TECHNICAL MEMORANDUM

85

Field presentation of verification statistics: summer forecasts of 1982 and 1983 compared

R. Nieminen

Research Department

December 1983

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Abstract

A data set containing monthly verification statistics at individual grid points of global fields for the Centre's forecasts from January 1980 to September 1983 has been produced. This verification result data base for grid points will be updated every month. The mean forecast and verifying analysis fields and a range of objective scores can be plotted from this data base for geopotential height, temperature and wind at eight standard pressure levels for different forecast days. The two summer seasons of 1982 and 1983 have been compared using the plotted results of this grid point verification data base.

1. Introduction

An operational system to compute verification scores from time series for individual grid points has been established at ECMWF. The verification results can be presented in chart form for geographical regions anywhere on the globe. The purpose of this paper is to introduce the system and give as examples the comparison of the results for summer 1982 and 1983.

The verification scores are explained in chapter 2, and in chapter 3 the verified products are described and a complete list of the available statistics is given. Global results for the two summer seasons 1982 and 1983 are presented in chapter 4.

2. Verification scores

The objective verification of numerical forecasts has traditionally been carried out for defined areas by comparing the predicted fields and the verifying analysis fields. The calculations are performed for the set of points in the area and the scores thus represent the whole area at a fixed time. The ensemble mean can also be computed for a sample of cases by combining area scores within some time period. A typical example of this is the limited area verification system at ECMWF which produces daily results for which time averaging can be applied to derive monthly mean scores (Nieminen, 1983).

In order to present a detailed geographical distribution of verification statistics, the scores are calculated from time series for individual grid points, and this is done in the system described here. The score now verifies the forecasts at each grid point for the specified time period, and the results for all grid points can be displayed in chart form. This kind of presentation has been used, for example, in the Centre for the systematic errors and rms-errors of height and temperature forecasts (ECMWF Forecast

Reports, 1980-1983). The area means can also be obtained from the grid point scores by averaging the results over defined areas. However, the values of correlation or standard deviation spatially averaged from the grid point scores are not comparable with the corresponding time averaged area results, because the deviations are taken from different mean values, one in time and the other in space, when these scores are calculated before averaging. The formulae for the scores used in this paper are defined in Table 1.

Mean error and root mean square error or standard deviation of error are verification parameters which can readily be displayed in chart form anywhere on the globe. The absolute values of these parameters will vary from one geographical area to another, and the persistence error is often used as a standard of comparison for the rms-error and standard deviation of error to determine the quality of forecasts.

It is more difficult to find a uniform score for the predictive skill, because this score should be able to verify different scales of motion in different areas. The correlation between the predicted and verifying point values within the specified time series estimates how well the observed daily variation at each point is predicted, and the correlation has produced as a measure of the predictive skill in this verification system. Another objective score which can be used for the evaluation of the quality of forecasts is the "normalised" error (defined as the standard deviation of forecast error divided by the persistence error). This score has been compared to the correlation in Figures 1 and 2 (combining the results for three summer months of 1982) and in Figures 3 and 4 (respectively for summer 1983) where similarities can be found in the main features; maximum and minimum values occur approximately in the same regions. However, both the correlation and the normalised error are sensitive to small scale motion represented by small persistence values. Figures 5 and 6 show the standard deviation of persistence error to be compared with the correlation and the normalised error presented in Figures 1 to 4. From all these figures we can see that the poor correlation and normalised error values often appear in the areas where persistence is small.

Table 1 Formulae used to compute verification scores from time series for individual grid points.

Using the symbols:

Ao = initialised analysis

Av = verifying analysis

F = forecast

n = number of cases in the time series (one month used; n varies from 28 to 31)

Σ = summation in time over the verified period

we get the following expressions:

$$1/n \, \Sigma \, Av = \overline{Av} = \text{mean of verifying analysis}$$

$$1/n \, \Sigma \, F = \overline{F} = \text{mean of forecast}$$

$$1/n \, \Sigma \, (F-Av) = (\overline{F-Av}) = \text{mean of forecast error}$$

$$\sqrt{1/n \, \Sigma \, (F-Av)^2} = \text{RMS-error of forecast}$$

$$\sqrt{1/n \, \Sigma \, (Ao-Av)^2} = \text{RMS-error of persistence}$$

$$\frac{\Sigma \left((F-\overline{F})(Av-\overline{Av})\right)}{\sqrt{\Sigma \, (F-\overline{F})^2 \Sigma (Av-\overline{Av})^2}} = \text{absolute correlation of forecast}$$

$$\frac{\Sigma \left((Ao-\overline{Ao})(Av-\overline{Av})\right)}{\sqrt{\Sigma \, (Ao-\overline{Ao})^2 \Sigma (Av-\overline{Av})^2}} = \text{absolute correlation for persistence}$$

The standard deviation of forecast error can be derived from the RMS-error and mean error of forecast as follows:

$$stdev = \sqrt{(RMSe)^2-(meane)^2}$$

The standard deviation of persistence error is assumed to be equal to the RMS-error of persistence (bias = 0 for persistence).

The normalised forecast error (in percentage) is defined by:

$$S_n = 100 \left[1 - \frac{S_f^{\,2}}{S_p^{\,2}} \right]$$

where S_f = standard deviation of forecast error

S_p = standard deviation of persistence error

For vector wind we combine the results of the wind components as follows:

$$\text{mean}(\bar{v}) = \sqrt{[\text{mean}(u)]^2 + [\text{mean}(v)]^2}$$

$$\text{rms}(\bar{v}) = \sqrt{[\text{rms}(u)]^2 + [\text{rms}(v)]^2}$$

$$\text{stdv}(\bar{v}) = \sqrt{[\text{stdv}(u)]^2 + [\text{stdv}(v)]^2}$$

$$\text{correlation } r(\bar{v}) \approx \sqrt{1/2*[r(u)^2 + r(v)^2]}$$

3. Verified products

Global fields on a 3 x 3 degree horizontal resolution are used for this grid point verification from time series. These upper air fields are selected from the post-processed data of the ECMWF model, and they are available in a special data base in the Centre (for description see Nieminen, 1983). In post-processing, forecast fields are interpolated vertically from the model sigma levels to pressure levels, and both the analysis and forecast fields are converted to a 1.5 degree resolution from the archived spectral components. It must also be remembered that in mountain areas data are extrapolated below the ground surface for the lower levels of the model.

Initialised analyses are used for verifying the forecasts of geopotential height, temperature and wind components at eight standard pressure levels from 1000 to 50 mb. The objective scores include the mean error for forecast, the rms-error for forecast and persistence and the correlation for forecast and persistence; mean forecast and analysis fields are also provided for different months. All these parameters are computed from monthly time series for individual grid points. From these basic results the standard deviation of forecast error can be derived from the values of mean error and rms-error. For persistence the standard deviation of error is assumed to be equal to the rms-error (bias = 0 for persistence). The normalised error score can also be provided from the standard deviations of forecast and persistence as was described in the previous chapter. The calculations are performed for the following forecast times: +24, +48, +72, +120, +168 and +240 hours.

Table 2 gives a complete list of the verification statistics for individual grid points which is available in the verification result data base. The results, which are global grid point fields, are stored in packed form on magnetic tapes. These monthly values for different meteorological and verification parameters are, at the time of writing, provided for each month from January 1980 to September 1983, and this data base will be updated after each calendar month.

Table 2: Monthly statistics available from the verification result data base for grid points.

Data form:
- global grid point fields
- resolution: 3 x 3 degree in the horizontal
- each field contains one verification parameter from one month time series for individual grid points
- averaged results for several months can be provided
- plots can be produced for any area in either cylindrical or polar stereographic projections

Levels: 1000, 850, 700, 500, 300, 200, 100 and 50 mb (up to 200 mb for 1980)

Parameters:
- geopotential height
- temperature
- u- and v-component
- vector wind (derived from the components)

Forecast times: +24, +48, +72, +120, +168 and 240 hours (wind to +120 hours for 1980)

Verification results:
- mean forecast field
- mean verifying analysis field
- mean error for forecast
- rms-error for forecast and persistence
- correlation for forecast and persistence
- standard deviation of forecast error (derived from rms-error and mean error)
- standard deviation of persistence error (assumed to be equal to rms-error for persistence)
- normalised error (derived from standard deviations of forecast and persistence error)

4. Some grid point verification statistic from the summer 1982 compared to the results for the summer 1983.

The verification result data base for grid points contains a large volume of global fields; at the time of writing, for each month from January 1980 to September 1983. For this paper some global fields of the results have been plotted for the two summer seasons of 1982 and 1983. First these charts provide a selection of examples of the grid point verification statistics, but on the other hand they also offer very relevant information about the performance of the Centre`s model, because the new spectral model with the envelope orography was implemented on 21 April 1983. The figures presented here are the averaged results of three summer months, June, July and August, for the years 1982 and 1983, and a cylindrical map form was used to plot the fields for this publication. A list of the figures included is given in Appendix 1.

Some major differences between the results for the two summer seasons in these example cases are listed below.

1. The systematic height error has changed at 200 mb. The positive error values have increased for the summer 1983 compared to the summer 1982, as Figures 11 and 12 show.

2. The mainly negative systematic temperature error at 850 mb over the continental areas in the middle and high latitudes has been reduced from the summer 1982 to the summer 1983 (see Figures 13 and 14). However, at the 200 mb level the positive systematic temperature error has increased in the tropical regions during the later summer period (see Figures 15 and 16).

3. The mean error of the vector wind at 850 mb in the tropics has increased in the area from the Indian Ocean to Southern Arabia (see Figures 17 and 18) but decreased in some other tropical regions. At 200 mb the mean error of the vector wind has generally decreased from the summer 1982 to the summer 1983 in the equatorial areas, as Figures 19 and 20 show.

4. Figures 25-30 present the vertical cross-sections of the systematic height error over the latitudes 30 and 45 north and 30 south. These figures show how the barotropic structure of the systematic error is related to the model mountains. From these figures we can also see the changes in the mean height error from the summer 1982 to the summer 1983. The pronounced negative mean error has almost disappeared from the stratospheric levels of the model over these considered latitudes. On the other hand the positive systematic height error has increased around the levels of 200 and 100 mb.

REFERENCES

Nieminen, R., 1983: Operational verification of ECMWF forecast fields and results for 1980-81. ECMWF Technical Report No. 36, 40pp.

ECMWF Forecast Reports, 1980-83: Published monthly (1980) and quarterly (from 1981) by ECMWF.

Appendix 1:

List of figures

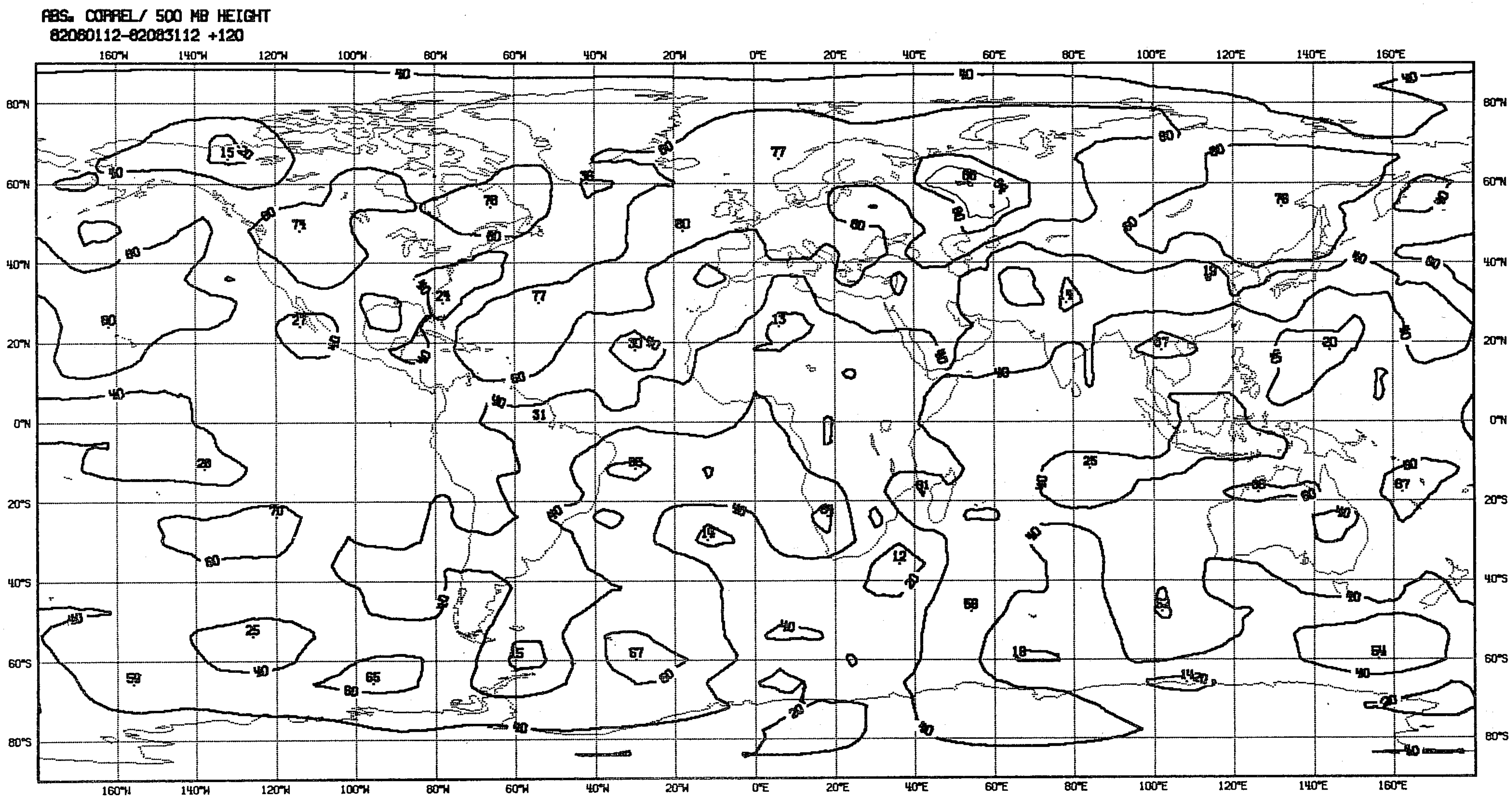

Fig. 1 Absolute correlation for the 5-day forecast of 500 mb height for the three month period of June, July and August 1982 (averaged from the monthly correlations).

