Short-Range Forecasts with the GISS Model of the Global Atmosphere

LEONARD M. DRUYAN

Institute for Space Studies, Goddard Space Flight Center, NASA, New York, N. Y. 10025

(Manuscript received 5 November 1973, in revised form 1 February 1974)

ABSTRACT

A nine-layer, primitive equation (PE) model of the global atmosphere developed at the Goddard Institute for Space Studies (GISS) has been used to generate six 48-hr forecasts during December 1972 and January 1973. Operational analyses north of 18N and experimental global analyses made available by the National Meteorological Center (NMC), NOAA, were used as the initial conditions; the operational analyses were used to verify the forecasts at 12-hr intervals over the northern hemisphere north of 22N. The combined analyses were used to verify the forecasts in the global domain.

Root-mean-square errors of the sea-level pressure, 1000-mb heights, and vector geostrophic winds, and 500-mb heights and vector geostrophic winds indicate that the GISS forecasts have skill comparable to those made by operational PE models.

A summary of the 36-hr evolution of extratropical cyclones shows that their speed of propagation is systematically too slow and their central pressures are systematically too high, as has already been documented for the NMC PE model forecasts.

Forecasts of the surface temperature, computed by vertical extrapolation from the model's two lowest levels, and verified quantitatively over North America and qualitatively over the United States, show considerable skill.

1. Introduction

A multi-level numerical primitive equation (PE) model of the global atmosphere developed at the Goddard Institute for Space Studies (GISS) was recently described in detail by Somerville *et al.* (1974). The version of the model that was tested has nine vertical levels and a horizontal grid-spacing of 4° in latitude and 5° in longitude for an effective grid point separation averaging slightly more than 400 km. (While grid points are less than 400 km apart in polar regions, internal smoothing needed to prevent high-latitude computational instability effectively decreases horizontal resolution.) The model has been used extensively for observing system simulation experiments, asynoptic data assimilation studies, and experimental long-range forecasting, largely in support of the Global Atmospheric Research Program (e.g., Jastrow and Halem, 1973). The validity of the simulation studies and the success of real-data assimilation techniques depends in large measure on the short and medium range forecast skill of the model.

In the original presentation of the model (Somerville *et al.*, 1974), some verification statistics were presented for five real-data, 48-hr forecasts and, in addition, a detailed discussion of the synoptic verification of a typical forecast was included. The purpose of this paper is to present a more thorough analysis of these forecasts during December 1972 and January 1973 (now increased to six in number) and to discuss their skill in the light of information currently available for other PE models.

The times of initial conditions for the forecasts are:

1200 GMT 12 December 1972
0000 GMT 22 December 1972
0000 GMT 9 January 1973
0000 GMT 11 January 1973
1200 GMT 12 January 1973
1200 GMT 13 January 1973.

The operational model at the National Meteorological Center (NMC) (Shuman and Hovermale, 1968), has a vertical resolution of six layers and a horizontal grid spacing at 60N of 381 km. The operational model at Fleet Numerical Weather Central (FNWC), U. S. Navy (Kessel and Winninghoff, 1972), has five vertical layers and a grid spacing at 60N also of 381 km. The general circulation model at the Geophysical Fluid Dynamics Laboratory (GFDL), NOAA, has also been used to generate a series of experimental extended-range predictions (Miyakoda *et al.*, 1972); the version of the GFDL model used has nine vertical levels and a horizontal grid spacing of 270 km at mid-latitudes.

2. Initialization and verification

The data sets used for the initial conditions for each of the six forecasts were supplied by NMC and were taken from NMC's operational data analysis north of latitude 18N and from an experimental global NMC data set south of 18N. These data did not include humidities for the model's three uppermost layers, so climatological values were used. The only other manip-

TABLE 1. Averages and standard deviations (in parentheses) of errors for the six GISS SLP and 1000-mb height forecasts, FNWC PE SLP forecasts and NMC and GFDL PE model 1000-mb height forecasts.

Type of verification error	Elapsed time 12 hr	24 hr	36 hr	48 hr
1. GISS SLP rms (mb)	3.7 (0.5)	4.7 (0.4)	6.1 (0.4)	7.2 (0.3)
2. FNWC SLP rms (mb)	3.0	4.5	5.8	7.0
3. GISS SLP S_1	46 (2)	56 (2)	66 (2)	73 (1)
4. GISS 1000-mb std. dev. (m)	28 (2)	36 (2)	47 (3)	54 (2)
5. NMC 1000-mb std. dev. (m)				
December 1972	23	33	44	53
January 1973	24	34	46	55
6. GFDL 1000-mb std. dev. (m)	—	47	54	60
7. GISS 1000-mb RMSVE (m/s)	6.7 (0.3)	8.0 (0.2)	9.5 (0.2)	10.5 (0.4)
8. NMC 1000-mb RMSVE (m/s)				
December 1972	5.4	7.4	8.9	10.2
January 1973	5.5	7.6	9.2	10.6
9. GISS 1000-mb correlation coefficient	0.74 (0.05)	0.81 (0.04)	0.77 (0.05)	0.74 (0.03)
10. NMC 1000-mb correlation coefficient				
December 1972	0.82	0.84	0.82	0.78
January 1973	0.80	0.83	0.79	0.75

ulation of these data was their horizontal and vertical interpolation to the GISS model's grid points. The state of the atmosphere predicted by the model after each of four subsequent 12-hr intervals beginning with the initial condition was compared to the NMC analysis over the Northern Hemisphere valid at the forecast times in order to compute the error statistics shown in Table 1 below. Global errors were computed by comparing the forecasts to the global data sets created by the same procedure used for the initial specification, but valid at the forecast times (see Table 3.)

3. Sea-level pressure (SLP) and 1000-mb height forecasts

GISS forecasts of the near surface can be compared to forecasts made by other PE models by reference to Table 1. Shown for the six GISS forecasts verified against the NMC objective analysis at 12-hr intervals over the Northern Hemisphere north of 22N are:

a) the average root-mean-square (rms) SLP error, line 1;
b) the average SLP S_1 score,[1] line 3;
c) the average standard deviation (std. dev.) of the 1000-mb height error, line 4 (since the arithmetic mean of the errors is small, the standard deviation is no more than a few percent smaller than the rms);
d) the average rms of the magnitude of the 1000-mb geostrophic vector wind error (RMSVE), line 7; and
e) the average correlation coefficient of the observed versus the forecast changes in the 1000-mb heights, line 9.

