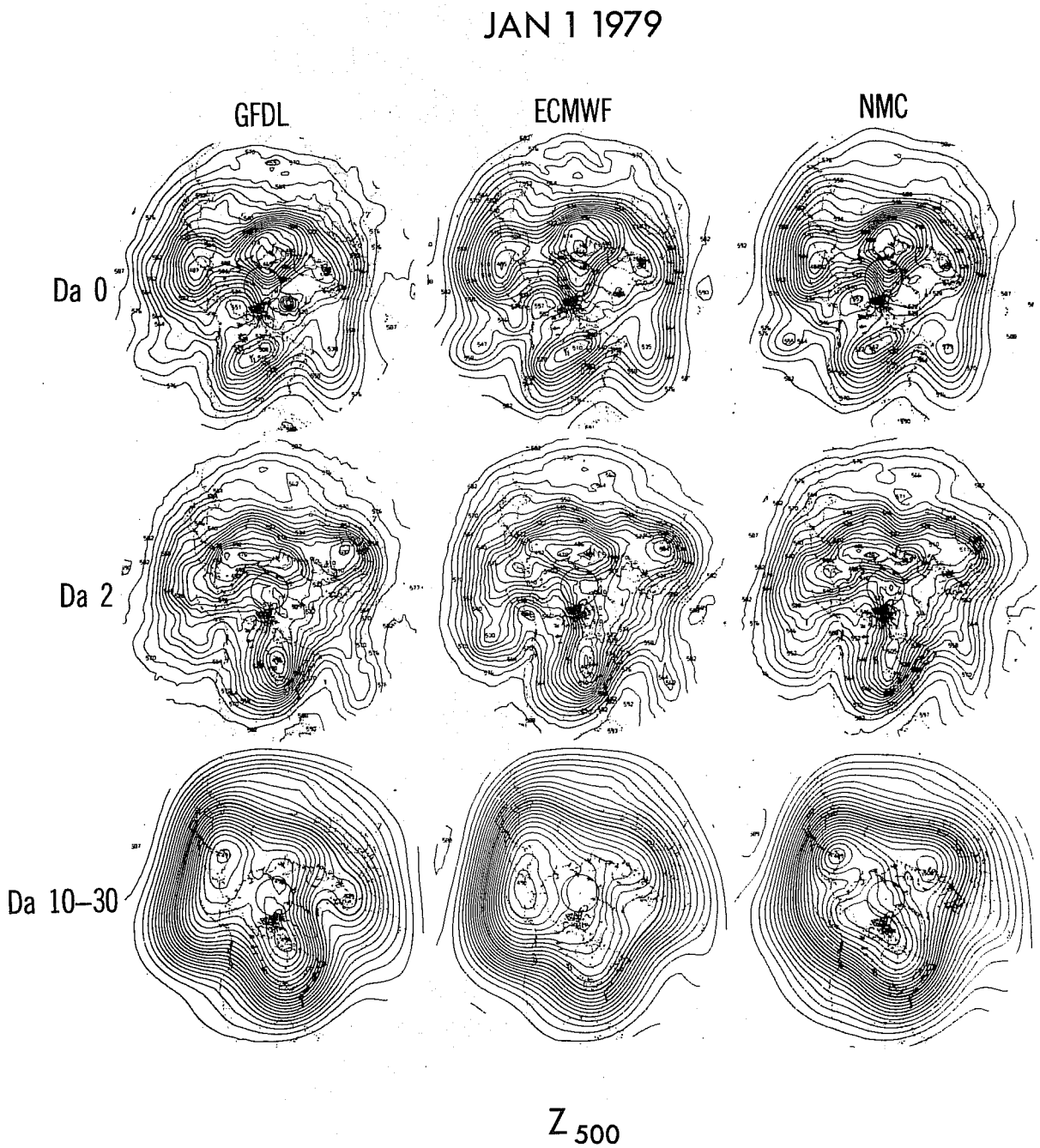

Fig. 12a -- Northern hemispheric maps to 20°N of 500 mb height from forecast case study using 00Z January 01, 1979 for initial conditions. Initial analyses are compared for GFDL III-B (top left), ECMWF III-B (top center), and NMC III-A (top right). Day 2 forecast analyses are shown in the middle row and day 10-30 time average forecasts are displayed in the bottom row. Contour interval is 60 m.