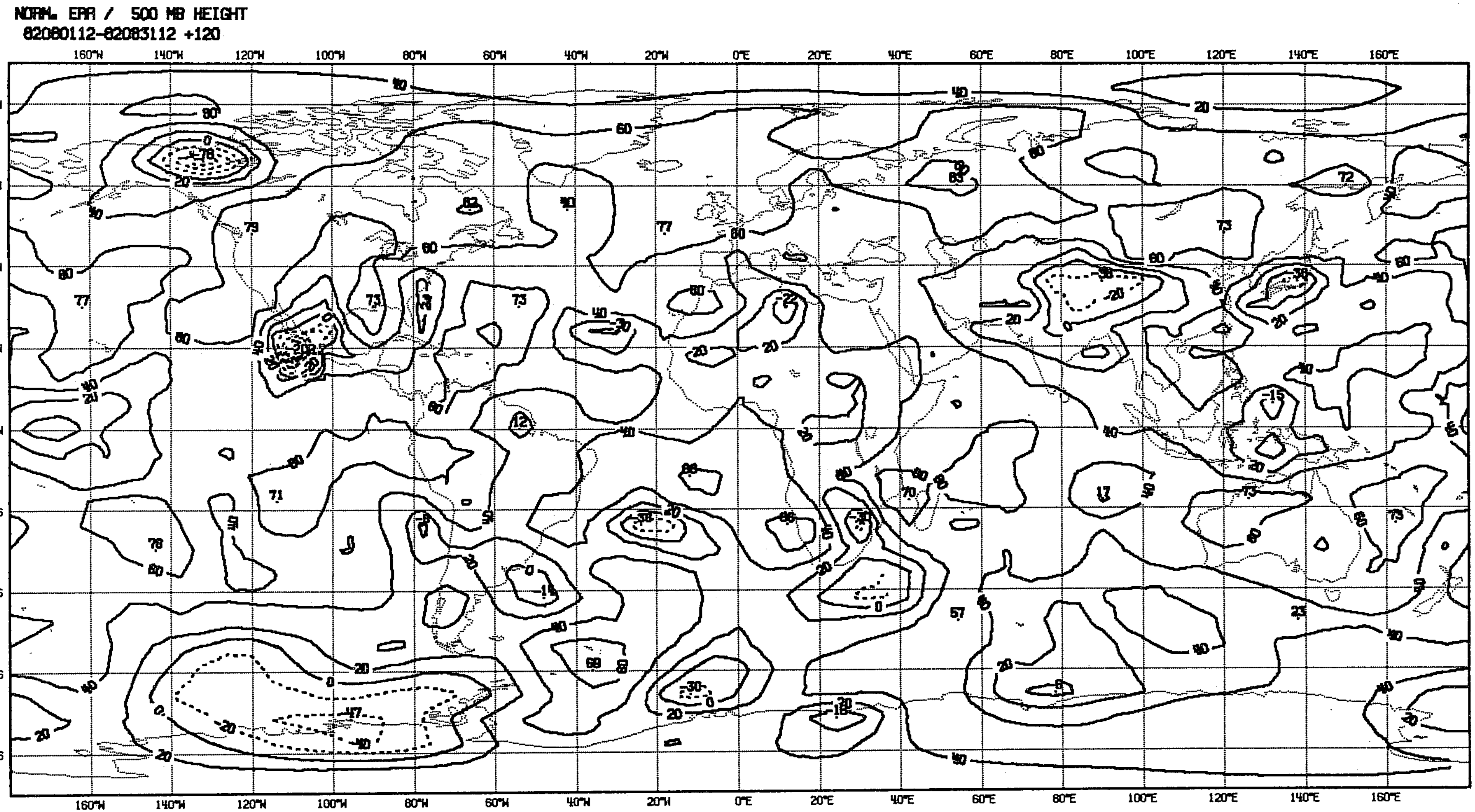

Fig. 2 Normalised error (standard deviation of forecast error divided by persistence error) for the 5-day forecast of 500 mb height for the three month period of June, July and August 1982 (averaged from the monthly scores).

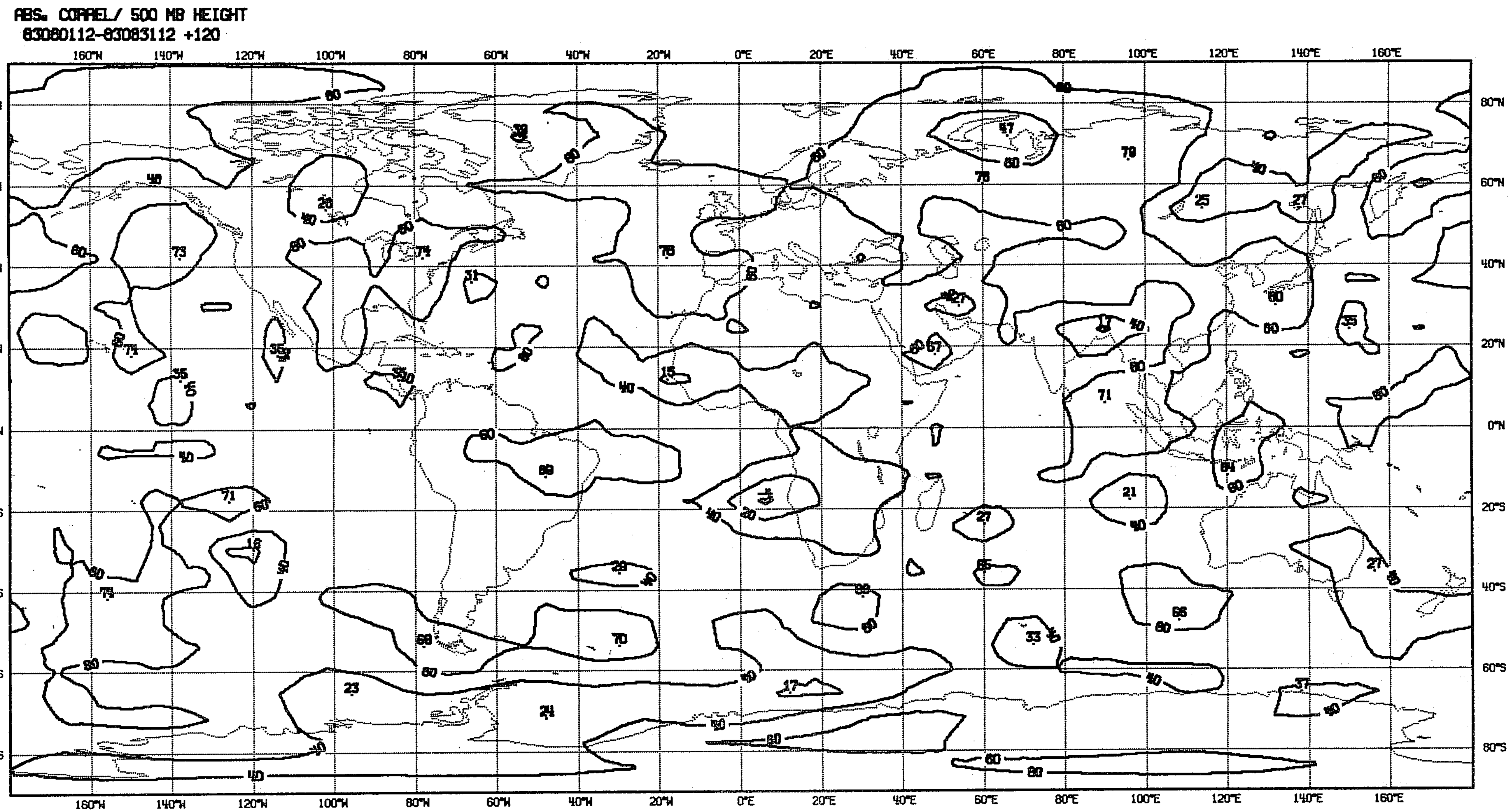

Fig. 3 Same as Fig. 1 but for 1983.

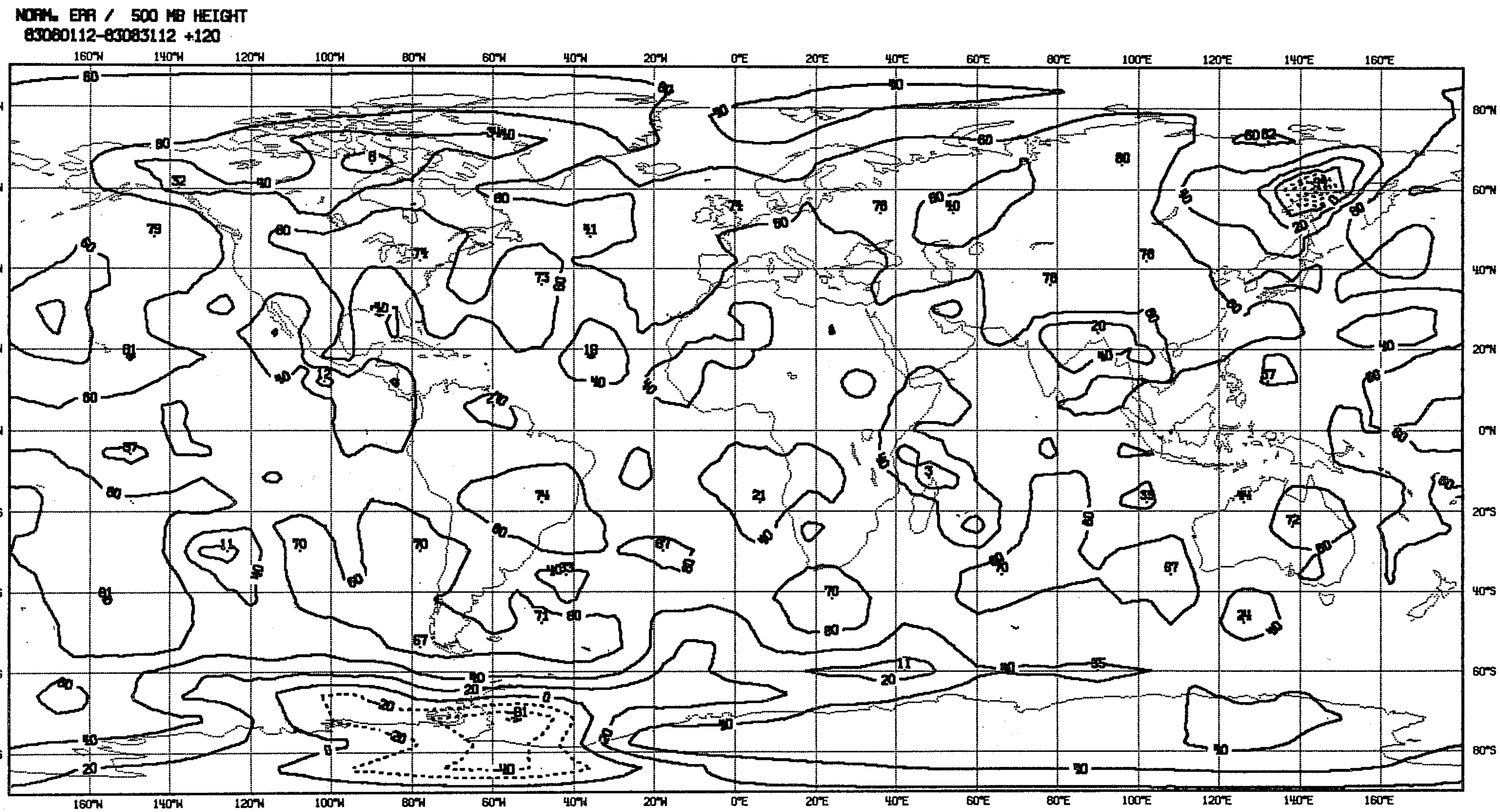

Fig. 4 Same as Fig. 2 but for 1983.

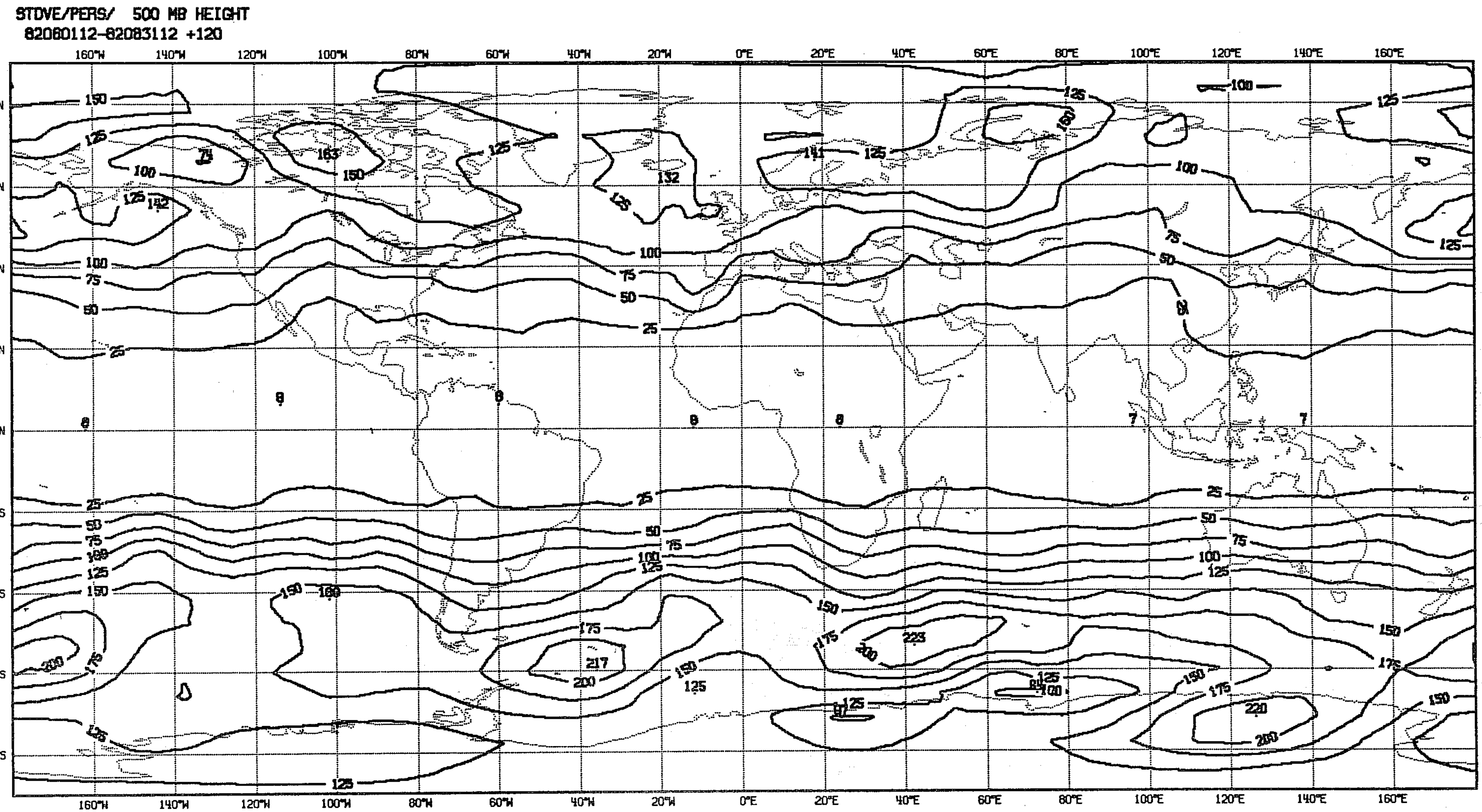

Fig. 5 Standard deviation of persistence error (metres) for the 5-day forecast of 500 mb height for the three month period of June, July and August 1982 (averaged from the monthly scores).