NMC has supplied us with monthly average error statistics for its PE model 1000-mb forecasts made during December 1972 and January 1973 and verified north of about 18N; these are cited by permission in Table 1. Also included are mean rms SLP errors for FNWC PE forecasts verified north of 20N and made during the period November 1970–January 1971 as shown by Kesel and Winninghoff (Fig. 2, 1972) and the mean standard deviations of 1000-mb forecast height errors for 12 GFDL PE model January forecasts as shown by Miyakoda *et al.* (Fig. 12A, 1972).

Comparison of the GISS average errors given in Table 1 to those cited above for other PE models indicates that the GISS prognostics of the SLP and/or 1000-mb pressure surface are of a quality comparable to those from the operational models at NMC and FNWC and appear to be superior to the experimental GFDL forecasts. It should be noted, however, that the verification of the SLP and the 1000-mb heights is very sensitive to the degree of smoothing implicit in each of the models. The RMSVE and the S_1 score, both based on the verification of gradients, are perhaps more reliable measures of skill.

GISS mean rms height and vector geostrophic wind errors improve with time (from 12 to 48 hr) relative to the corresponding NMC monthly means. This may be due to errors introduced into the initial conditions by the interpolation from the NMC to the GISS grid, possibly adversely affecting the early part of the integration. It is also possible that the advantages of the

[1] The GISS computation of the S_1 score is a computerized adaptation of the procedure originally suggested by Teweles and Wobus (1954). S_1 scores are now computed by NMC on a uniform grid of 5° latitude by 10° longitude (Brown, personal communication, 1974) and they have been cited as an NMC forecast diagnostic by Shuman and Hovermale (1968) and, more recently, by Brown and Fawcett (1972). In effect, the S_1 score is a measure of the error in forecasting the orthogonal components of the horizontal pressure gradient; low S_1 scores are therefore desirable.

The GISS calculation of the S_1 score uses the model's computational grid of 4° latitude by 5° longitude. Thus far there is an insufficient number of GISS model forecasts to be able to assign levels of forecast skill to the range of GISS S_1 scores.

Arakawa (1972) numerical scheme are felt increasingly with time or that both models are approaching the same assymptotic error level. Special attention has been given to the parameterization of sub-grid processes such as the radiation and the moist convection in the GISS model (see Somerville *et al.*, 1974), but it is not known whether these have a significant impact on verification scores within the first 48 hr.

Brown and Fawcett (Fig. 1, 1972) show a skill score that corresponds to an average S_1 score of about 55 for 24-hr NMC SLP forecasts verified only over the 48 conterminous states during the period October 1971–March 1972. S_1 scores computed for the comparable GISS forecasts over the United States averaged 57 and had a standard deviation of 3. Since the S_1 score is sensitive to the distance over which differences are taken and since the GISS grid is different from that used by NMC, a clean comparison between GISS and NMC S_1 scores cannot be made; also, results computed for a region as small as the United States based on only six cases may not be representative.

Table 2 summarizes the model's performance in forecasting the evolution of the 18 extratropical cyclones whose history could be traced for the first 36 hr of the six forecasts. The summary is based on the examination of machine-made plots of the forecast and observed sea-level pressure fields at 12-hr intervals.

The predicted downstream movement of the cyclones averaged about 75% of the observed distance traversed, although the forecast direction of cyclone movement was usually quite accurate. The cross-track error (computed as the perpendicular distance from the forecast cyclone position to the observed trajectory, reconstructed in 12-hr segments) was never much greater than one grid length and was usually less. The tabulated results show a definite negative correlation between the observed lengths of the cyclone trajectories and the percent of the downstream movement that was forecast: i.e., the model forecasts a smaller percent of the movement of fast moving storms.

The tendency for PE forecasts of the sea-level pressure field to underestimate the downstream movement of cyclones has been documented for the NMC operational forecasts for the winter of 1969–1970 by Leary (1971). From a total of 323 cyclones whose evolution was traced for 36 hr, 55% were forecast west of the observed position, 7% due north or south and only 38% east of the observed position at the end of the forecast period. These results were taken to indicate a systematic slowness in the NMC PE forecasts of cyclone movement.

Miyakoda *et al.* (1971) have shown that the speed of traveling waves as forecast by numerical techniques increases with the horizontal grid resolution; underresolving such waves causes underestimation of their phase speed. It follows that, for a given grid resolution, this sluggishness increases with decreasing wavelength.

TABLE 2. Summary of the verification of the 36-hr evolution of extratropical cyclones.

Cyclone number	Total distance moved (km)	Percent of downstream movement that was forecast	Cross-track position error (km)	Total change in central pressure (mb)	Percent of deepening or filling that was forecast
1	3100	42	0	+10	140
2	3000	57	0	−27	52
3	0	*	0	0	*
4	3400	50	0	−11	0
5	1800	61	0	+4	150
6	1500	100	0	−10	0
7	650	92	0	+5	260
8	2000	65	0	−12	33
9	800	100	0	−4	0
10	1800	72	450	−34	29
11	950	79	0	+4	250
12	2800	46	200	+11	127
13	2100	82	150	−54	43
14	500	100	0	+11	64
15	1500	100	500	+5	80
16	750	100	500	−14	29
17	1700	77	200	−14	0
18	550	45	500	+8	200

* Correctly forecast as stationary but with a central pressure error of 8 mb.

Thus, in an evaluation of NMC 36-hr PE forecasts of the 500-mb height field for December 1967, Brown and Fawcett (Fig. 4, 1972) indicate that 13 out of 16 of the shorter waves (of wavelength 15–25° latitude) were indeed forecast to propagate too slowly as compared with observation, whereas an average phase speed closer to the observed was computed for waves of longer wavelength (35–45° latitude). Inasmuch as the fast-moving cyclones are usually associated with the short wavelength perturbations of the midtroposphere (Godske *et al.*, 1957, p. 563), some of the systematic slowness of their forecast movements by the GISS model can be attributed to the limited horizontal resolution of the grid.