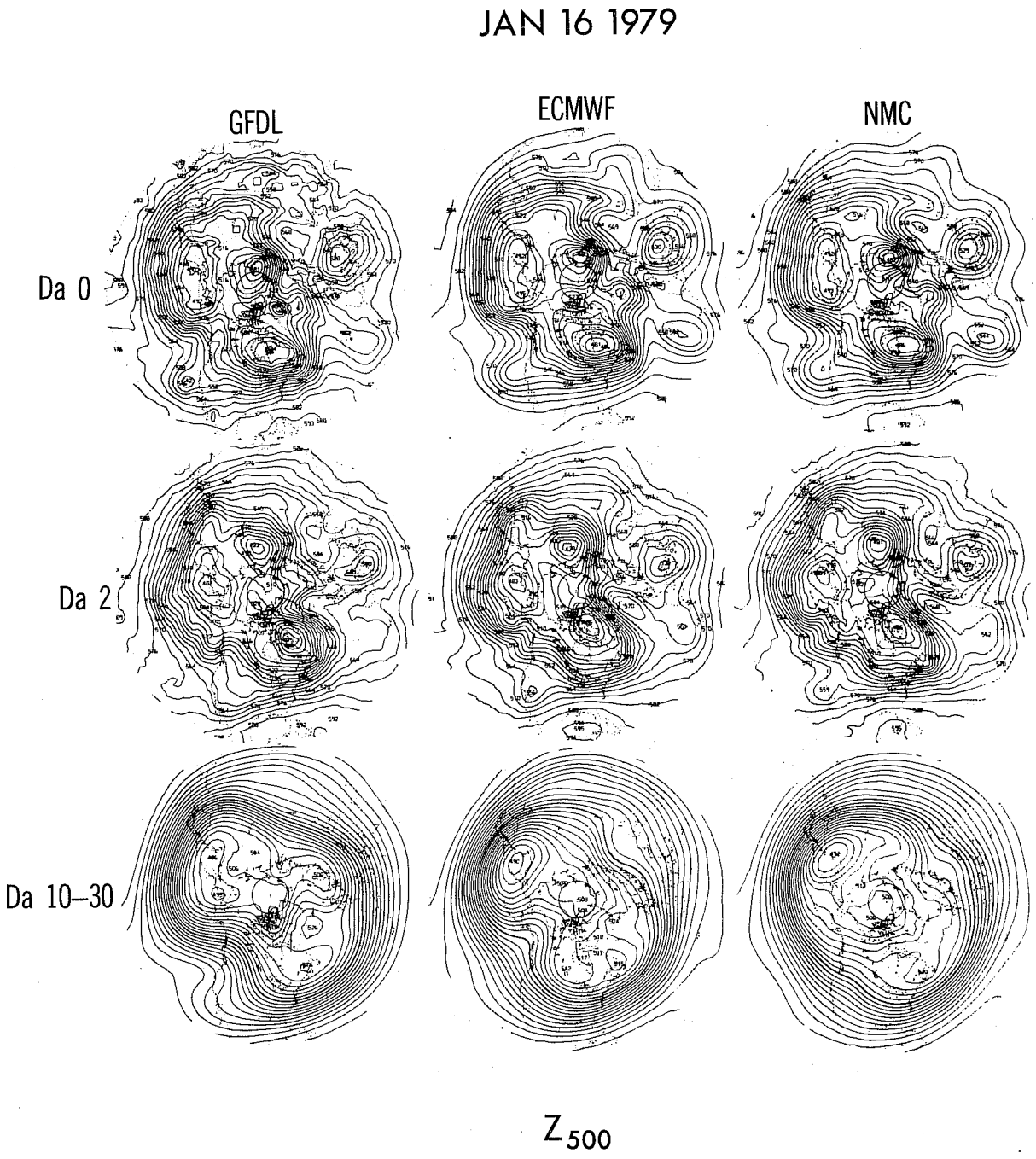

Fig. 12b -- Same as in Figure 12a, except for forecast starting from 00Z January 16, 1979.