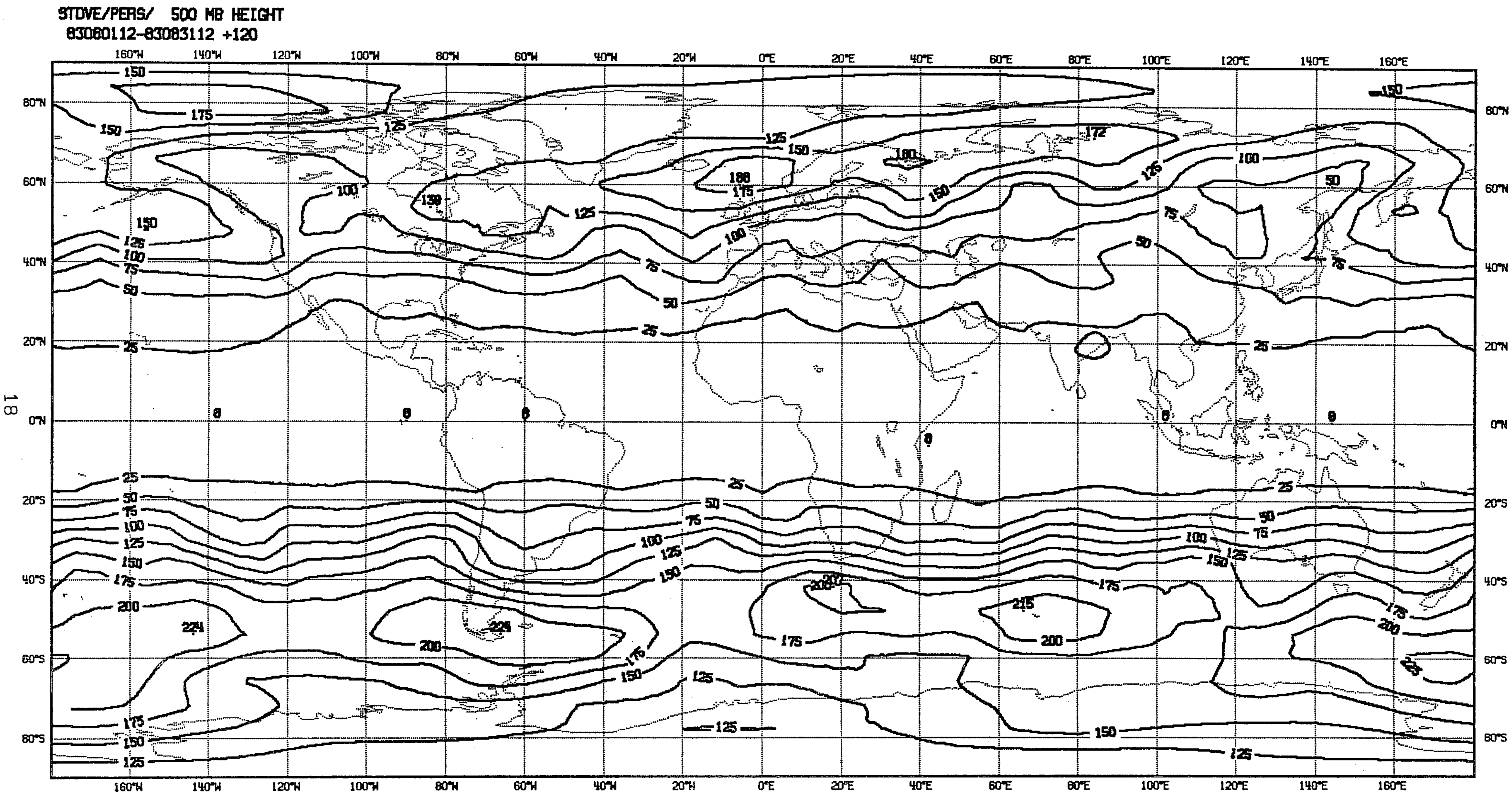

Fig. 6 Same as Fig. 5 but for 1983.

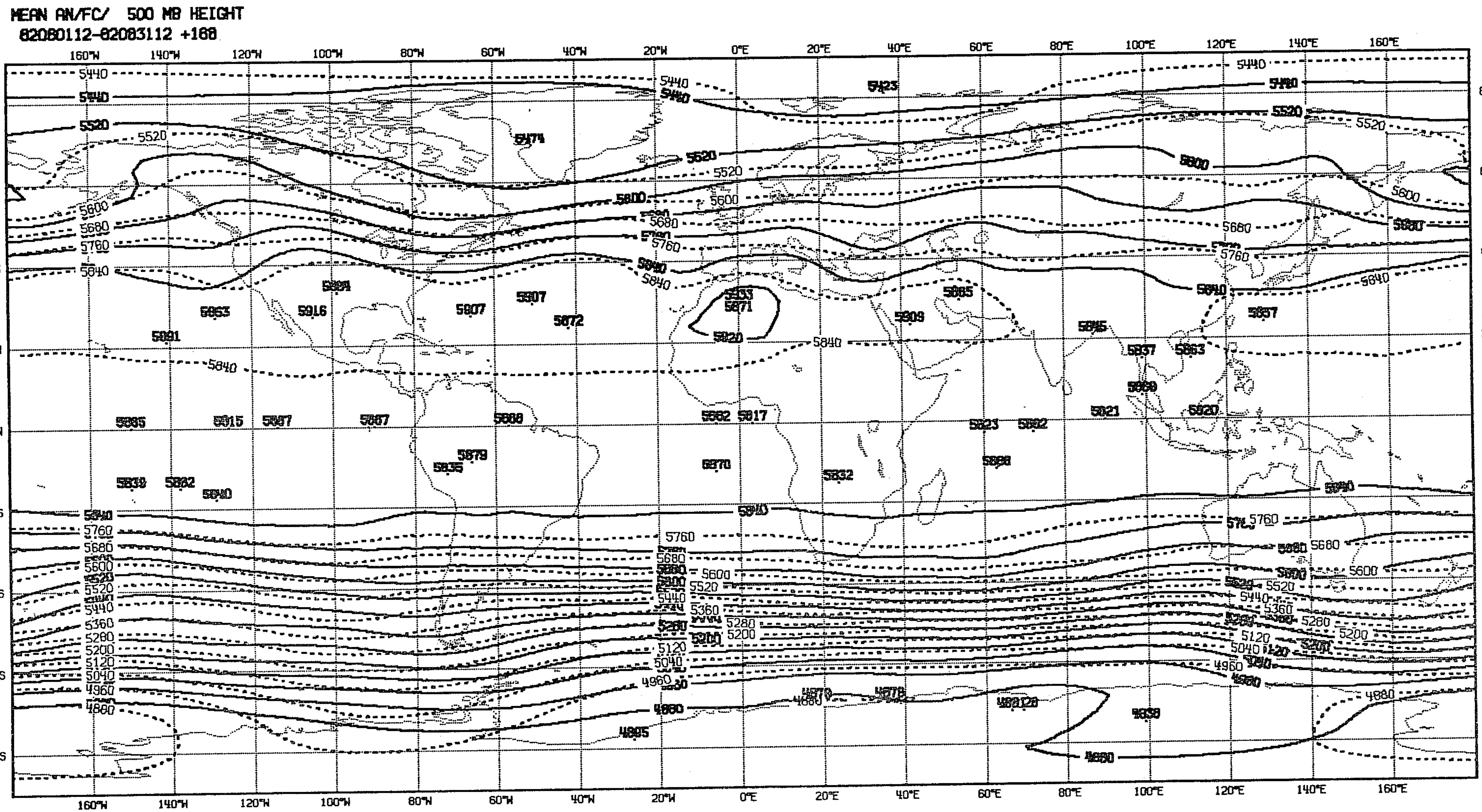

Fig. 7 Mean of the 7-day forecast (dashed line) and the verifying analysis (solid line) of 500 mb height for the three month period of June, July and August 1982.

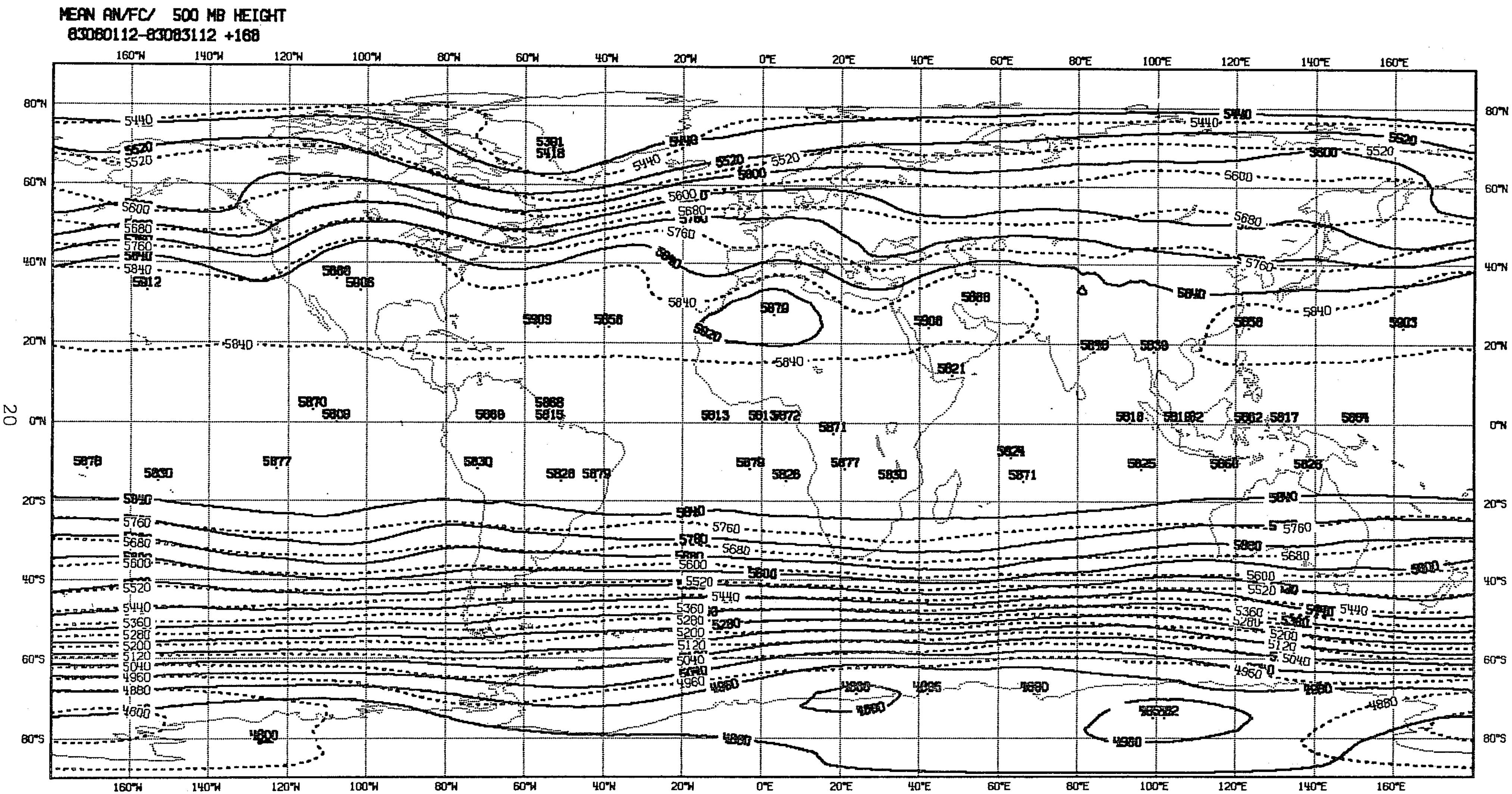

Fig. 8 Same as Fig. 7 but for 1983.

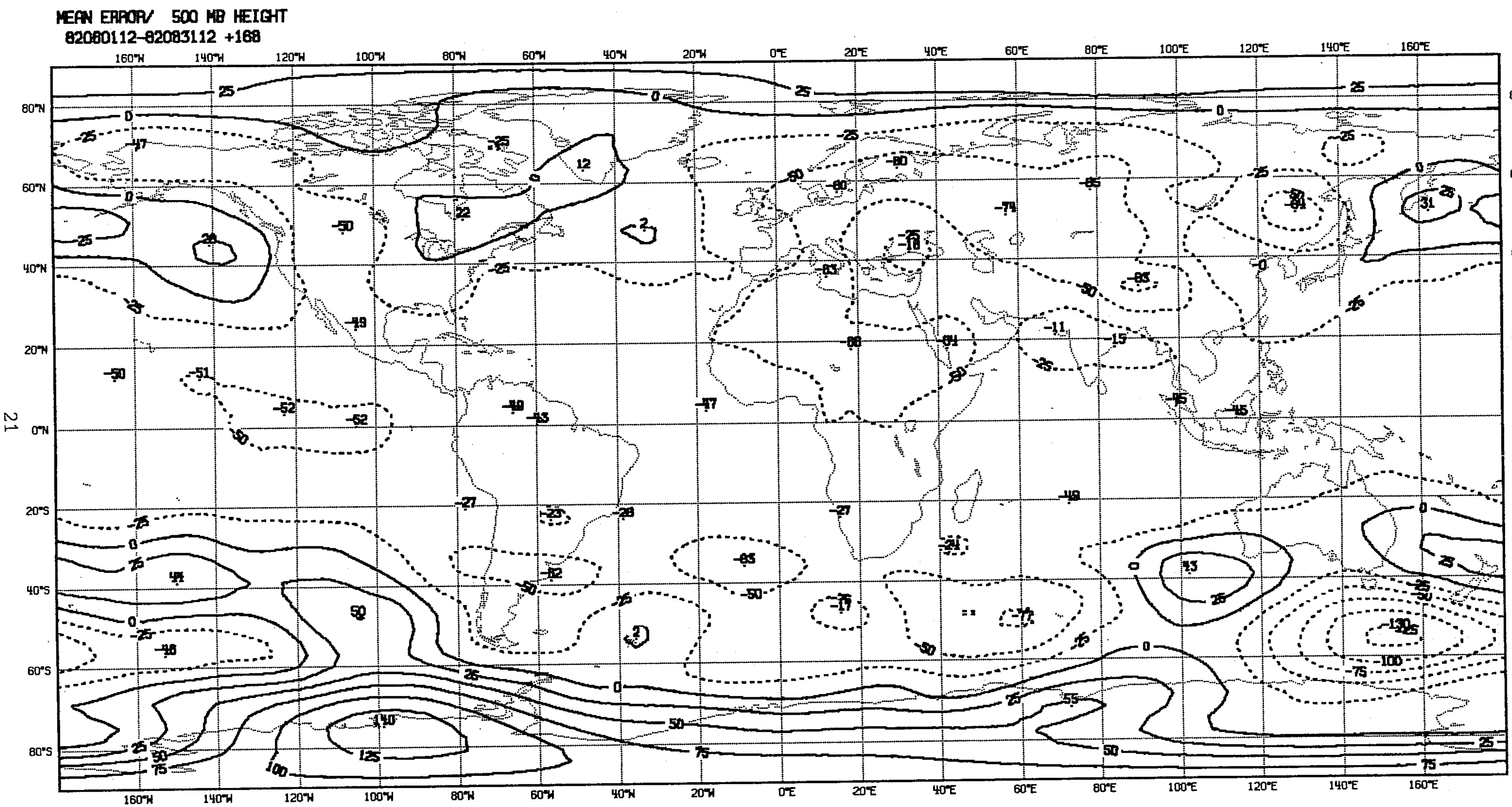

Fig. 9 Mean error (metres) for the 7-day forecast of 500 mb height for the three month period of June, July and August 1982.

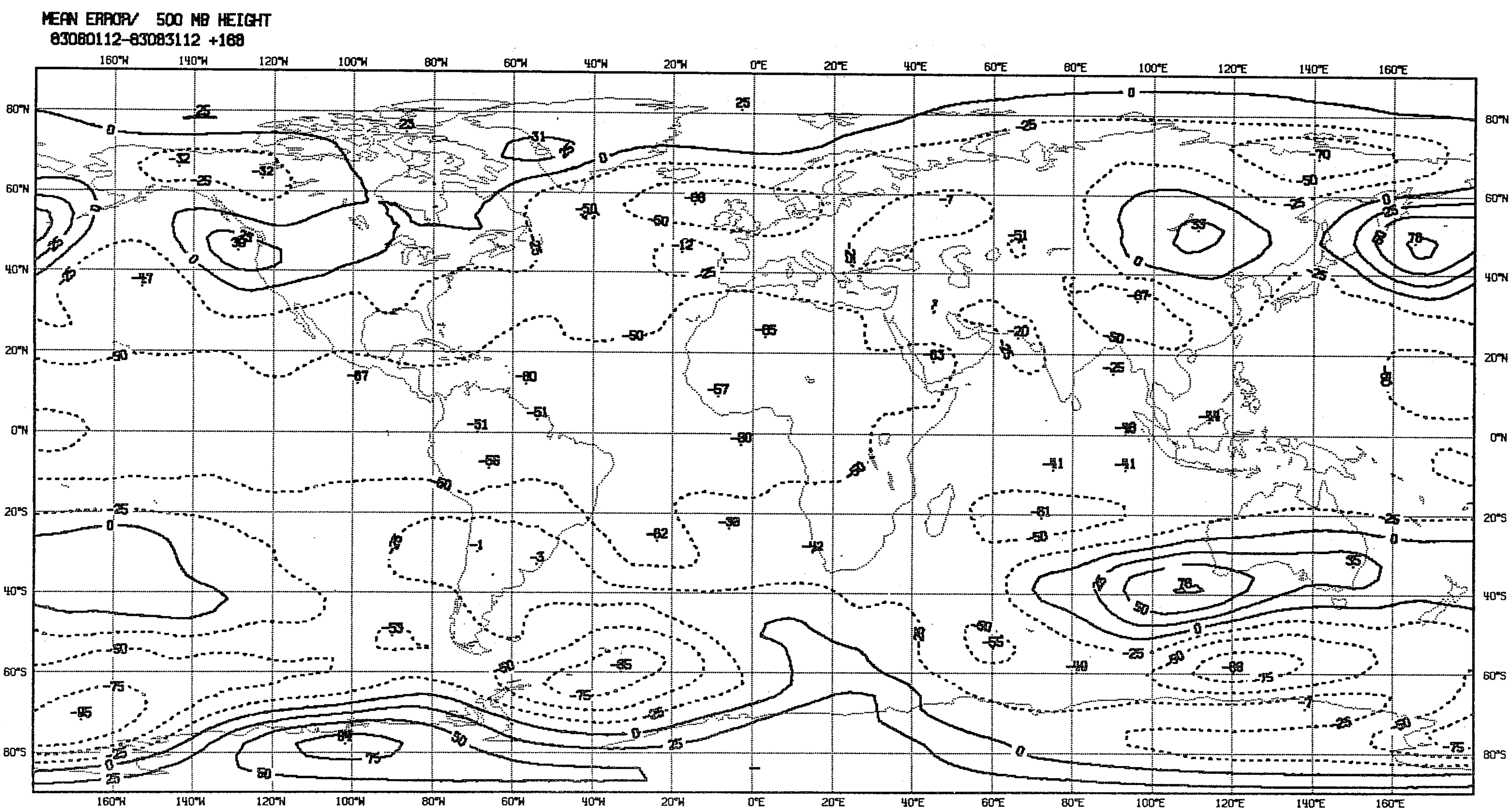

Fig. 10 Same as Fig. 9 but for 1983.