Table 2 also shows that, on the average, only 21% of the observed deepening of cyclone central pressure is forecast by the model. (A forecast for filling when deepening is observed is scored as zero percent.) On the other hand, for those cyclones that filled during the 36-hr period, the average forecast increase in cyclone central pressure was 159% of the observed. (A forecast for deepening when filling is observed is scored as zero percent.) Thus, there is a tendency for cyclones to be forecast more shallow than observed; the forecasts do not deepen them enough and fill them too much. This characteristic was also documented by Leary (Table 2, 1971) for the NMC 6-layer PE model forecasts. Based on 190 cyclones whose 36-hr evolution was traced for the winter of 1969–1970, the tabulated data show that NMC forecasts accounted for an average of only 23% of the observed deepening but 125% of the observed filling. This model deficiency may also result from the

limitations on horizontal resolution imposed by the grid spacing. In experiments by Miyakoda *et al.* (Fig. 4, 1971), it was evident that increasing the midlatitude horizontal resolution from 540 km to 270 km had the effect of maintaining deeper cyclones near the surface.

4. 500-mb height forecasts

Table 3 shows the mean error statistics at 12-hr intervals for the six GISS forecasts of the 500-mb height field verified over the Northern Hemisphere north of 22N. The means and standard deviations of the rms height error, the standard deviations of the height errors, the S_1 scores (see Section 3), the rms of the magnitudes of the vector geostrophic wind errors, and the correlation coefficients of the forecast versus the observed height changes are included.

To provide a comparison, several verification statistics computed for 500-mb forecasts from other PE models have been compiled and are also shown in Table 3. Included are mean forecast errors for the models at NMC (quoted by permission) for December 1972 and January 1973 over the Northern Hemisphere north of 18N, at FNWC (Kesel and Winninghoff, 1972) between November 1970 and January 1971 over the Northern Hemisphere north of 20N and at GFDL (Miyakoda *et al.*, 1972) for twelve January forecasts over the entire Northern Hemisphere.

Comparison of the mean forecast errors for other PE models with those from the GISS model (Table 3) indicates that GISS forecasts of the 500-mb height field show about the same relative skill as those near the surface as discussed in Section 3.

An average S_1 score of about 40 corresponds to the skill score offered by Brown and Fawcett (Fig. 1, 1972) for NMC 36-hr 500-mb forecasts over the 48 conterminous states during 1971. The S_1 scores computed for the comparable GISS forecasts over the United States also averaged about 40 with a standard deviation of 3; nevertheless, the values for the GISS forecasts may not be directly comparable for reasons discussed in Section 3.

5. Global verification

Each of the six forecasts of SLP and 500-mb height were verified against the global data sets (see Section 2) at 12-hr intervals valid at the appropriate forecast times. Table 4 shows the average rms SLP and 500-mb height errors computed for these verifications.

The very sparse coverage of meteorological observations in the Southern Hemisphere renders the verification analysis there highly inaccurate as compared with better observed areas and the "forecast errors" of Table 4 should therefore be interpreted cautiously. Nevertheless, it is perhaps useful to establish some benchmark values for global real-data numerical forecasts.

Comparison of the global errors with those discussed above for the Northern Hemisphere north of 22N shows that the former are consistently smaller. This is because the low-gradient regimes of the tropics and, to a lesser extent, the summertime Southern Hemisphere contribute proportionally lower errors to the global rms.

Baumhefner (1970) has verified global forecasts made on the two-layer PE model developed at the National Center for Atmospheric Research. He reports a growth in the rms SLP error of 4.4 mb after 24 hr to 6.3 mb after 48 hr for one case study based on January 1958 data. These statistics must also suffer from the uncertainties of the tropical and Southern Hemisphere verification analyses which, in his case, were assembled from a number of independent sources.

6. Temperature forecasts

a. Upper air

Fig. 1 shows the time evolution of the average rms temperature errors at all heights for the six forecasts verified over the Northern Hemisphere north of 22N.

TABLE 3. Averages and standard deviations (in parentheses) of errors for the six GISS 500-mb height forecasts, and averages of errors for 500-mb height forecasts made with NMC, FNWC, and GFDL PE models.

Verification errors	Elapsed time 12 hr	24 hr	36 hr	48 hr
1. GISS rms height errors (m)	32 (4)	40 (3)	55 (4)	62 (6)
2. FNWC rms height errors (m)	32	43	58	67
3. GISS std. dev. of height errors (m)	31	40	54	62
4. NMC std. dev. (m)				
December 1972	30	38	49	60
January 1973	32	39	53	65
5. GFDL std. dev. (m)	—	54	70	80
6. GISS S_1 scores	29 (1)	35 (1)	41 (2)	45 (2)
7. GISS RMSVE (m sec^{-1})	7.2 (0.5)	8.7 (0.5)	10.4 (0.6)	11.4 (0.9)
8. NMC RMSVE (m sec^{-1})				
December 1972	6.0	8.0	9.5	10.9
January 1973	6.3	8.0	9.7	11.1
9. GISS correlation coefficients	0.75 (0.05)	0.82 (0.03)	0.79 (0.04)	0.78 (0.04)
10. NMC correlation coefficients				
December 1972	0.79	0.85	0.84	0.81
January 1973	0.76	0.85	0.83	0.80

TABLE 4. Averages and standard deviations (in parentheses) of global errors for the six GISS SLP and 500-mb height forecasts.

Type of verification	Elapsed time 12 hr	24 hr	36 hr	48 hr
1. Global rms SLP error (mb)	3.4* (0.5)	4.4 (0.5)	5.0 (0.3)	5.7* (0.4)
2. Global rms 500-mb height error (m)	32* (10)	40 (8)	46 (4)	51* (4)

* Based on only five cases due to missing verification data in the Southern Hemisphere.

