EUROPEAN CENTRE FOR MEDIUM RANGE WEATHER FORECASTS

TECHNICAL REPORT

NO. 12

****

October 1978

DATA ASSIMILATION EXPERIMENTS

BY

R. SEAMAN

EUROPEAN CENTRE FOR MEDIUM RANGE WEATHER FORECASTS

SHINFIELD PARK

READING, BERKS.

U.K.

CONTENTS PAGE NUMBER

A B S T R A C T

This report describes the first series of experiments with the ECMWF global data assimilation system. Using the Data Systems Test data base, analyses of the mass and wind fields, at a resolution of N24/9 levels, were produced at six-hourly intervals from 4 to 10 February 1976. Quantitative and/or synoptic comparisons were made with analyses from three other analysis centres. Additional experiments assessed the response of the system to the omission of satellite observations, to the omission of rawinsonde observations, and to different guess fields at the start.

1. Introduction

The purpose of this report is to describe the first series of experiments performed with the ECMWF global data assimilation system. The main objective was to assess the performance of the complete system using an observational data base of a quality approaching that expected to be available operationally. Prior to these experiments, the individual components of the system had been separately developed and tested, but their combined performance had still to be evaluated.

The scope of this report is confined to assessment of the analyses produced by the assimilation system; forecasts resulting from the analyses as initial states may be the subject of a later report. Particular attention is focussed upon, (i) comparisons with analyses from other analysis centres and, (ii) the response of the system to the inclusion or omission of specified types of observations.

The configuration of the assimilation system used in these tests was determined by the state of development of the various components at the time, and by the constraints of a CYBER 175. The system is undergoing continual improvement. A number of the shortcomings to be mentioned are expected to be ameliorated by refinements which were already well advanced at the time of this report. The results of these tests should therefore be regarded as a benchmark which is confidently expected to be surpassed in operational practice.

2. The assimilation system

The overall plan for the envisaged operational assimilation system at ECMWF has been described by Lorenc, Rutherford and Larsen (1977), henceforth denoted as TR6. It is an intermittent, forward scheme based upon a six-hour analysis-forecast cycle. Fig. 1 illustrates the inter-relation of the primary functional components and data files in one cycle of these experiments. The functions of the various components, and relevant details of their configurations, are set out below.

2.1 Analysis

The analysis method was a multivariate three-dimensional, statistical interpolation, to a regular global grid, of the deviation between observed and predicted values. Full details are contained in TR6. A 3.75 degree latitude-longitude grid, at nine pressure levels (Fig. 2), was used. The analysed variables were geo-potential, and the zonal and meridional wind components. Humidity was not analysed; the guess field provided by the prediction was unmodified.

The forecast error variances and spatial correlations, and the corresponding parameters for observational error, were specified in a similar way to that in TR6. For details, see Larsen et al (1977). The forecast error variances were allowed to vary with latitude and level, but were not allowed to vary with longitude, or from one cycle to the next. The horizontal correlation of forecast error depended upon the separation s, according to the Gaussian function $\exp(-bs^2)$, where b was a globally invariant constant of 2×10^{-12} m^{-2}. The vertical correlation matrix for forecast errors was also the same for every vertical column. These obviously severe

restrictions upon the forecast error covariance structure will certainly be relaxed in the operational configuration. The exact way in which the forecast error variances depended upon latitude, differed from TR6 to the extent that larger forecast error variances were assumed for the southern hemisphere than for the northern hemisphere at the same latitude (Fig.3). Another minor difference from TR6 was the assumption of an observational error of 7.2 m (root mean square) for a 1000 mb geopotential computed from a SYNOP (cf. 12 m in TR6). This had the effect of increasing the weight of SYNOP pressures relative to winds, and to the forecast, in the 1000 mb analysis.
A datum was rejected when either, (i) it differed from the guess field by more than five standard deviations of the prediction error, or (ii) it differed from the interpolated value (using the surrounding observations and the prediction) by more than four standard deviations of the interpolation error.

2.2 Initialisation

The purpose of this step was to prevent the growth of inertia-gravity oscillations during the subsequent prediction, by separating the initial state into Rossby and gravity modes, and setting the time derivatives of the latter to zero. For details, see Temperton and Williamson (1978). Only the first five vertical modes were included. Preliminary experiments indicated that the use of such a scheme obviated the need for heavy time smoothing during the prediction, provided that the initialisation was performed at every cycle.