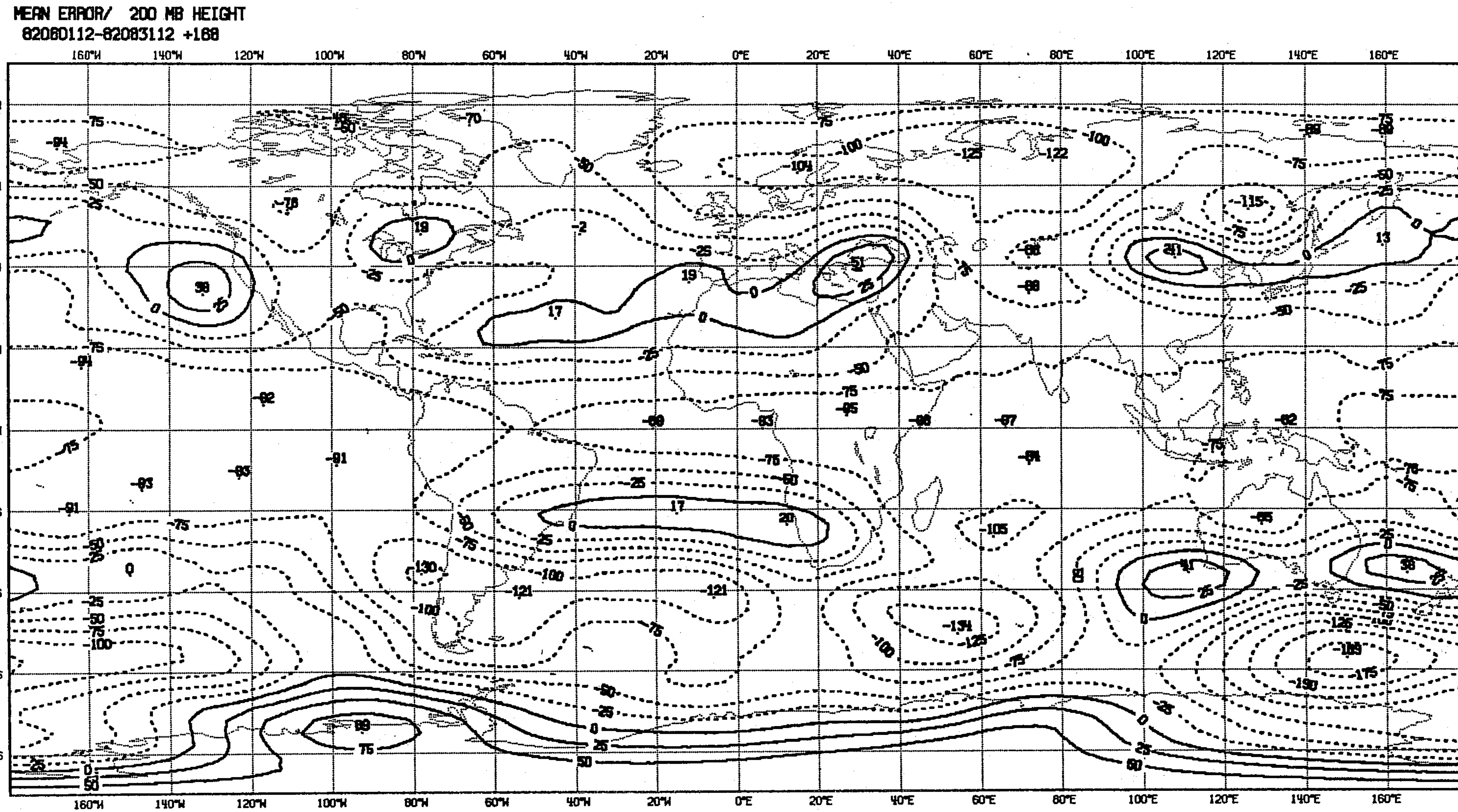

Fig. 11 Mean error (metres) for the 7-day forecast of 200 mb height for the three month period of June, July and August 1982.

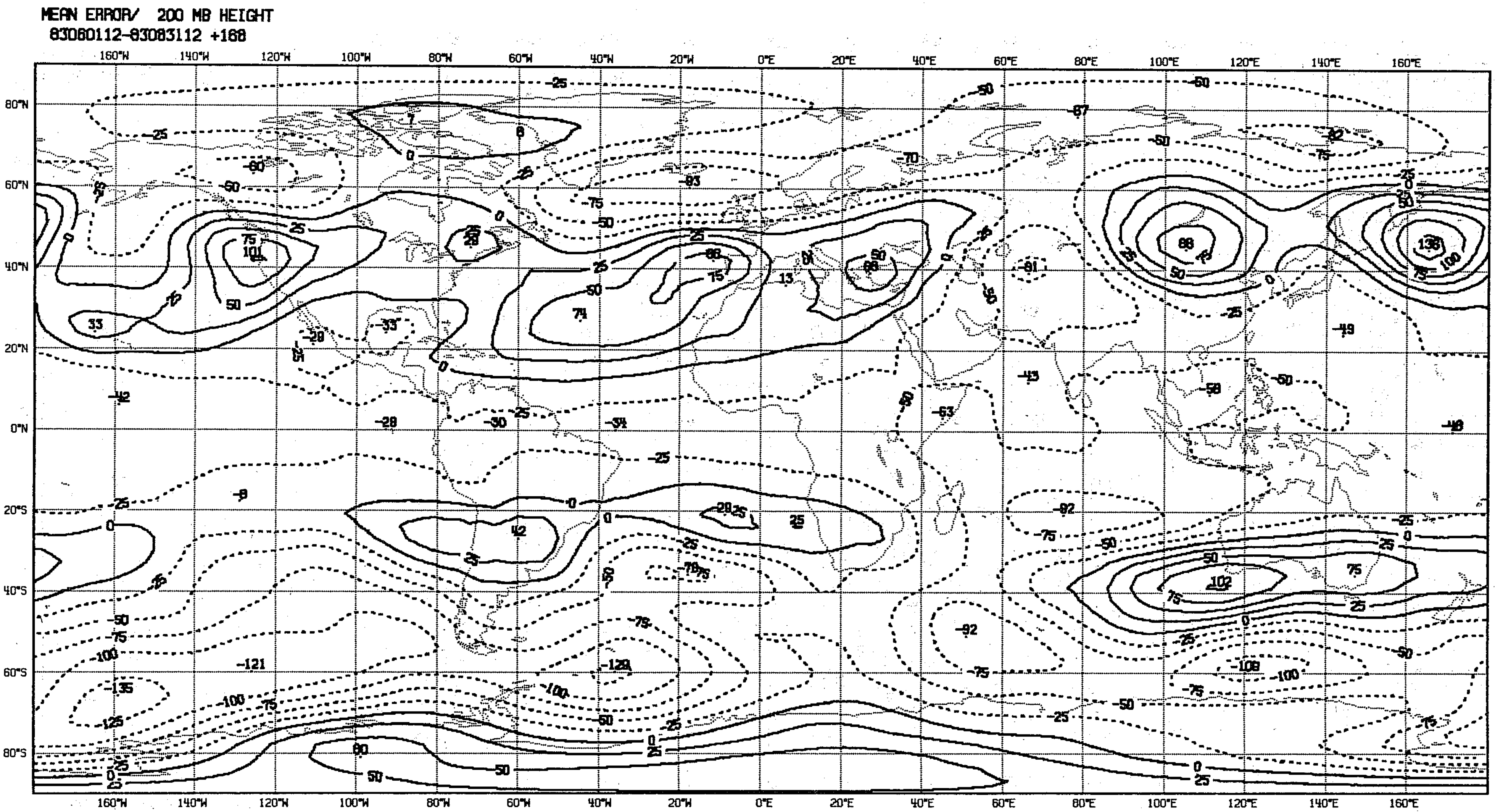

Fig. 12 Same as Fig. 11 but for 1983.

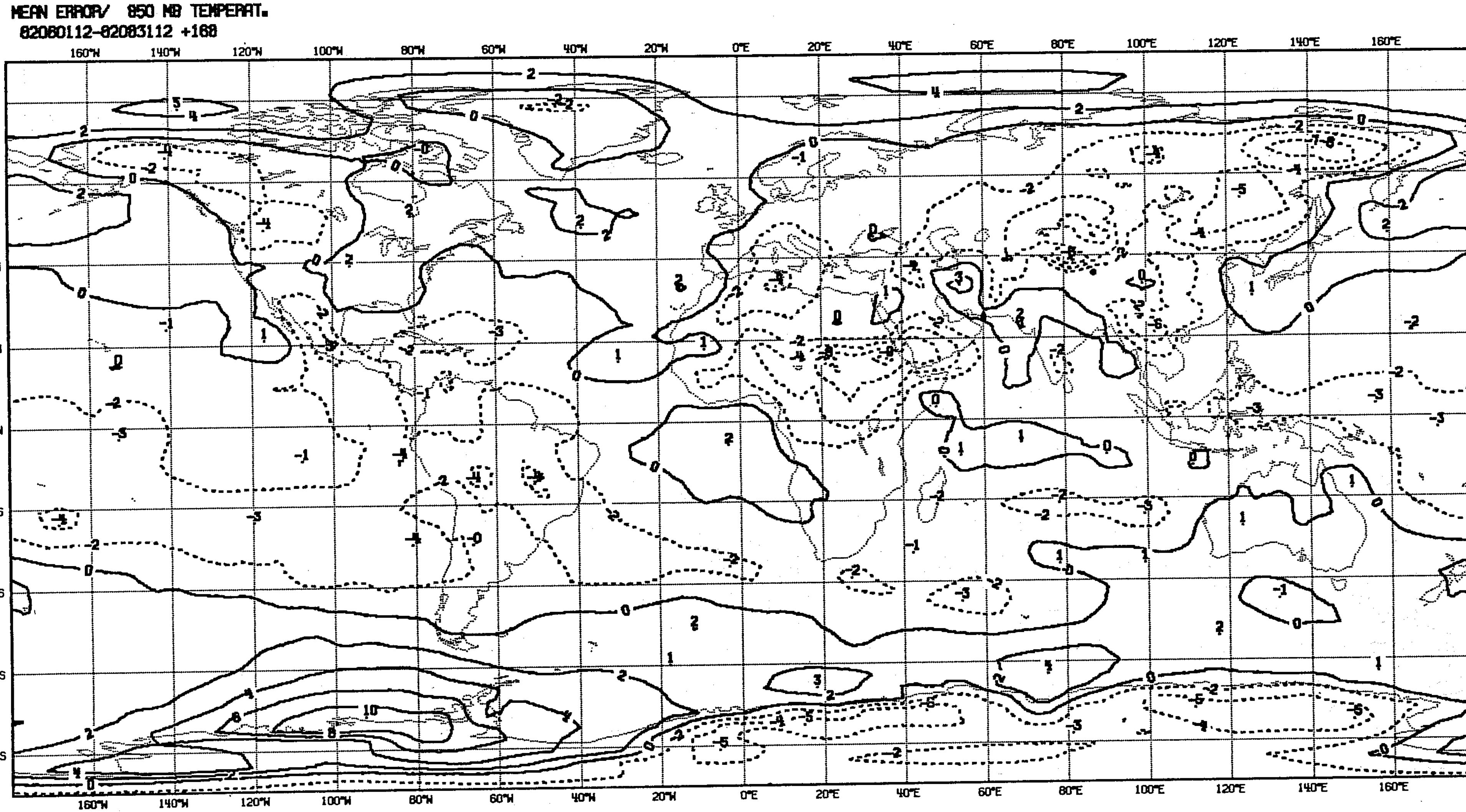

Fig. 13 Mean error (°C) for the 7-day forecast of 850 mb temperature for the three month period of June, July and August 1982.

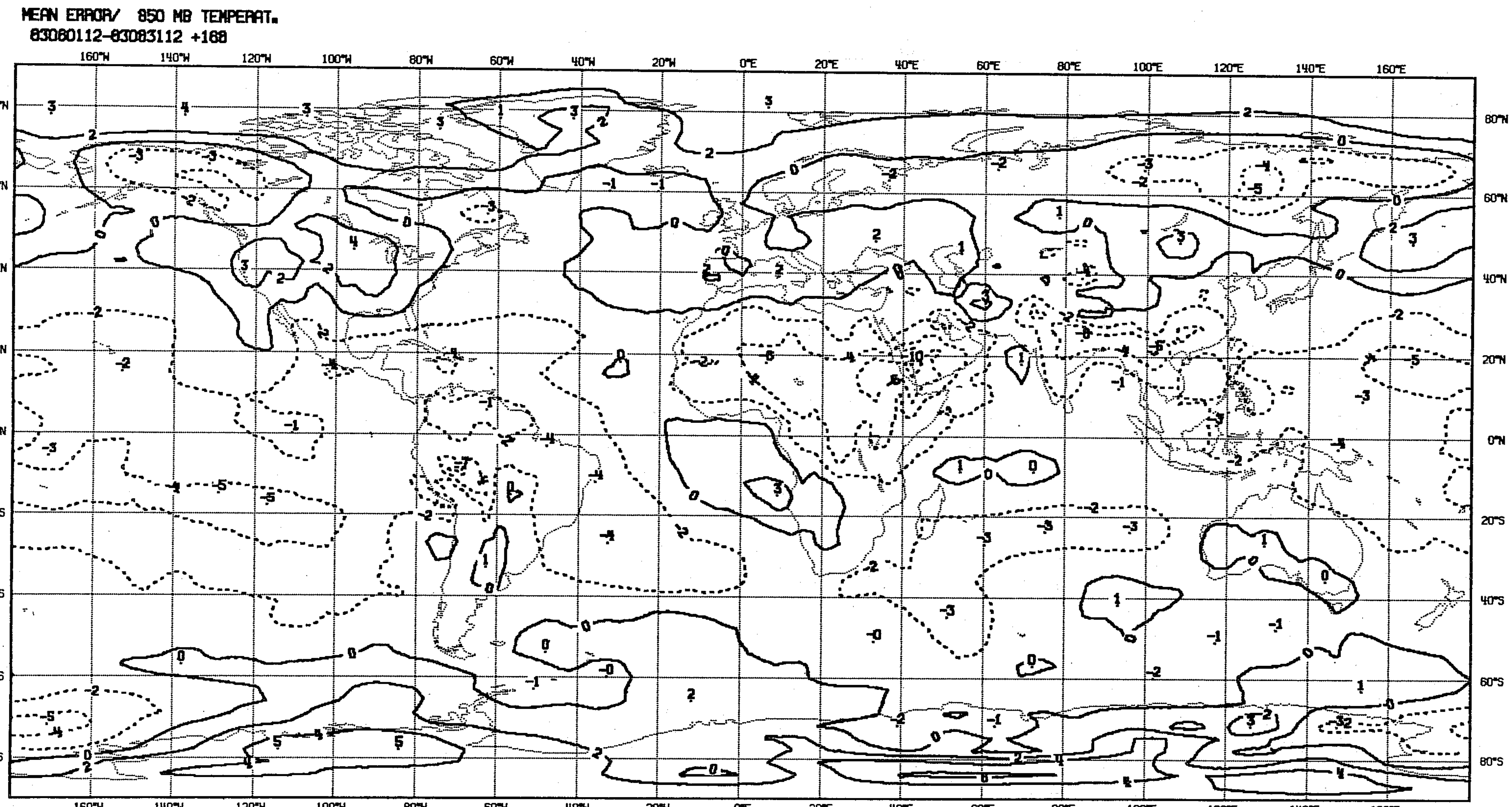

Fig. 14 Same as Fig. 13 but for 1983.