At any forecast time the minimum error is at about 600 mb and the errors increase above and below that level. These results can be compared to those presented by Miyakoda *et al.* (1972), which included a similar vertical-time section of average rms temperature errors based on six GFDL PE model forecasts. The level of minimum temperature error for the GFDL forecasts is about 300 mb (after 48 hr) and a secondary minimum appears at about 70 mb. The GISS errors are lower than those computed for the GFDL model in the middle and lower troposphere while the opposite is true above 400 mb. These results are probably due in part to the disparate vertical resolution of the GISS and GFDL models. The former model has only three computational levels above 350 mb and six below whereas the latter model has four levels above 350 mb and five below. At the lowest model level (~945 mb for GISS) after 48 hr the GISS average rms error of 5K is about 1K smaller than the GFDL one; at about 600 mb the GISS error is about 0.7K smaller; and above the 300-mb level the GFDL average rms temperature errors after 48 hr are more than 1K smaller than the corresponding GISS values.

b. Surface

The GISS model forecasts of the surface temperature are computed by extrapolating the temperatures at the two lowest model levels to the surface pressure by assuming that the temperature varies linearly with the logarithm of pressure. This procedure has proved superior to deriving the surface temperature based on the model's present parameterization of the boundary layer heat budget.

A search was made for the most suitable data for verifying forecasts of the surface temperature. The NMC numerical procedure for specifying the surface temperature field uses values that are horizontally interpolated from station observations in data-rich areas and the equivalent of a numerical forecast of 1000-mb temperatures elsewhere. The resulting gridded temperatures usually contain a positive bias because neither component of the procedure accounts for the designated grid point topography elevation. This is true even for data-rich areas with uneven topography because the majority of the stations are usually situated in valleys so that their observations reflect temperatures warmer than at an altitude more representative of the region as a whole. The 1000-mb temperatures are, of course, too warm over any area of high terrain.

For these reasons, GISS initialization and verification surface temperatures are derived in part from the NMC operational surface temperature analysis and in part from the NMC upper air analysis. The resulting "hybrid" specification uses surface temperatures extrapolated linearly with the logarithm of pressure from the

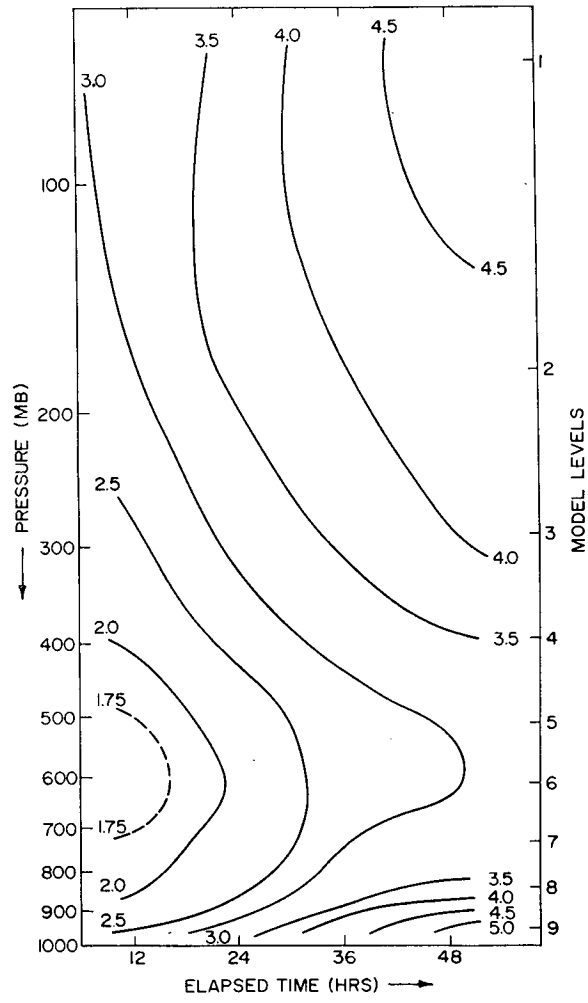

FIG. 1. Vertical-time section showing the temporal evolution of the average root-mean-square temperature errors (°K) verified over the Northern Hemisphere north of 22N for each model level, labeled also according to the usual corresponding atmospheric pressure.

TABLE 5. Average rms errors (°K) and standard deviations (in parentheses) of the six forecasts of surface temperature verified over North America north of 30N.

Elapsed time			
12 hr	24 hr	36 hr	48 hr
3.4 (0.5)	5.0 (0.6)	6.5 (0.7)	7.6 (1.1)

upper air data for grid points whose designated topography elevations are above 500 m and NMC surface temperatures interpolated to GISS grid point elevations at the remaining grid points.

Table 5 shows the average rms surface temperature errors at 12-hr intervals for the six forecasts verified for all land grid points over North America north of 30N.

A separate verification was made for those grid points which are over the eastern United States. Because this area has relatively low topography and a relatively dense network of observing stations, verification temperatures are derived from actual surface observations as explained above; because it is a region of flat topography the observations are quite representative of the designated ground altitudes at each grid point. As a consequence, interpolation errors are at a minimum in the specification of verification surface temperatures over the eastern United States. The rms surface temperature errors computed after 12 hr for this region averaged 3.2K and the average mean absolute error was 2.5K. These results can be compared to those reported by Glahn and Lowry (1972) for computing "today's" maximum temperature at 16 eastern United States stations during December 1970 and January 1971. The mean absolute error for such forecasts, usually made by 0400EST each day, was about 1.8K for man-modified forecasts but 2.4K for those made by the "Model Output Statistics" (MOS) procedure. The skill of the GISS forecasts, by comparison, is particularly noteworthy because the MOS forecasts do not have to verify at a precise time of day. Also, the MOS procedure applies empirical experience, tailored to each observing station, to the output from dynamic models. It is conceded, however, that the GISS forecast, verified against the observed field gridded at 400-km resolution, would tend to yield a lower rms error than if verified by interpolation to the actual observing stations.