2.3 Prediction

The six-hour prediction, which was the first guess for the next analysis, was provided by the adiabatic, grid point, primitive equations model of Burridge and Haseler (1977). The horizontal resolution was a staggered N24, and there were nine sigma levels in the vertical (Fig.2). The finite difference scheme conserved potential absolute enstrophy but not energy. Surface drag (C_D = .002), and dry convective adjustment were included. The linear time filter used a coefficient α of 0.005.

2.4 Grid transformations

The different coordinate systems used in the analysis and prediction steps, necessitated interpolations back and forth at each cycle. Additionally, the mass field was described in the analysis by geopotential, and in the prediction by the temperature. A hydrostatic transformation in one or the other coordinate system was therefore required.

The envisaged operational procedure will interpolate and transform analysis increments rather than complete fields (Rutherford 1977). Such a scheme had not been developed at the time of these tests, and the available programs used methods essentially the same as Rutherford's, but applied to the complete fields rather than to increments. It soon became apparent that such repeated transformations had two undesirable effects. Fristly, an excessive number of super-adiabatic lapse rates occurred in the lowest sigma layers. Secondly, the 1000 mb geopotential fields resulting from pressure reduction procedures over areas of high terrain had a very noisy appearance,with some extremely high and low values.

The first effect was traced to the use of cubic spline interpolation of geopotentials close to the earth's surface. The second effect, admittedly only a cosmetic problem, was a result of the sensitivity of pressure reduction procedures to the temperatures implied by cubic spline interpolation.

Accordingly, the interpolation methods were modified so as to use linear interpolation of geopotential between the earth's surface and the lowest pressure level above the surface. Also, the temperature at any sigma level below the midpoint of the lowest pressure layer, was computed by dry adiabatic extrapolation. Additionally, it was decided for interpolation purposes not to use geopotentials on pressure surfaces which had resulted from reduction procedures below the earth's surface. The earlier undesirable effects were overcome by the modified procedures, but as will be seen, the new methods were not without their own shortcomings.

3. The observational data base

The U.S. National Aeronautics and Space Administration provided the Data System Test 6 (DST-6), Level II data set, and also provided U.S. National Meteorological Centre (NMC) analyses based upon the Level II data. A description of the content, and a discussion of the quality of the Level II data set is given by Desmarais et al (1978). Because of processing problems, constant level balloon (TWERLE) data from the DST-6 data set were not used. Coverages typical of the major observation times (00 and 12 GMT), and the intermediate times (06 and 18 GMT), all ±3 hours, are shown in Figs. 4 and 5. (The rawinsonde coverage over Australia at 00 GMT, was much better than that shown for 12 GMT.). The data base during the First GARP Global Experiment (FGGE), and operationally thereafter, almost certainly will be

better than this DST data set. Additionally, no attempt was made in these tests to incorporate the potentially valuable information available over oceanic areas from the manual interpretation of satellite imagery. However, it is envisaged that in the future, digitised versions of southern hemisphere sea level pressure and 1000 - 500 mb thickness analyses produced manually at the Australian Bureau of Meteorology, will also be available in time for operational use.

4. Design of the experiments

The experiments consisted of a CONTROL run, and other runs differing in only one respect from the CONTROL.

The CONTROL run used the complete data base. It began at 00 GMT 4 February 1976, using as a first guess the monthly mean climatology, together with climatological covariances as in TR6. The six-hour analysis-initialisation-forecast cycle, described in section 2, was continued up to day six (00 GMT, 10 February).

The NOSAT run did not use any observational data from satellites (remote soundings or cloud winds).

The NOSONDE run did not use any rawinsonde data.

The WARMSTART run began with a guess obtained using a six-hour forecast from an NMC analysis.

In non of the runs was any manual intervention attempted. Observational data which were rejected according to the tolerances of section 2.1, simply did not affect the analysis, even if (as sometimes happened) they were thought subjectively to be correct, and vice versa.

5. Results and discussion

5.1 General remarks

The overall performance of the assimilation system in the CONTROL run was encouragingly stable. This was indicated (i) by the synoptically realistic way in which the observations were incorporated, (ii) by the small and unchanging proportion (1 to 2 per cent) of data rejected, and (iii) by the steady behaviour of integral quantities. Shortcomings of the system were mainly associated with (i) the crude specification of the critical observational and forecast error variance and covariance parameters required by the multivariate interpolation scheme, (ii) systematic biases introduced by the interpolation and transformation interfaces and (iii) lack of manual intervention to reinsert important data occasionally wrongly rejected.