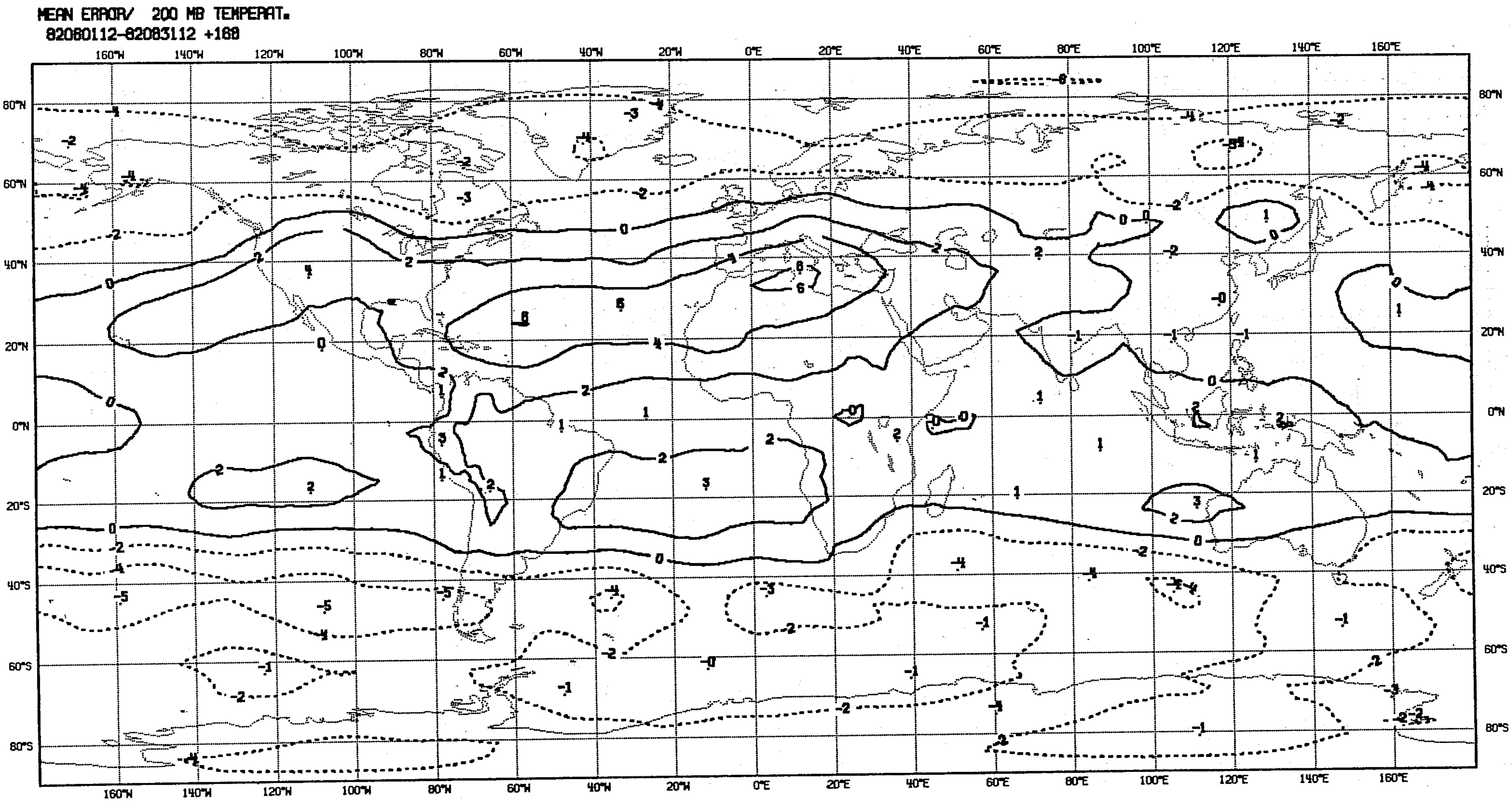

Fig. 15 Mean error (°C) for the 7-day forecast of 200 mb temperature for the three month period of June, July and August 1982.

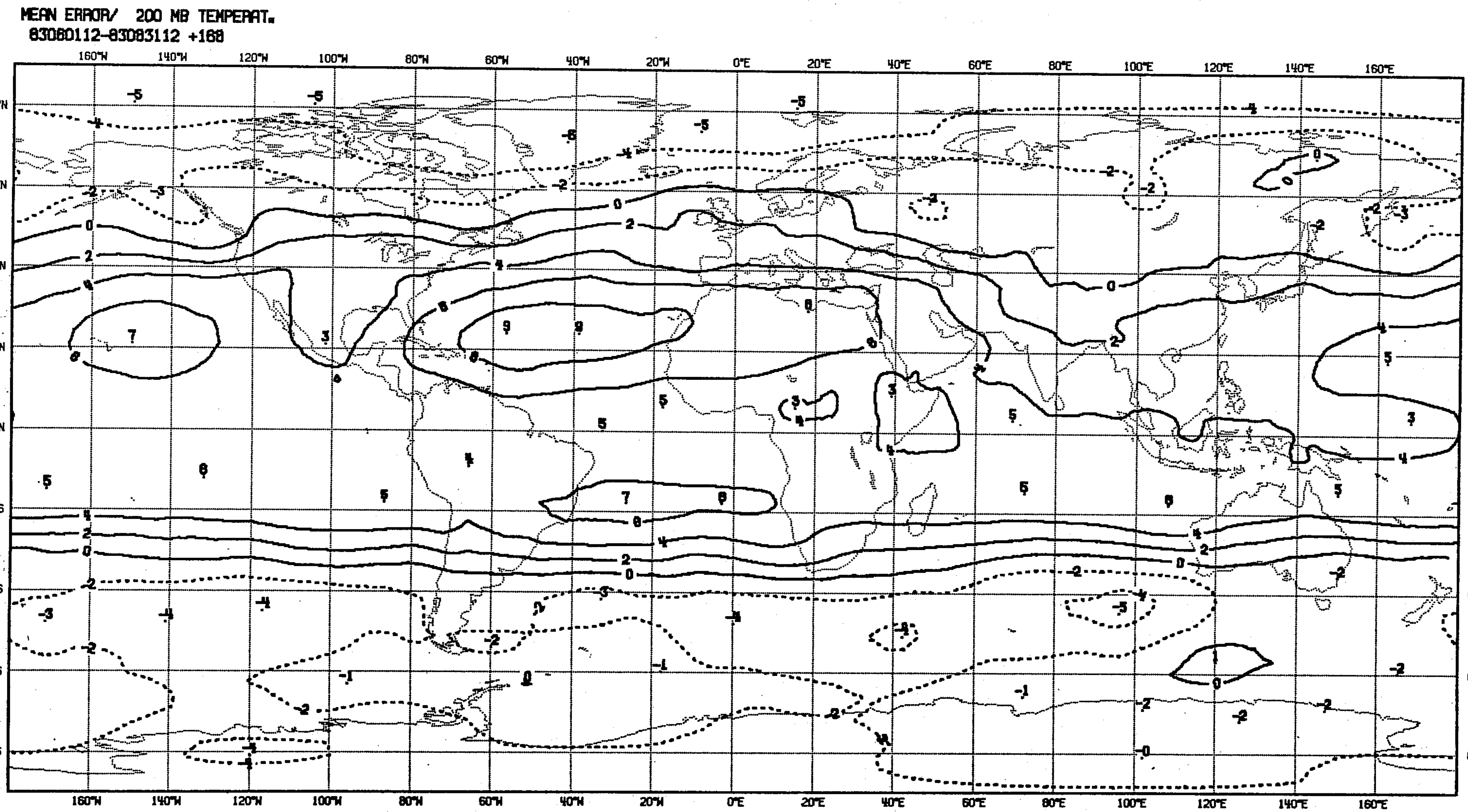

Fig. 16 Same as Fig. 15 but for 1983.

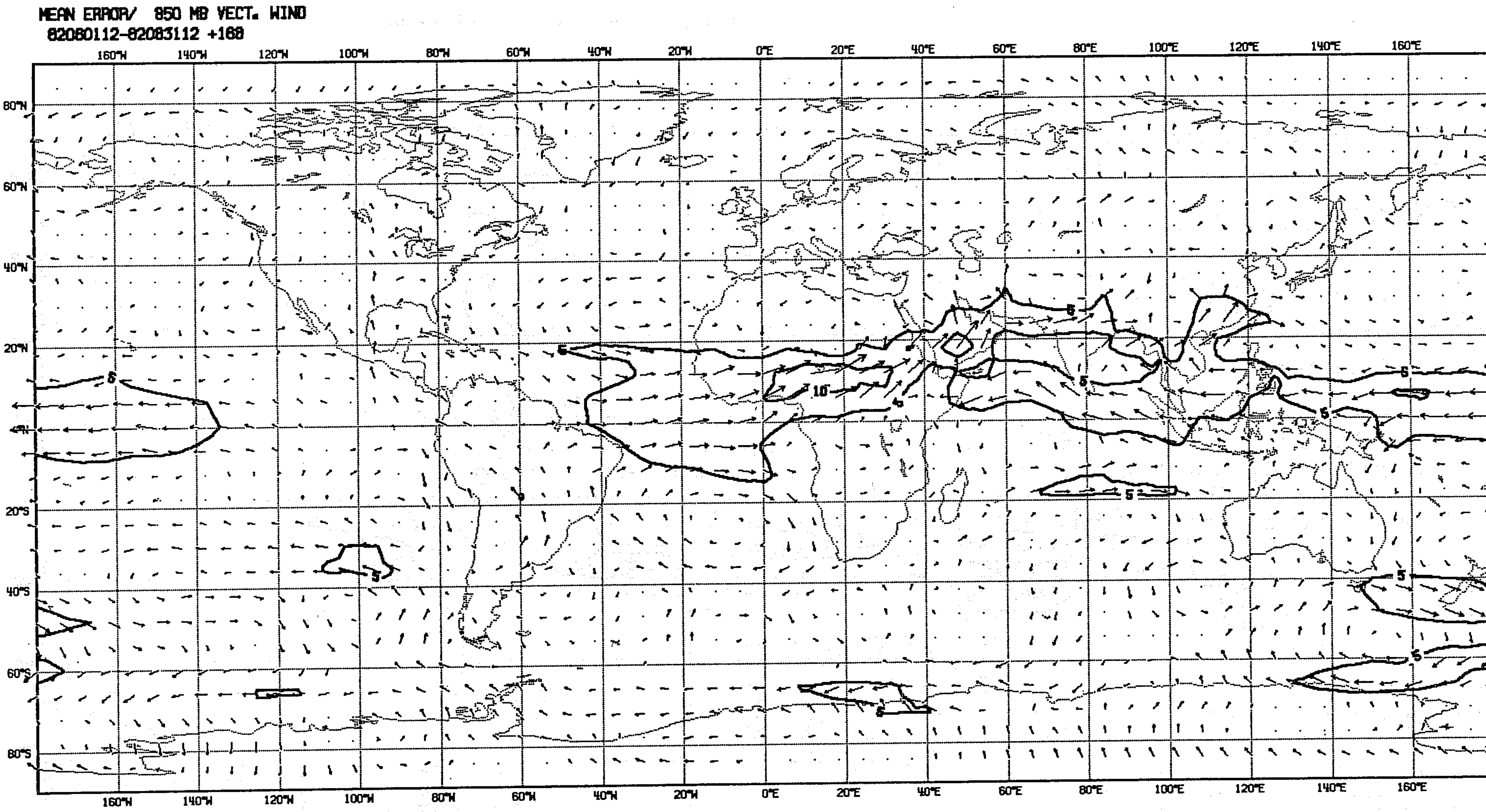

Fig. 17 Mean error (isotachs in m/s) for the 7-day forecast of 850 mb vector wind for the three month period of June, July and August 1982.

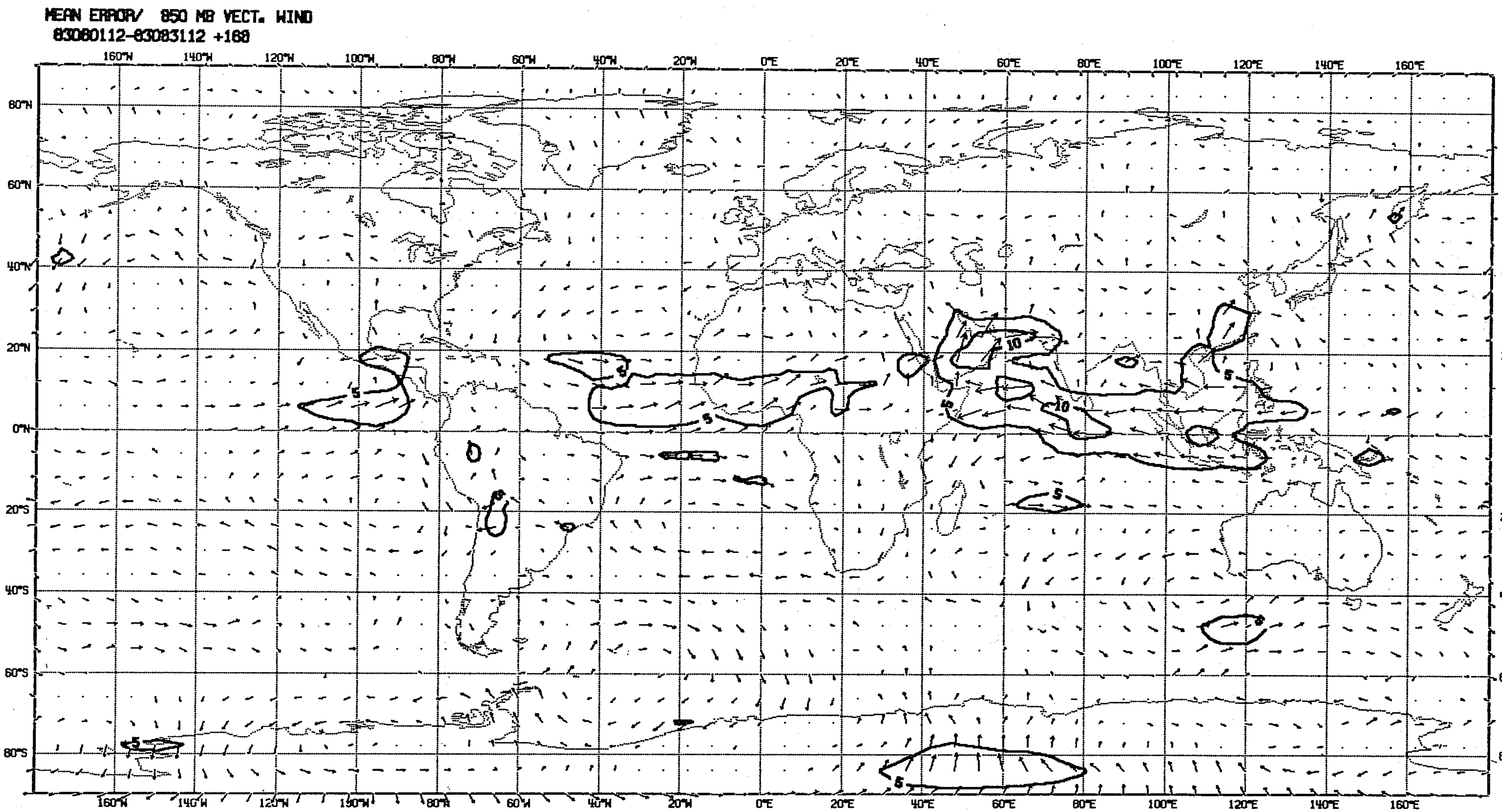

Fig. 18 Same as Fig. 17 but for 1983.