Fig. 2 shows the computed distribution of surface temperature and the forecast errors (°F) over the United States for 1200 GMT 13 December 1972, which is 24 hr after the initial condition of one of the six forecasts selected at random as a case study. The observed distribution used to compute the errors has been analyzed from the actual station observations as plotted on the Daily Weather Map (Weekly Series, Environmental Data Service, NOAA) for that date. The 24-hr forecast of surface temperature verifies fairly well over most of the United States' northern border region, the Northeast, the Southern Plains, and over the Rocky Mountains. A too slow forecast movement of the cold front through Louisiana has resulted in computed temperatures there and to the northeast that are about 20F too warm. Errors greater than 10F spread eastward from the front to the central Appalachians and the Louisiana coast; errors of similar magnitude occur over the Central and Northern Plains and the Pacific Coast. The two areas where the computed temperatures are more than 20F too warm over Colorado and Nevada correspond to observed minima below 0F that were not resolved by the model forecast.

In Fig. 3, (a) and (b) show the distribution of maximum and minimum surface temperatures over the United States, respectively, forecast from the initial conditions at 1200 GMT 12 December 1972 for the 24-hr period ending 0600 GMT 14 December 1972. Also shown are the forecast errors, computed from observed maxima and minima analyzed from the actual station observations as plotted on the Daily Weather Map (Weekly Series, Environmental Data Service, NOAA) for 14 December 1972. Figs. 3e and 3d show the distribution of normal daily maximum and minimum surface temperatures for December derived from the *Climatic Atlas of the United States* (U. S. Department of Commerce, 1968) and their departures from the observed values for 14 December 1972, which can be regarded as the errors of a climatological forecast.

A very skillful forecast of the maximum temperatures (Fig. 3a) has been made over most of eastern United States, the Great Lakes region, the northern Rockies, and the West Coast. The model forecasts of maximum temperature are superior to climatology (closer to the observed than the normal maxima, shown in Fig. 3c) in all other regions except over the central Appalachians, where they are too warm, and extreme southern California and eastward to Texas where they are too cold.

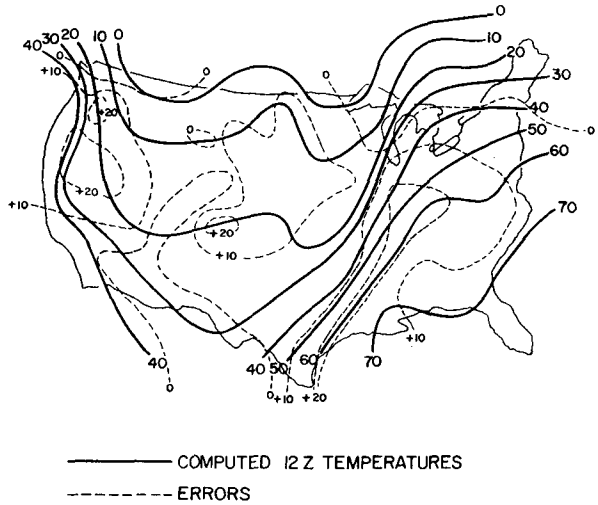

FIG. 2. The 24-hr forecast surface temperatures and errors (°F) over the U. S. made from the initial conditions of 1200 GMT 12 December 1972.