The subsequent presentation will focus upon two main aspects. Firstly, comparisons will be made between the CONTROL run and analyses from other centres, and the systematic differences discussed. Secondly the CONTROL run will be compared with the other runs described in section 4, and conclusions drawn.
In the course of the presentation some of the points mentioned in the preceding paragraph will be illustrated.

5.2 Comparisons with analyses from other centres

Comparisons were made with (i) NMC's global analysis (see Desmarais et al. (1978) for details of the NMC assimilation system at that time), (ii) the Deutscher Wetterdienst (DW) northern hemisphere analyses, and (iii) the Australian Bureau of Meteorology southern

hemisphere analyses. Since the NMC analyses were known to have used an observational data base similar to the CONTROL run, and were already available on magnetic tape, quantitative comparisons of NMC versus CONTROL were performed. The analyses from all three centres were used for subjective synoptic assessments.

5.2.1 Synoptic comparisons

5.2.1.1 Northern hemisphere

The general similarities and differences between the CONTROL and NMC 1000 mb analyses, and the DW sea level pressure analyses, for the northern hemisphere can be seen from the representative example of Fig. 6. The analyses from all three sources were usually very similar, both in the location and in the intensity of synoptic systems. The CONTROL analyses were characterised by slightly more intense depressions, and a greater concentration of gradients, than were the NMC analyses. The manual DW analyses contained more detail than either the NMC or the CONTROL analyses, and the extrapolated central pressures were more extreme. The details present in the CONTROL, but not in the NMC analyses, were often, but not always, confirmed by observations. There appeared to be some spurious detail introduced by (i) the use of different sets of observations in neighbouring analysis boxes (see TR6), and (ii) in thickness patterns over areas well served by rawinsondes, due to the quasi two-dimensional nature of the geopotential analyses in these areas as a result of observation selection procedures.

Similar systematic differences between the three sets of analyses were apparent in the 1000 - 500 mb thickness patterns, particularly over dense rawinsonde areas.

In Fig. 7, CONTROL and DW show more concentrated zones of baroclinicity than does NMC. As would be expected, the foregoing differences were also reflected in a greater intensity of jet streams in CONTROL than in NMC.

The characteristic differences between the DW analyses on the one hand, and the CONTROL and NMC analyses on the other, appear to be largely the result of the deliberately chosen scales of interest; the grid lengths of the CONTROL and NMC systems would in any case be too coarse to represent some of the features in the DW manual sea level analysis. However, the characteristic differences between the CONTROL and NMC analyses appear to be associated with the different assimilation methods, and not with the different grid resolutions, since the NMC N36 grid was in fact finer than the CONTROL grid.

Although the CONTROL and NMC assimilation systems are believed to have had access to the same observational data base, it is apparent that the different quality control procedures occasionally resulted in different data actually being used at the analysis stage. Examples occurred in different parts of the hemisphere in Fig.8.The system near (72N,55W) in the CONTROL analysis, was substantiated by a SYNOP sea level pressure report at Upernavik (02210). The central pressures of the systems at (29N,180E) and (45N,157W) in the NMC analysis were substantiated by pressure reports from ships close to the depression centres. In the latter case, the CONTROL data checking procedures rejected these ships (the prediction had both centres far too weak),and perhaps the same happened in the NMC system in the former case. Also of interest is the central pressure of the system at (36N,14W) in the CONTROL. This is confirmed by a pressure and wind from a ship at (36N,12W); but it is not known whether this difference between CONTROL and NMC arose from the rejection of this report in the latter system, or

whether it is simply a reflection of a characteristically "smoother" analysis scheme.

5.2.1.2 Southern hemisphere

The amount of observational data available in the southern hemisphere for these tests, was considerably less than is anticipated during and possibly after the FGGE. In particular, only two geostationary satellites were operating and there were no drifting buoys to augment the surface network. In view of these constraints, it was encouraging that the CONTROL run remained stable without any manual intervention, and that there was some correspondence both with the NMC and the Australian analyses. As indicated in Fig. 9 (after five days of cycling in the CONTROL run), the broad scale features of the 1000 - 500 mb thickness field agreed quite well in the analyses from all three sources. However, at 1000 mb there was considerable variation between all three analyses in the detail of depression centres (e.g. Fig. 10), and the differences between CONTROL and NMC were in general greater than in the northern hemisphere. During this period, the Australian analyses at 00 GMT were based upon manually drawn charts of sea level pressure and 1000 - 500 mb thickness. The analysts incorporated their interpretations of satellite imagery, and also had access to VTPR (but not Nimbus) remote sounding data. The adverse effect of the neglect of satellite imagery on the quality of the CONTROL analyses was apparent, particularly at 1000 mb during the first two days.