forecast, the amplitude of the transient features in the GFDL analyses appears to be somewhat weaker than in the ECMWF and NMC analyses. However, for the extended range forecasts shown, it is fairly clear that the forecasts from the GFDL analyses have greater amplitude in the quasi-stationary long waves. This same tendency has been seen in other northern hemisphere winter cases where GFDL and NMC analyses have been used. Although the number of cases involved is far too few to draw any statistically significant conclusions, there is enough evidence to justify further study to better understand the reasons for these differences. At this time, our speculation is that the tropical analysis is quite important for the extended range mid-latitude winter forecasts and that the greater amount of tropical divergent flow retained in the GFDL analyses may be responsible for the stronger long waves in the extra-tropics.

4. CONCLUSIONS

Continuous data assimilation at GFDL during FGGE appears to be a capable system which has produced useful III-B analyses, though there are some shortcomings. Its main assets are: (1) the potential to achieve a consistency among variables which utilizes the full physics of the model (including diabatic processes), (2) the control of error growth in the model solution, which may help to improve analyses in data voids, and (3) an ability to assimilate asynoptic data closer to their actual observation time.

General circulation features were shown to be quite realistic when compared to more traditional analyses, and in some instances the III-B analyses appeared to have improved upon the station analyses. However, the differences between the GFDL and the ECMWF analyses do seem to

warrant another look at the data selection and screening criteria.

Although the lack of constraints imposed during OI and initialization appear to be an asset in allowing for the strong tropical divergence, the GFDL analyses show a fairly high level of fast modes and a tendency to reject unbalanced mass data, especially surface pressure. Linear normal mode initialization of the data increments (Puri et al., 1982) and using a geostrophic correction technique to produce winds where only mass data exists (Bourke et al., 1982) will be investigated in a later study as possible remedies to these deficiencies. Other proposed areas of study include: (1) increasing the influence range in the OI to be more consistent with the assimilating model's resolution, which could also reduce the amount of small scale structure seen in the analyses, (2) increasing the model resolution as well as the analysis grid resolution, and (3) reducing in the cycle time from 12 hours to 6 or perhaps 2 hours so as to use a more current first guess during OI.

5. ACKNOWLEDGEMENTS

The authors are grateful to Drs. R. Rosen and N. Lau for their helpful discussions and diagrams. Thanks also to the GFDL drafting group and photographer John Conners for their part in preparing the figures, and to Joan Pege for word processing this paper.

6. REFERENCES

Ballish, B., 1980: Initialization, Theory and Application to the NMC Spectral Model. Ph.D. Thesis, Dept. of Meteorology, University of Maryland, 151 pp.

Bengtsson, L., M. Kanamitsu, P. Kallberg, and S. Uppala, 1982: FGGE 4-dimensional data assimilation at ECMWF. Bull. Amer. Meteorol. Soc., 63, 29-43.

Bergman, K. H., 1978: Role of observational errors in optimum interpolation analysis. Bull. Amer. Meteor. Soc., 59, 1603-1611.

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

Bourke, W., K. Puri, R. Seaman, B. McAvaney and J. LeMarshall, 1982: ANMRC data assimilation for the southern hemisphere. Mon. Wea. Rev., 110, 1749-71.

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.

Gandin, L., 1963: Objective Analysis of Meteorological Fields. Gidrometeorologichesckoe Izdatel Stro, Leningrad, translated from Russian to English, Israel Program for Scientific Translation, Jerusalem, 1968, 242 pp.

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

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model with application to normal mode initialization. Beitrage Zur Physik der Atmosphare, 50, 253-271.

Miyakoda, K and J. Sirutis, 1977: Comparative integrations of global models with various parameterized processes of subgrid scale vertical transports: Description of the parameterizations. Contrib. Atmos. Phys., 50, 445-487.

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.

Puri, K., W. Bourke, and R. Seaman, 1982: Incremental linear normal mode initialisation in four dimensional assimilation. Mon. Wea. Rev., 110, 1773-85.

Rosen, R., D. Salstein, J. Peixoto, A. Oort and N. Lau, 1984: Circulation statistics derived from Level III-b and station-based analyses during FGGE. (Submitted for publication in Mon. Wea. Rev.)