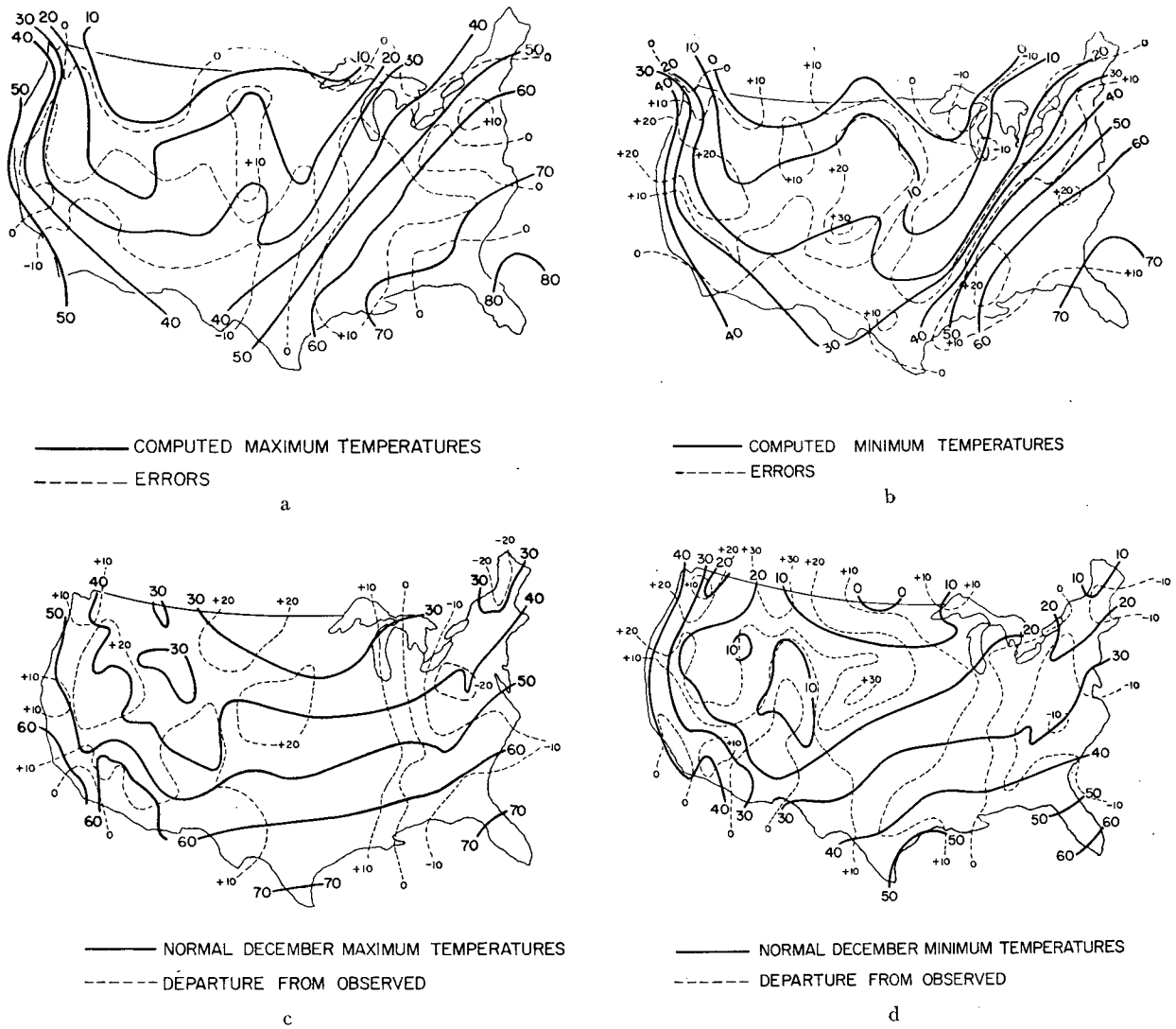

FIG. 3. (a) Forecast maximum surface temperatures and errors (°F) over the U. S. for the 24-hr period ending 0600 GMT 14 December 1972 from the initial conditions of 1200 GMT 12 December 1972. (b) Forecast minimum surface temperatures and errors (°F), otherwise the same as for (a). (c) Normal maximum surface temperatures (°F) for December and their departure from the observed maxima. (d) Normal minimum surface temperatures (°F) for December and their departures from the observed minima.

A very skillful minimum temperature forecast (Fig. 3b) has been achieved over the St. Lawrence River—eastern Great Lakes region, the central and northern Mississippi valley, the Southwest, and the extreme Northwest. The minimum surface temperature forecast is also better than climatology over the central plains states; it is no better than climatology over the southeastern states, the lower Mississippi valley, the Northern Plains, the Rocky Mountains, and the Pacific Coast.

7. Case study

The following discussion pertains to the forecast made from the initial conditions of 1200 GMT 12 January 1973. Fig. 4 (a–e) shows the initial 500-mb height field and the 24-hr and 48-hr forecast and verification 500-mb height fields over the northwestern quadrant of the global domain. Fig. 5(a–e) shows the initial sea-level pressure distribution and the 24-hr and 48-hr forecast and verification sea-level pressure charts for the same region. In Fig. 5 (b and d), the pressure at each grid point has been smoothed by weighting the unsmoothed grid point value equally with the arithmetic mean of the unsmoothed values at the four surrounding grid points.

a. 500 mb

A short wave that moved eastward and formed a trough over northwest Africa and the Iberian peninsula was skillfully forecast 24 hr after the initial conditions

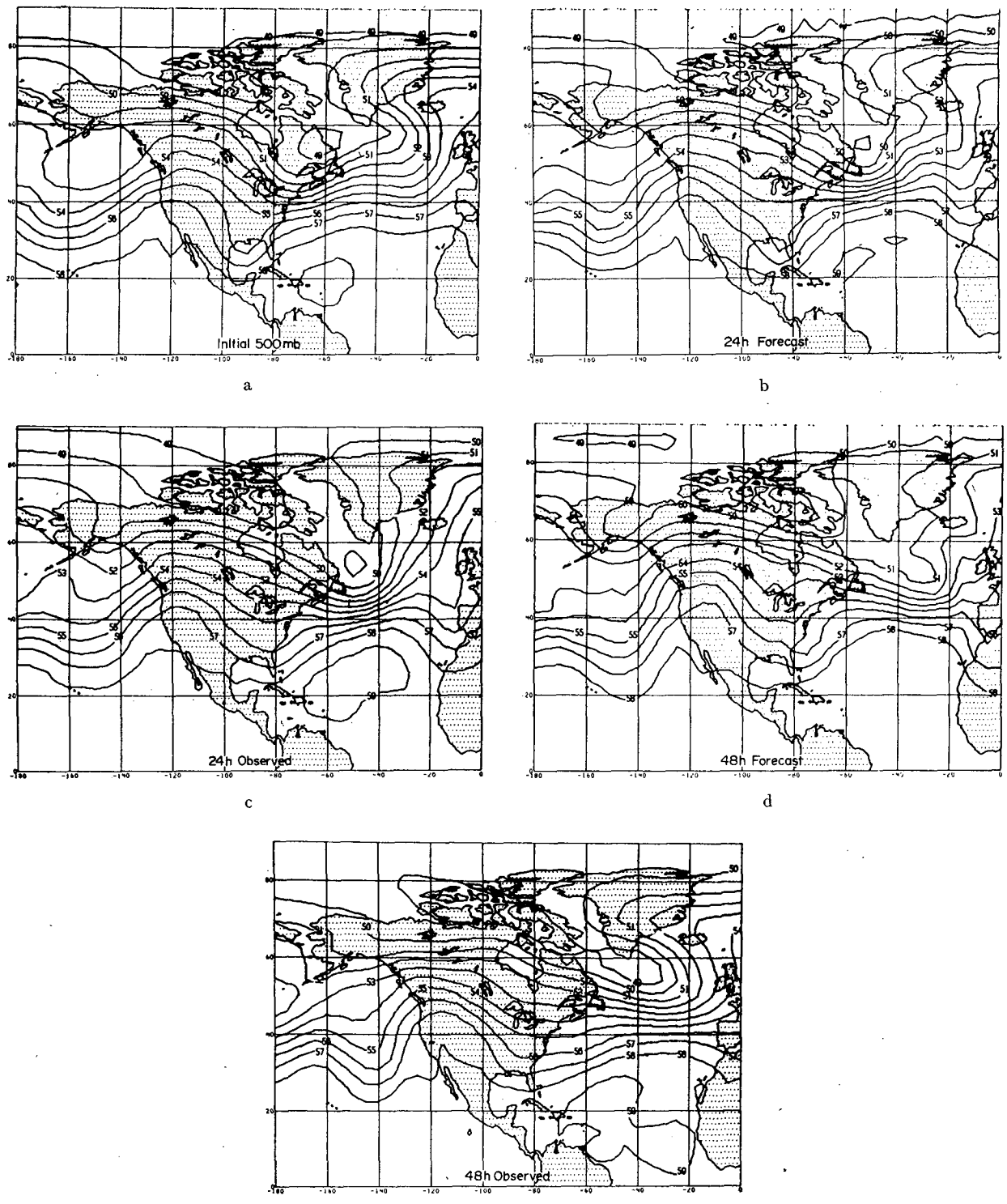

FIG. 4. (a) Initial 500-mb heights (10^2 m) valid 1200 GMT 12 January 1973. (b) The 24-hr forecast 500-mb heights (10^2 m) valid 1200 GMT 13 January 1973. (c) The 24-hr observed 500-mb heights (10^2 m) valid 1200 GMT 13 January 1973. (d) The 48-hr forecast 500-mb heights (10^2 m) valid 1200 GMT 14 January 1973. (e) The 48-hr observed 500-mb heights (10^2 m) valid 1200 GMT 14 January 1973.