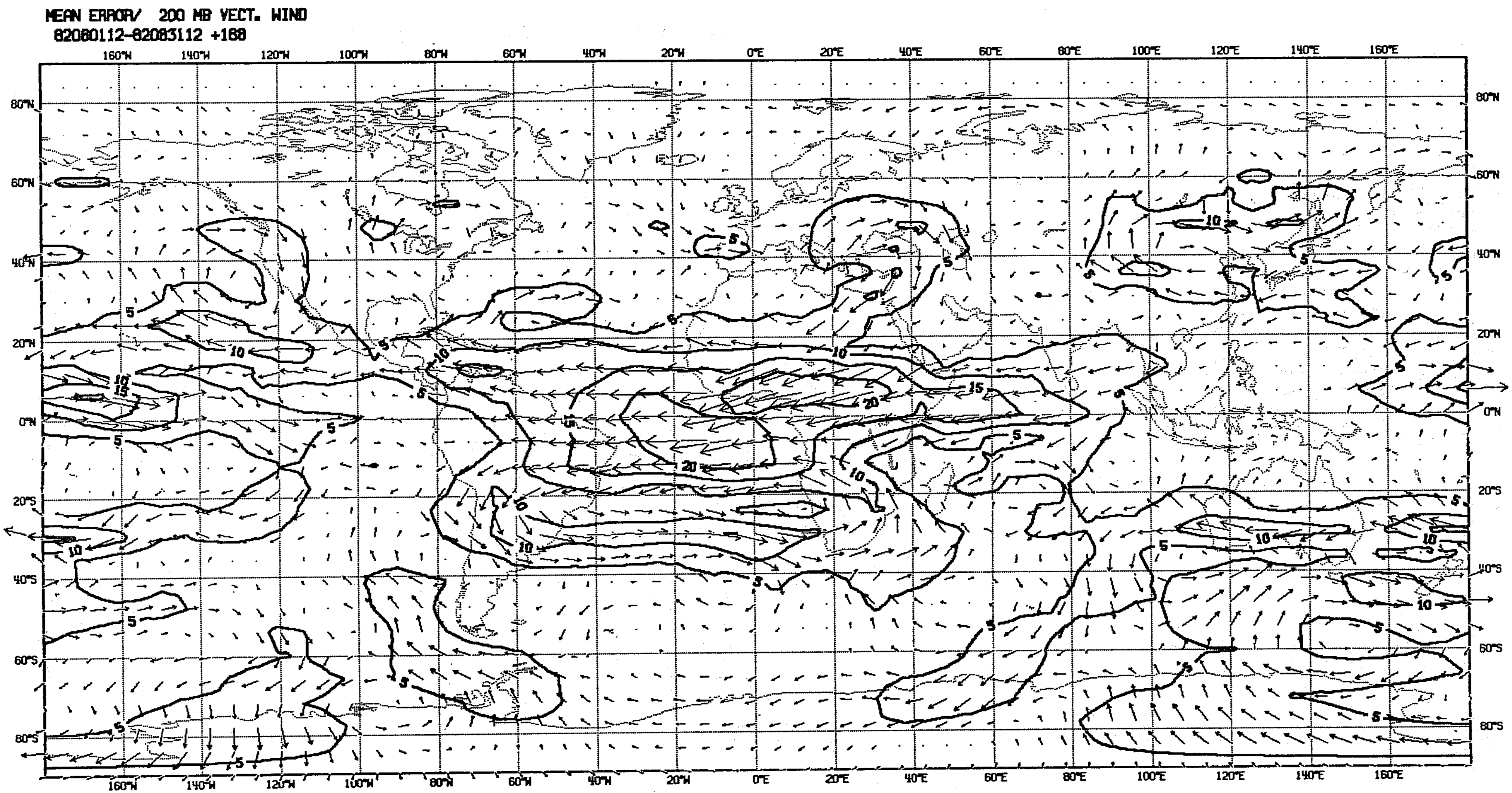

Fig. 19 Mean error (isotachs in m/s) for the 7-day forecast of 200 mb vector wind for the three month period of June, July and August 1982.

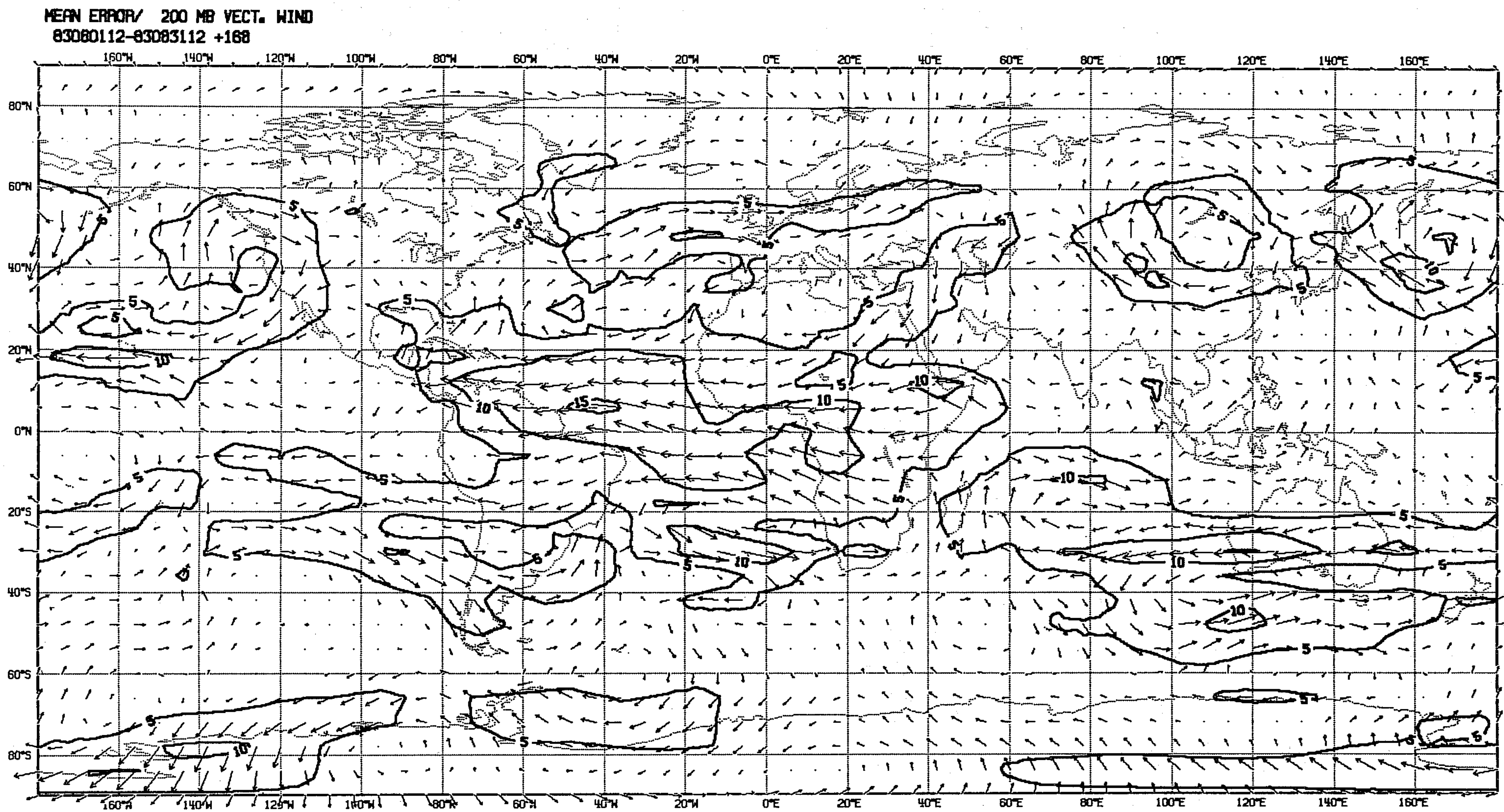

Fig. 20 Same as Fig. 19 but for 1983.

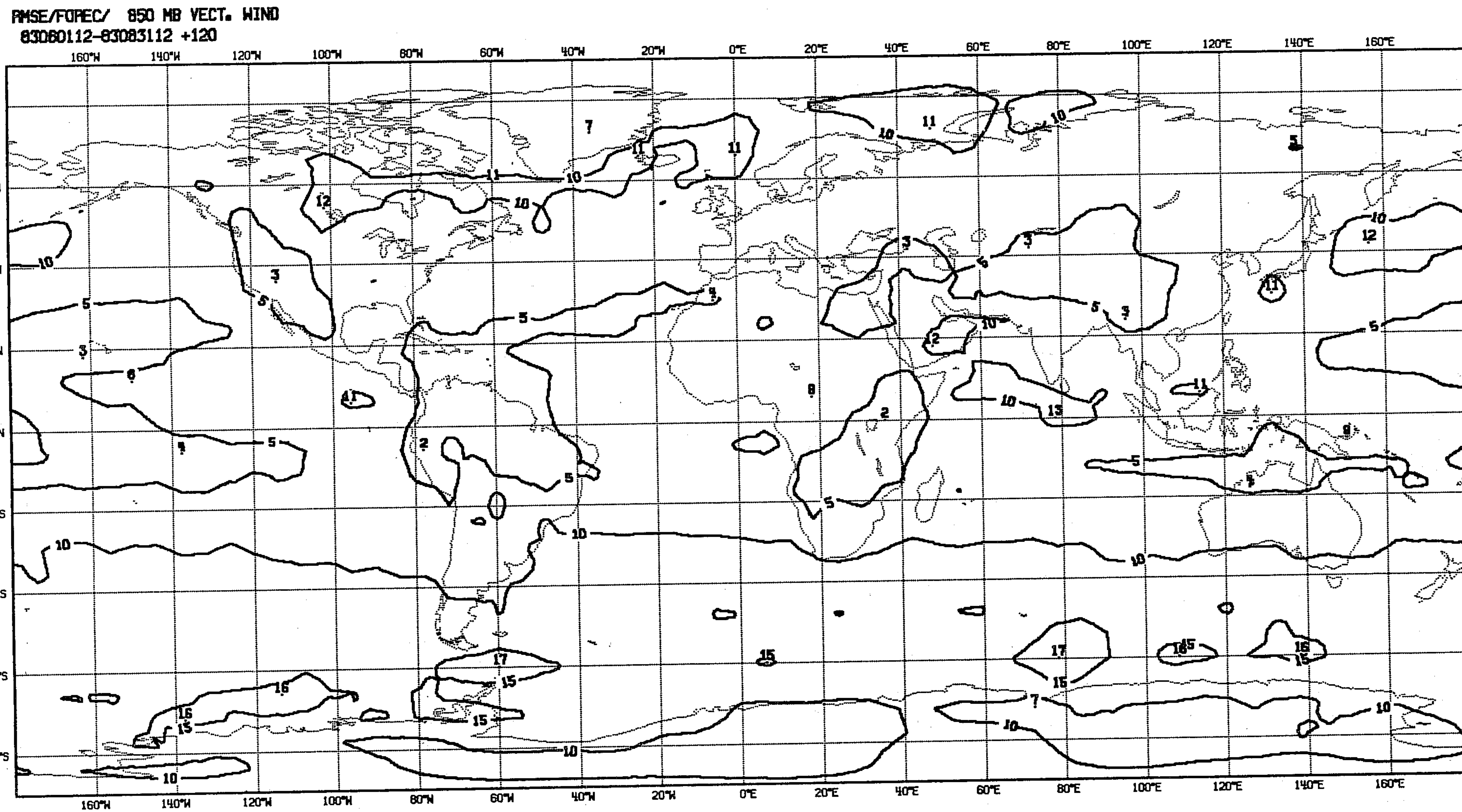

Fig. 21 Rms-error (m/s) for the 5-day forecast of 850 mb vector wind for the three month period of June, July and August 1983.

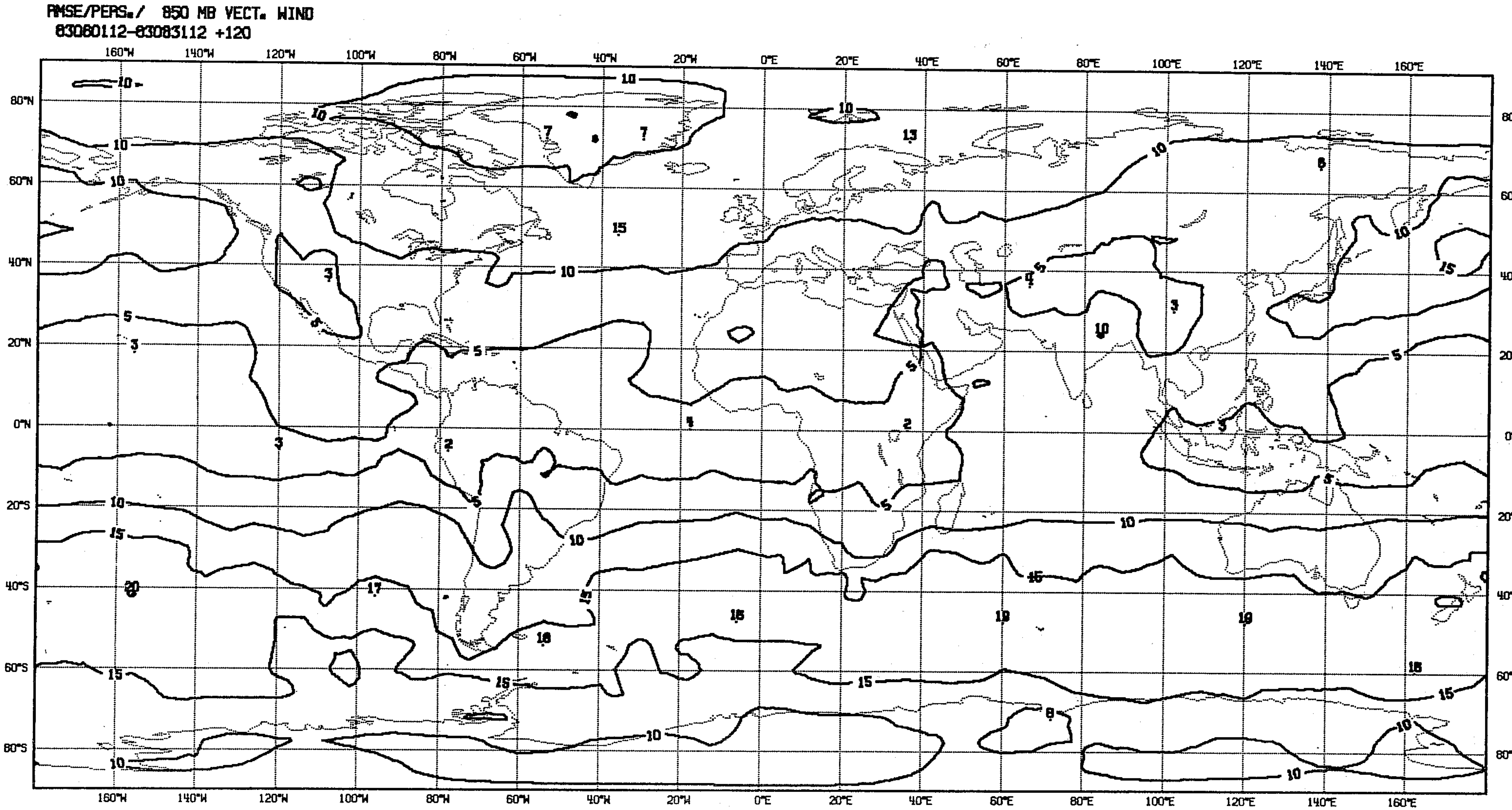

Fig. 22 Same as Fig. 21 but for persistence.