The systematically different way in which the CONTROL and NMC systems handled remote temperature soundings observations, is indicated in Fig. 11.(Apologies to southern hemisphere readers for the unusual chart orientation !). It is clear

that the NMC analysis fits the absolute values of the SATEM thicknesses much more closely than does the CONTROL. This characteristic of the CONTROL analysis is quite in accordance with the properties of a statistical interpolation scheme. The amount by which the statistical interpolation method changes the first guess field towards an observed value, depends upon the (assumed) error variances of the observations and first guess, and upon the way in which these errors are spatially correlated. In this case, the error variances of the SATEM thicknesses were assumed to be of the same order as the error of the guess field, and it was also assumed that the SATEM errors were spatially correlated. A consequence of these assumptions was that the CONTROL analysis was strongly influenced by the absolute value of the thickness guess field, when SATEMS were the only observations in an area. However, the SATEMS did appear to be effective in determining the horizontal gradients of the thickness field; this is also in theoretical agreement with the way in which a statistical interpolation scheme should utilise observations with spatially correlated errors.

In those areas where one or two rawinsondes and many SATEMS were present, another systematic difference between the CONTROL and NMC was apparent. In the CONTROL analysis, with its assumed observational error covariances, the rawinsondes were fitted closely at the expense of the surrounding SATEMS. In the NMC analysis, there was a greater tendency to compromise between the two types of observation. This is evident from the two thickness analyses in the vicinity of most of the rawinsondes in Fig. 11.

Phillips (1976) has suggested that, provided the critical error covariance parameters for a statistical interpolation method are well enough known, the error in a prediction based upon an analysis by this method, should be less

than the error in a prediction based upon a "credulous" analysis method which fits the observed data closely without explicit regard to the relative errors of the observation and guess field. No such claim is made for these tests, since both the assumed observational error and prediction error parameters were subject to considerable doubt, in particular the latter which may well have been too small in the southern hemisphere. The example above is presented simply to illustrate the characteristic differences between the two analysis methods.

5.2.1.3 Tropics

Over much of the tropics, the observational data base fell very far short of that likely during FGGE and thereafter. Where there were many SATOB's and few other observations there was generally good correspondence between CONTROL and NMC wind fields at or near the observation levels (e.g. Figs. 12, 13). Again one can see the tendency of the statistical interpolation method to compromise between the guess field and the observations. (The error variances of each were of the same order but the observational errors were assumed to be random). However, in contrast with the situation for SATEMS, the CONTROL analyses appeared to fit the observed SATOBS somewhat more closely than did the NMC analyses.

There was poor correspondence between the CONTROL and NMC geopotential and thickness fields, particularly equatorward of about 15 degrees. At these latitudes, the corrections to the geopotential and wind guess fields are essentially decoupled in the CONTROL analysis. The geopotential differences appear to be associated with (i) the very small impact which remote soundings (with their assumed error characteristics) had upon the absolute

values of the thickness guess field at tropical latitudes in the CONTROL scheme, and (ii) a systematic bias in the CONTROL geopotential guess fields which, in the tropics, was of the same order of magnitude as the natural variance. The latter shortcoming will be elaborated in section 5.6.

5.2.2 Quantitative comparisons between the CONTROL and NMC analyses

Since the CONTROL analyses were performed on a 3.75 degree grid, and the NMC analyses on a 2.5 degree grid, a horizontal interpolation using cubic splines was necessary in order to compute statistics of the differences between the two sets. Except when otherwise specified, the following computations were performed on the 3.75 degree grid, the NMC analyses being interpolated. It was verified, however, that it made little difference which analysis was interpolated. The comparisons were made for the complete globe, each hemisphere, and particular smaller areas. The statistics were weighted to account for the closer spacing of grid points at higher latitudes.