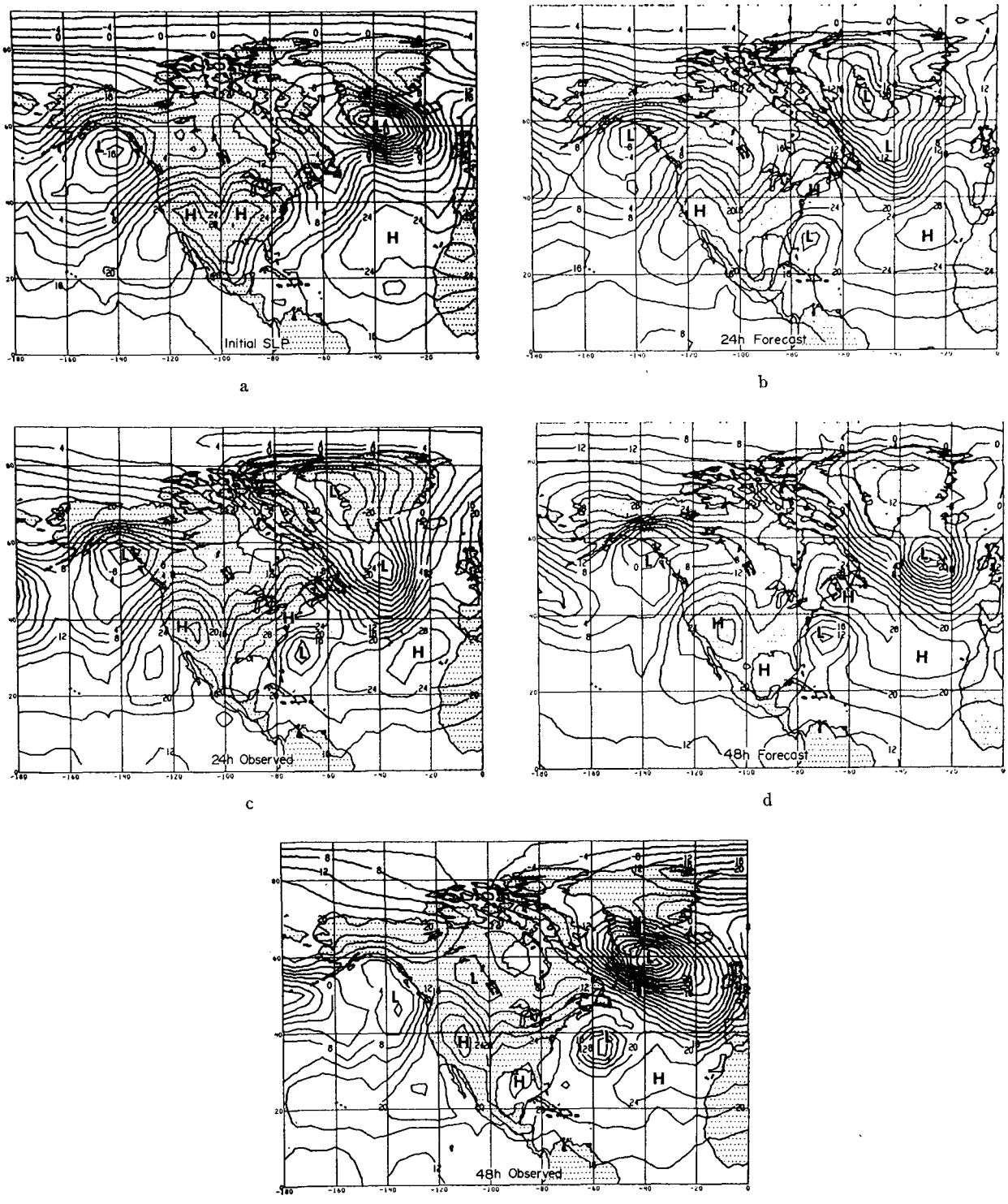

FIG. 5. (a) Initial sea-level pressures (minus 1000 mb) valid 1200 GMT 12 January 1973. (b) The 24-hr forecast sea-level pressures (minus 1000 mb) valid 1200 GMT 13 January 1973. (c) The 24-hr observed sea-level pressures (minus 1000 mb) valid 1200 GMT 13 January 1973. (d) The 48-hr forecast sea-level pressures (minus 1000 mb) valid 1200 GMT 14 January 1973. (e) The 48-hr observed sea-level pressures (minus 1000 mb) valid 1200 GMT 14 January 1973.

(see Fig. 4 (a, b, and c). The breakdown of the ridge over Greenland was not forecast because the propagation speed predicted for the wave southeast of Greenland was too slow. The closed low was correctly moved eastward from Labrador and the trough over the United States Gulf States was translated eastward by slightly less than the correct distance, although the observed height increases along the axis were not forecast. The northwesterly flow persisting over most of the United States and all of Canada was well duplicated by the 24-hr forecast, as was the eastward movement of the shallow trough in the eastern Pacific Ocean.