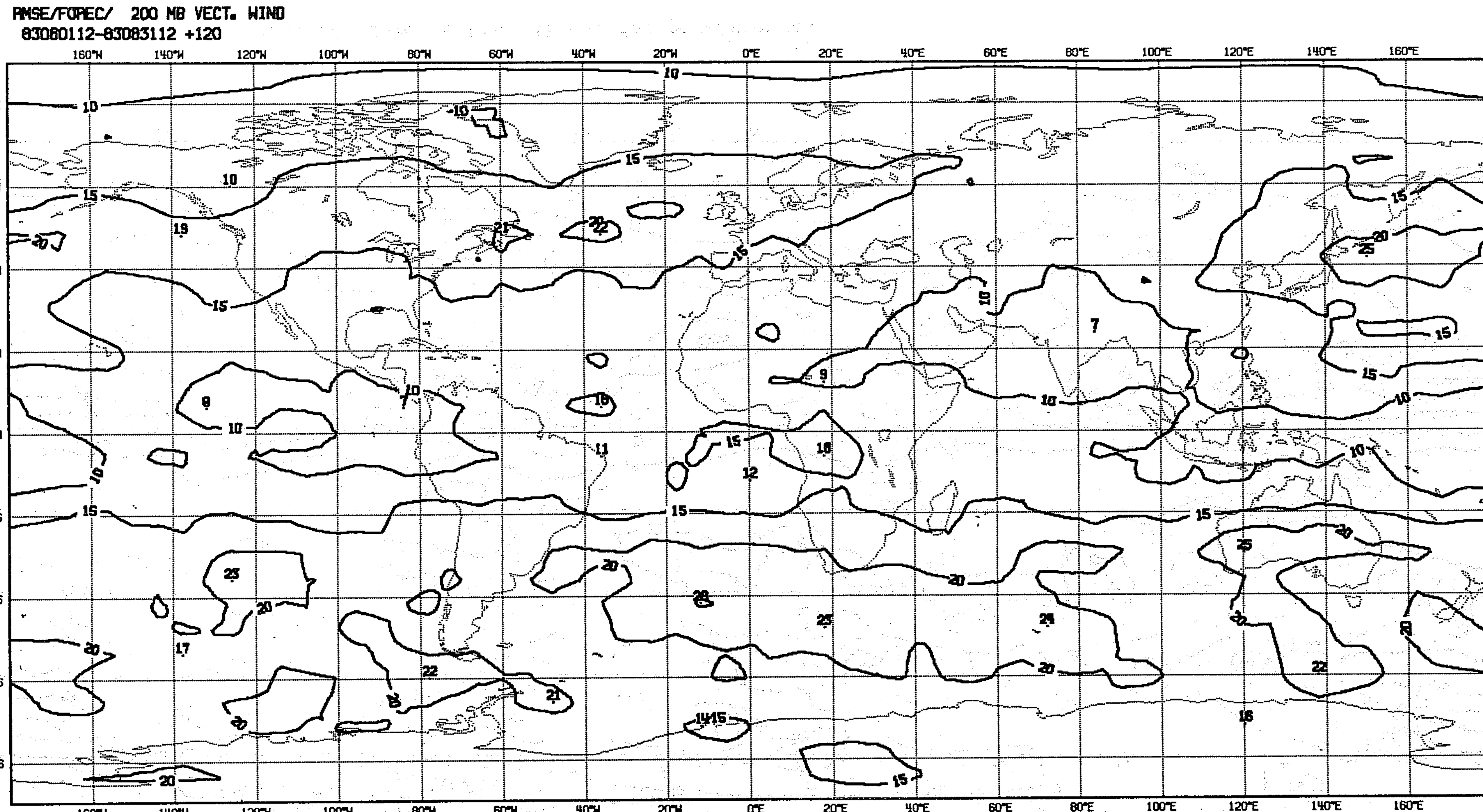

Fig. 23 Rms-error (m/s) for the 5-day forecast of 200 mb vector wind for the three month period of June, July and August 1983.

RMSE/PERS./ 200 MB VECT. WIND
83060112-83083112 +120

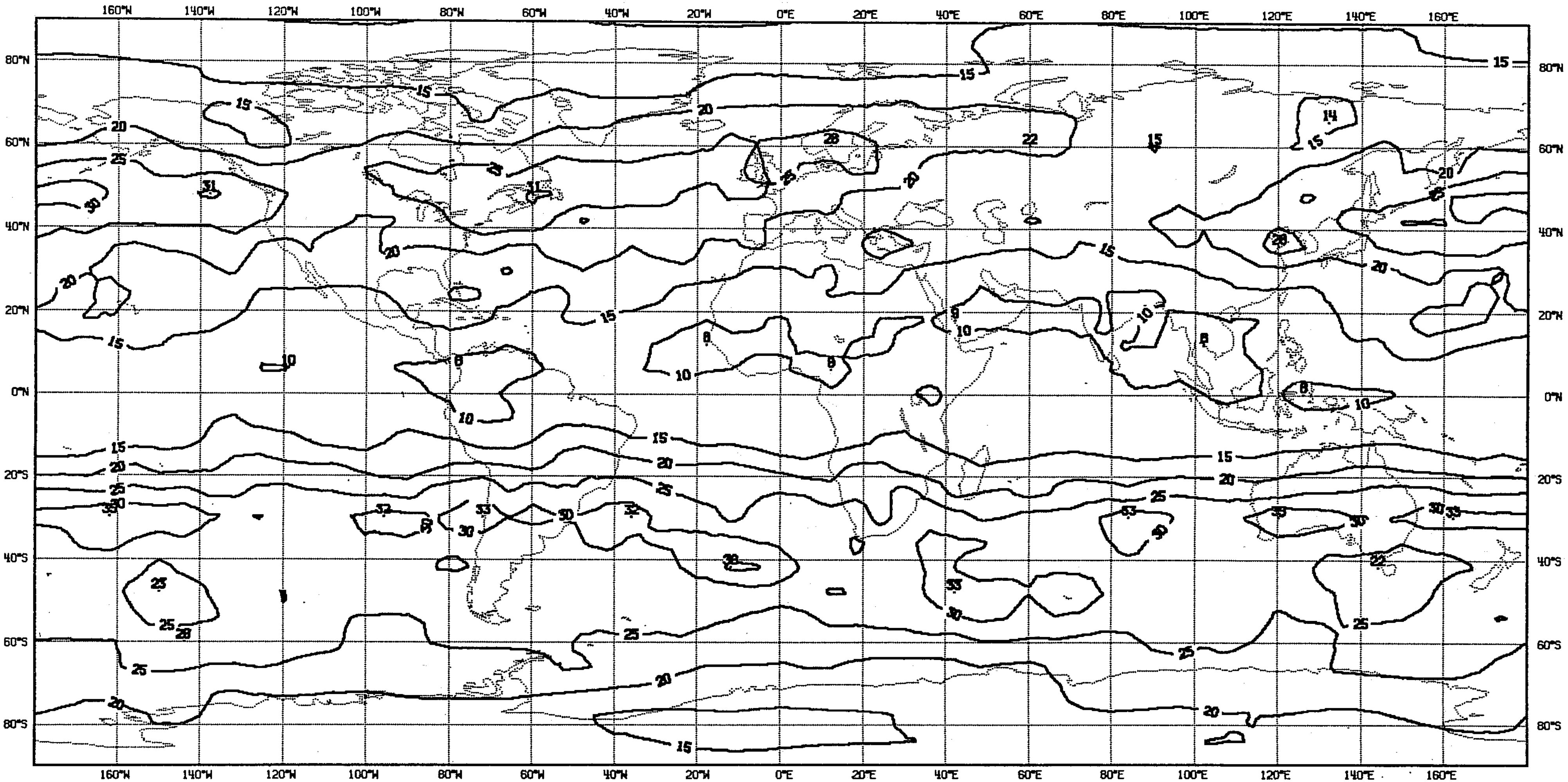

Fig. 24 Same as Fig. 23 but for persistence.

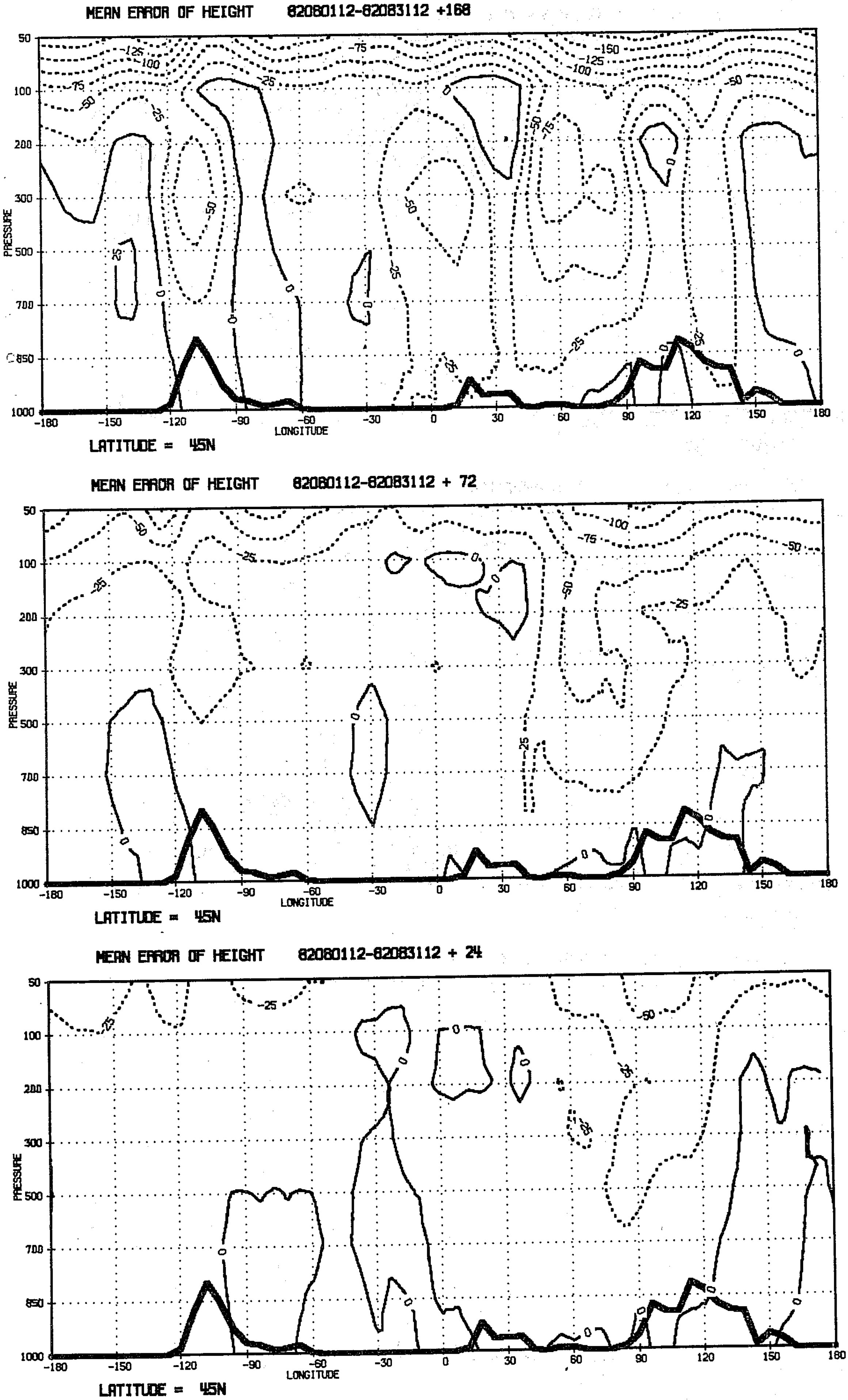

Fig. 25 Vertical cross-section of the mean height error (metres) along latitude 45 north for forecast days 1,3 and 7 for the three month period of June, July and August 1982.

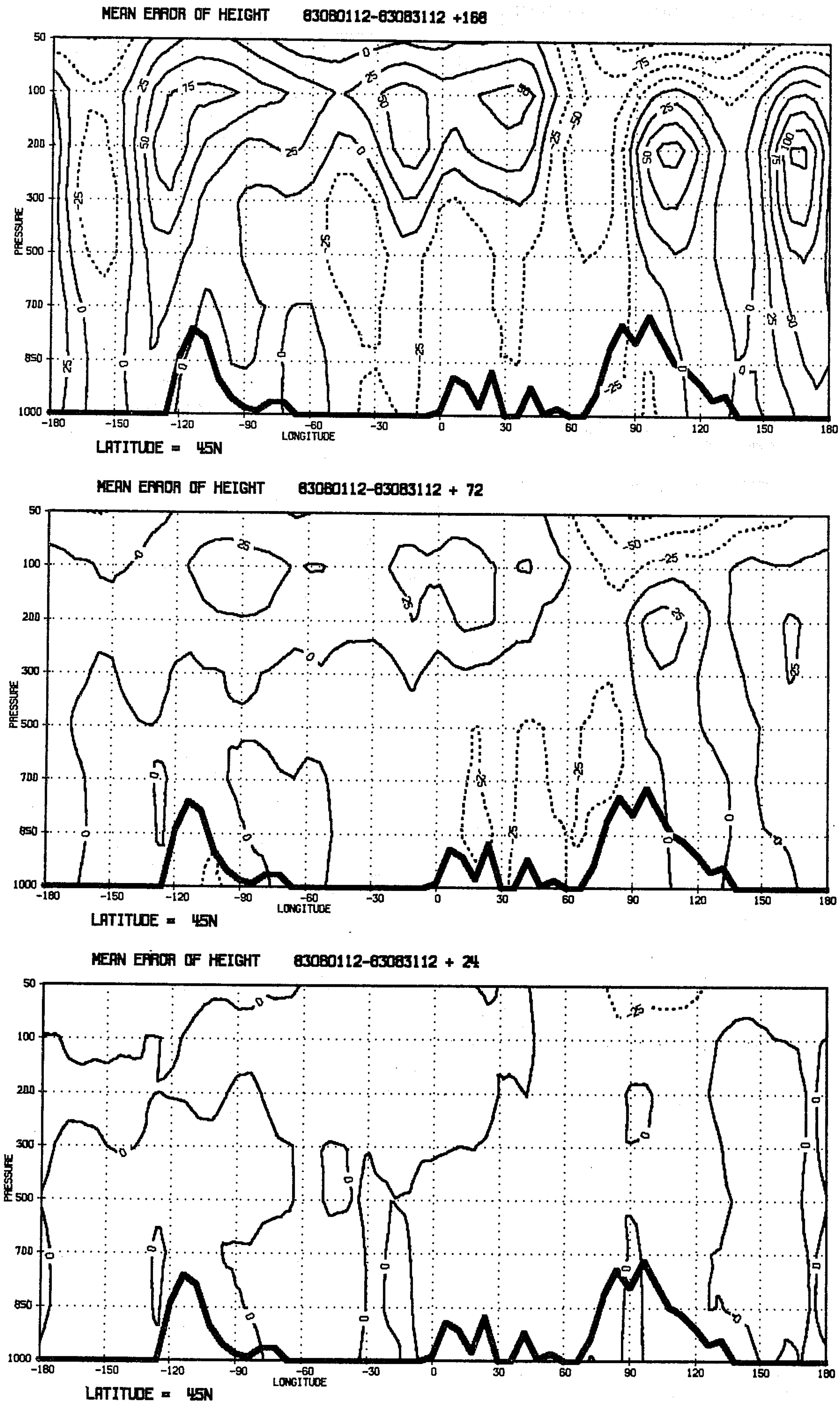

Fig. 26 Same as Fig. 25 but for 1983.