Fig. 4(d and e) shows the 48-hr forecast and the verification 500-mb height field. The Atlantic Ocean low has developed into an elongated northwest-southeast trough whose apex is in the vicinity of 50N, 25W on both the forecast and verification charts; on the former, however, the low appears to be at least 200 m too shallow. The strong ridge that redeveloped over Greenland during the second 24-hr period was not forecast. The trough that had been over southeastern United States can be traced to two eastward moving short waves—one located at about 65W and the other at 87W. The corresponding longitudes for their forecast positions are 75W and 90W, indicating some slowness in the prediction of their phase speed. The eastward drift of both the ridge over the northwestern United States and western Canada and the shallow trough over the eastern Pacific Ocean was accurately forecast during the second 24-hr prognostic period.

b. Sea level

Fig. 5(a) shows the initial sea-level pressure distribution for the forecast described above, Figs. 5(b) and (c) show, respectively, the 24-hr forecast and verification sea-level pressure charts. A deep cyclone (number 12 in Table 2), initially southeast of Greenland and imbedded in the swift southeast flow aloft, has reached the northwest Greenland coast after 24 hr and has filled about 8 mb. The forecast has moved it only part of the total distance but in the correct direction and has correctly increased the central pressure. A new low (number 13 in Table 2) has formed over the northern Atlantic Ocean from the trough that had been over Newfoundland. The forecast has propagated it to the correct location but decreased the central pressure only 8 mb out of an observed decrease of more than 20 mb. As a result, the pressure pattern over the middle and eastern Atlantic Ocean was forecast quite well, but the strength of the flow south of the cyclone underestimated. Development of an incipient storm was correctly forecast over the Atlantic Ocean just east of Florida. At 24 hr after the initial conditions the forecast position of the new cyclone is only slightly southwest of verification and the pressure minimum was exactly forecast. The small changes in the pattern over the North American continent are fairly well duplicated on the 24-hr forecast. The model predicted a slightly too rapid filling of the low over southeastern Alaska (number 14 in Table 2) but correctly kept its position stationary.

The 48-hr forecast of the sea-level pressure in Fig. 5(d) reproduced the major features of the observed field in Fig. 5(e). The northern Atlantic Ocean cyclone continued to deepen and the forecast does indicate the same trend; however, the predicted deepening rate continued to be only about one-half of the observed, so that the forecast central pressure is about 24 mb too shallow. The strength of the northwesterly flow across the Atlantic Ocean and the easterly circulation over southern Greenland is consequently grossly underestimated. The forecast position of this low lags slightly to the southeast of the observed position. The low over the western Atlantic Ocean was forecast considerably west of its verified position and the predicted decrease in the central pressure is again only about half of the observed. The prognostic maintained the ridge over the New England states and consequently failed to forecast the tight pressure gradient observed over that region due to pressure falls. The persistent pattern of high pressure from the Gulf of Mexico to the northwestern states appears on both charts. The 48-hr forecast of the sea-level pressure did not properly cut off the area of low pressure that has drifted eastward into central Canada; it was also too slow in filling the low southeast of Alaska.

8. Summary

Six 48-hr forecasts have been carried out with the GISS PE model of the global atmosphere during December 1972 and January 1973. Verification of forecast sea-level pressures, 1000-mb heights, and 500-mb heights, as well as 1000-mb and 500-mb vector geostrophic winds, shows that the model has forecast skill comparable with that of operational PE models.

Based on the 36-hr evolution of 18 extratropical cyclones, the model forecasts exhibit a systematic tendency toward underestimating their propagation speeds and overestimating their central pressures. Both deficiencies are probably due in part to inadequate horizontal grid resolution.

Quantitative verification of forecast surface temperatures over the eastern United States shows a forecast skill comparable to that achieved by combined dynamical-statistical procedures (MOS). Qualitative evaluation of a typical forecast of maximum and minimum surface temperatures shows a model skill that surpasses climatology over most regions of the United States within the 48-hr period.

Acknowledgments. The author is indebted to Dr. Peter H. Stone and Dr. Richard C. J. Somerville for their review of the manuscript and their wise counsel. Data processing was carried out by a Computer Sci-

ences Corporation group under the able supervision of Mr. Donald Berman. The author also wishes to thank Mr. Hugh O'Neill of NMC, who facilitated GISS acquisition of NMC data and error statistics, and Dr. Frederick Shuman, director of NMC, who consented to have the latter published. The continuing interest of Dr. Shuman in GISS modeling research is gratefully acknowledged.

REFERENCES

Arakawa, A., 1972. Design of the UCLA atmospheric general circulation model. Technical Report No. 7, Dept. of Meteorology, University of California at Los Angeles.

Baumhefner, D. P., 1970: Global real-data forecasts with the NCAR two-layer general circulation model. *Mon. Wea. Rev.*, **98**, 92–99.

Brown, H. E., and E. B. Fawcett, 1972: Use of numerical guidance at the National Meteorological Center. *J. Appl. Meteor.*, **11**, 1175–1182.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.

Godske, C. L., T. Bergeron, J. Bjerknes, and R. C. Bundgaard, 1957: *Dynamic Meteorology and Weather Forecasting*. Boston, American Meteorological Society, and Washington, Carnegie Institute of Washington, 800 pp.

Jastrow, R., and M. Halem, 1973: Simulation studies and the design of the first GARP global experiment. *Bull. Amer. Meteor. Soc.*, **54**, 13–21.

Kesel, P. G., and F. J. Winninghoff, 1972: The Fleet Numerical Weather Central operational primitive-equation model. *Mon. Wea. Rev.*, **100**, 360–373.

Leary, C., 1971: Systematic errors in operational National Meteorological Center primitive-equation surface prognoses. *Mon. Wea. Rev.*, **99**, 409–413.

Miyakoda, K., R. F. Strickler, C. J. Nappo, P. L. Baker, and G. D. Hembree, 1971: The effect of horizontal grid resolution in an atmospheric circulation model. *J. Atmos. Sci.*, **28**, 481–499.

Miyakoda, K., G. D. Hembree, R. F. Strickler, and I. Shulman, 1972: Cumulative results of extended forecast experiments, I. Model performance for winter cases. *Mon. Wea. Rev.*, **100**, 836–855.

Somerville, R. C. J., P. H. Stone, M. Halem, J. E. Hansen, J. S. Hogan, L. M. Druyan, G. Russell, A. A. Lacis, W. J. Quirk, and J. Tanenbaum, 1974: The GISS model of the global atmosphere. *J. Atmos. Sci.*, **31**, 84–117.

Shuman, F. G., and J. B. Hovermale, 1968: An operational six-layer primitive-equation model. *J. Appl. Meteor.*, **7**, 525–547.

Teweles, S., and H. Wobus, 1954: Verification of prognostic charts. *Bull. Amer. Meteor. Soc.*, **35**, 455–463.

U. S. Department of Commerce, 1968: *Climatic Atlas of the United States*, Environmental Data Service, NOAA, 80 pp.