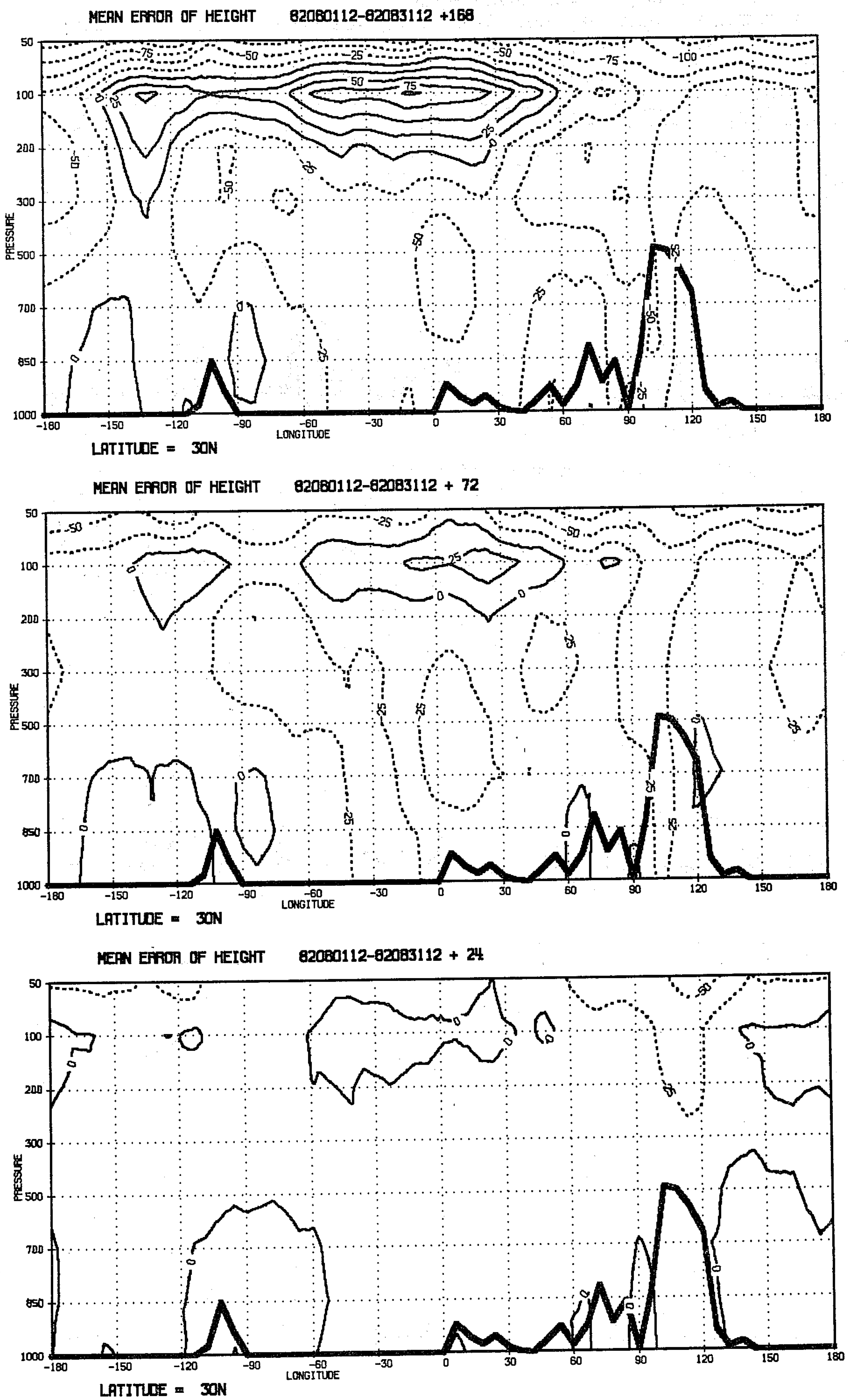

Fig. 27 Same as Fig. 25 but along latitude 30 north.

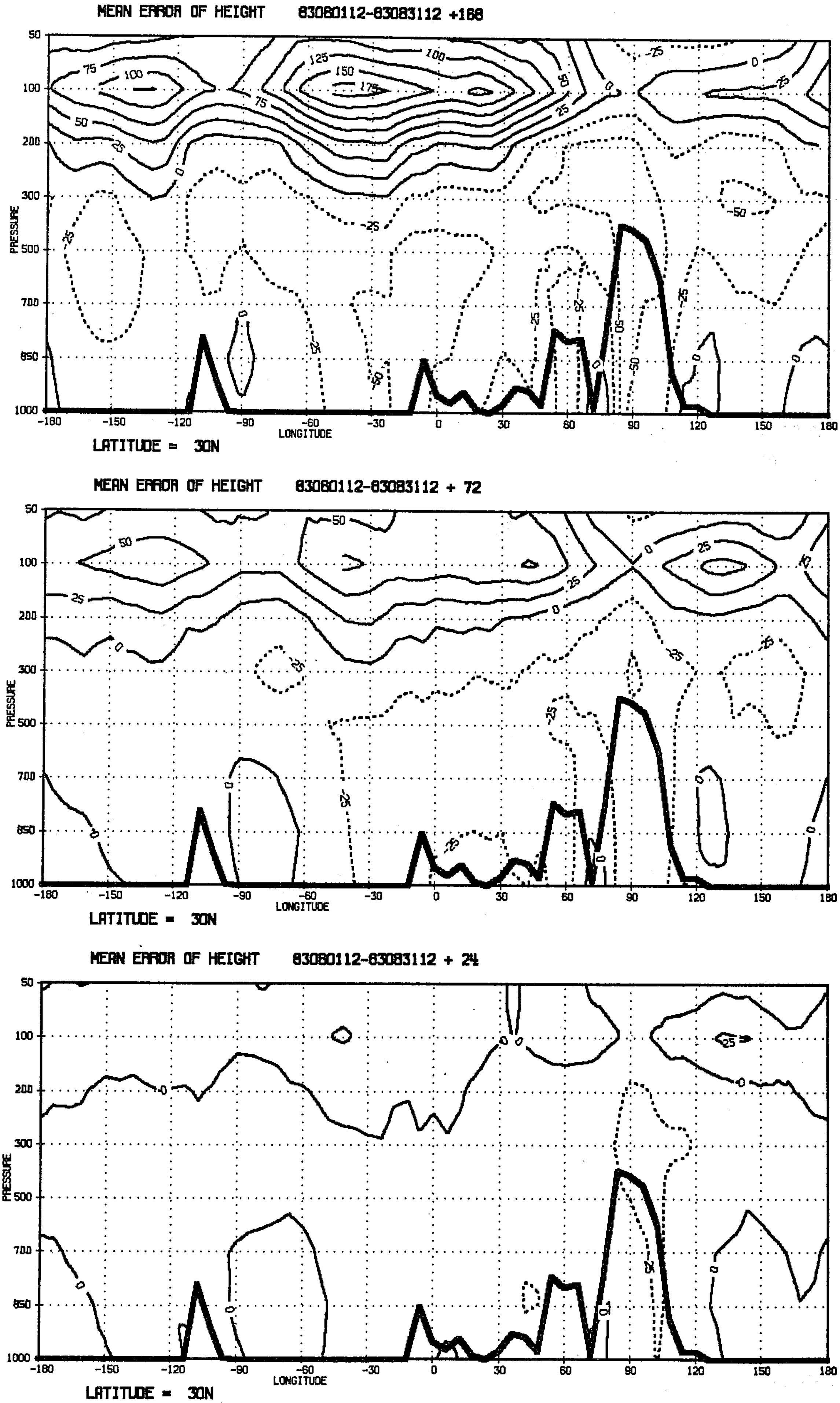

Fig. 28 Same as Fig. 27 but for 1983.

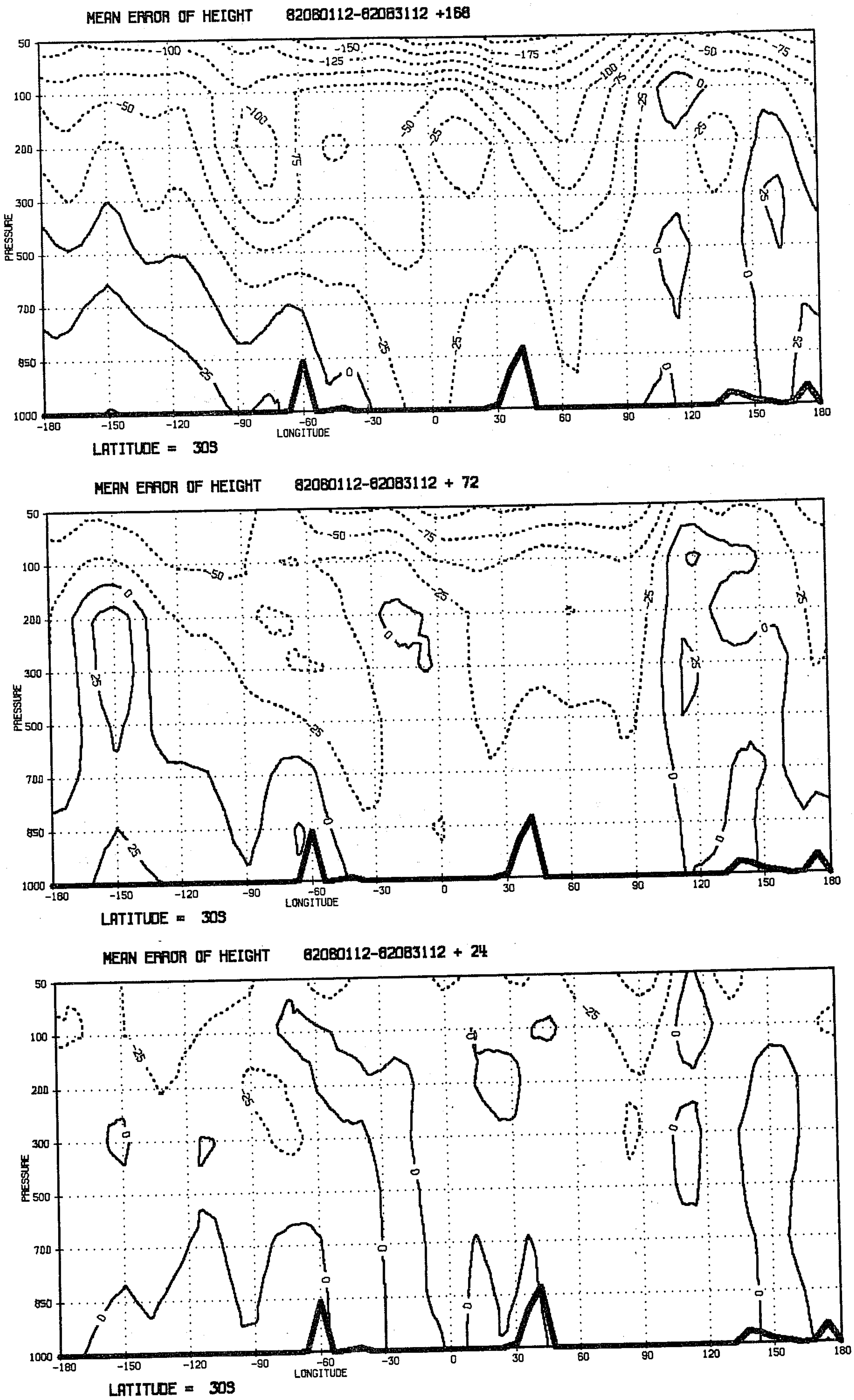

Fig. 29 Same as Fig. 25 but along latitude 30 south.

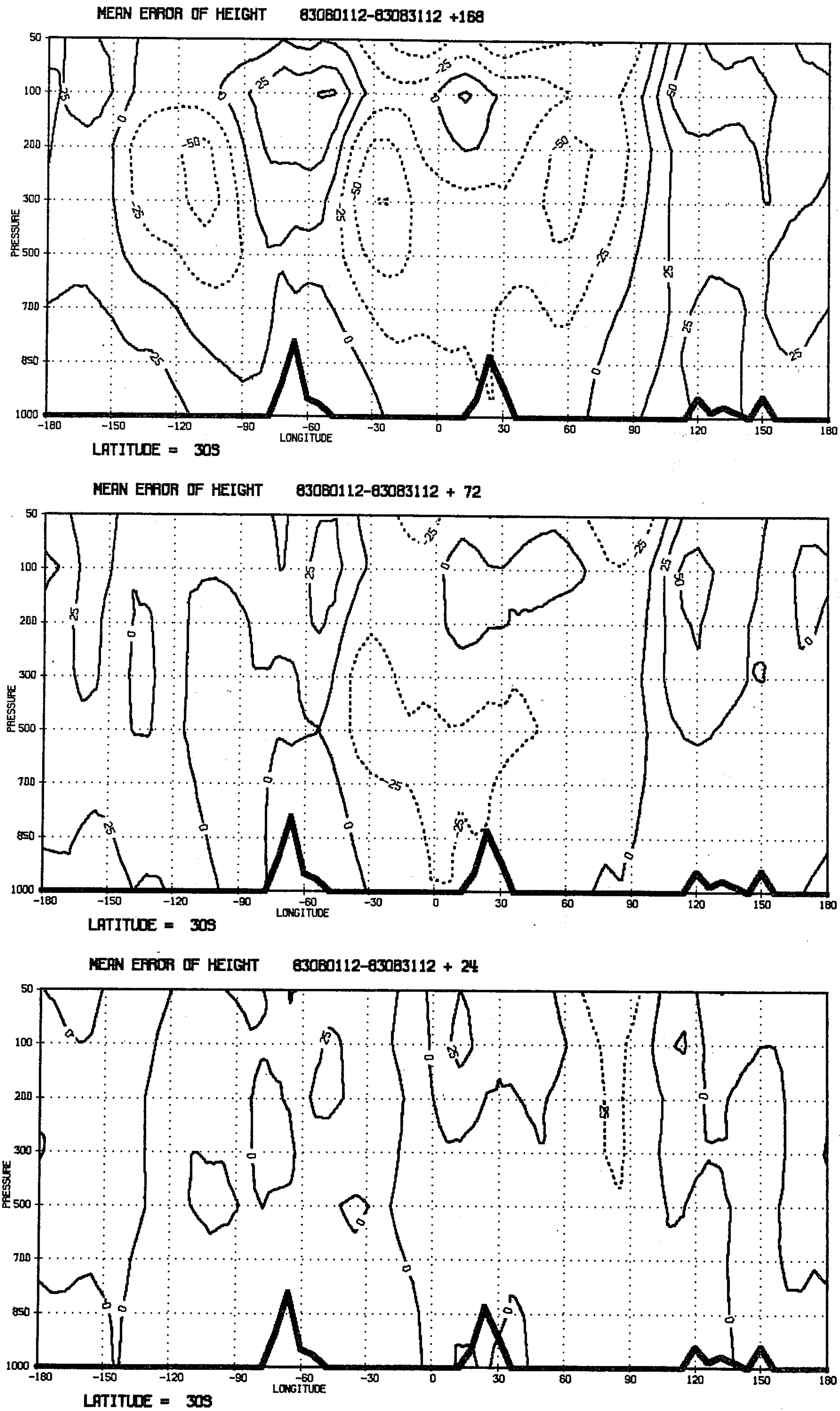

Fig. 30 Same as Fig. 29 but for 1983.