Probably the most marked systematic differences revealed by quantitative comparison, were the greater wind speeds and greater variances of geopotential in the CONTROL analyses (Figs. 14a and 14b) . Such differences were evident at all levels (except 1000 mb for wind speed), and all areas irrespective of observational data coverage. Desmarais et al.(1978) noted that analyses from NMC's then experimental assimilation system, which used a multivariate statistical interpolation scheme, also had stronger winds than the NMC analyses used above. They suggested that in their tests, the stronger wind speeds obtained using the multivariate interpolation were due in part to its formulation on the prediction model sigma surfaces, and the

consequential avoidance of kinetic energy losses during back and forth transformation from pressure to sigma coordinates. It is clear that the differences between the CONTROL and NMC analyses cannot be explained in this way, since the CONTROL multivariate interpolation scheme was formulated on pressure surfaces. Possible contributory factors may be less severe time filtering during the predictions in the CONTROL scheme, or the different scale characteristics of the CONTROL multivariate interpolation scheme vis-à-vis the Hough analysis scheme used by NMC.

The root mean square (RMS) differences, between the CONTROL and NMC analyses, of the 1000 mb geopotentials and 1000 - 500 mb thicknesses over several areas (Figs. 15a and 15b), were generally consistent with the variations in the observational data base. In particular, the differences in 1000 - 500 mb thickness were of similar magnitude in each hemisphere, but the differences in the 1000 mb geopotential were significantly greater in the southern than in the northern hemisphere. This appears to reflect the uncertainties arising from the lack of surface observations in the southern hemisphere, and supports the FGGE requirement for a network of ocean buoys. The differences in 1000 mb geopotential over elevated topography were artificially inflated by differences in reduction methods. In all these comparisons, the different natural variances in the summer and winter hemispheres should be borne in mind.

Some differences in the spectral characteristics of the CONTROL and NMC analyses are indicated in Figs. 16 and 17. These were computed using the diagnostic package of Arpe et al.(1976). In this package, both sets of analyses were interpolated to an N48 grid, using cubic splines, before the spectral decompositions were performed. The differences between the analyses were relatively greatest in the middle and high wave number ranges, where the CONTROL

analyses had the larger amplitude. These results are in agreement with the qualitative conclusions drawn in the synoptic comparisons.
Only the 00 and 12 GMT analyses from NMC were available for these comparisons. However, Desmarais et al.(1978) have noted that the NMC analyses at 06 and 18 GMT lacked continuity with their 00 and 12 GMT analyses over those areas, such as North America, where rawinsondes predominated at the latter times, and SATEMS at the former times.
No such effect was evident in the CONTROL analyses, either synoptically, or from the time variation of means and variances over Europe and North America.

5.3 The effect of leaving out satellite observations

The differences between the CONTROL and NOSAT analyses reflected the geographical and temporal variations in the relative amounts of satellite-based and conventional rawinsonde observations. When and where the latter were plentiful (northern hemisphere land areas at 00 and 12 GMT), there was very little quantitative difference, and no discernable synoptic difference between the two sets of analyses. In this circumstance, the very small impact of satellite data was, in any case, ensured by the data selection scheme which gave preference to rawinsonde data over satellite data when there were many of both types.
At the other extreme, in much of the southern hemisphere, where satellite observations comprised almost the complete data base, the differences between the CONTROL and NOSAT analyses were substantial. Indeed, as was anticipated at the outset, the observational data base in the NOSAT cycle was inadequate to sustain a viable system in most of the southern hemisphere; the analyses and guess fields were synoptically unrealistic, and vital observations in data sparse areas were sometimes rejected on the basis of large deviations from poor guess fields. The problem of erroneous

rejection probably could have been avoided by specifying larger variances of the first guess error. However, the data coverages in Figs. 4 and 5 clearly indicate the major cause of the southern hemisphere result in the NOSAT cycle.

Between the two extremes of the northern hemisphere continents and southern hemisphere oceans, were oceanic areas of the northern hemisphere, where satellite observations might have been expected to add information to the relatively sparse conventional network. Here, the quantitative differences between the CONTROL and NOSAT analyses were greater than over the continents and differences in synoptic detail were also apparent (e.g. from 20 to 50 N in the mid Pacific in Fig. 18). On the other hand, even in the North Pacific, which was the area of greatest difference between the CONTROL and NOSAT analyses in the northern hemisphere, the impact of the satellite data was still only of the same order as the difference between the CONTROL and NMC analyses over the same area. The geographical variability in the quantitative effect of satellite observations, relative to the quantitative effect of the different assimilation systems, is apparent from Figs. 19a and 19b.

The northern hemisphere differences between the CONTROL and NOSAT analyses, although generally small, were nevertheless systematic in several respects. In the NOSAT analyses, the winds were stronger, and the variances of thermal and geopotential fields were greater, than in the CONTROL (Figs. 20,21a,21b). The differences in variance were relatively greatest in the higher wave numbers. These results are consistent with the tendency which has been noted by several authors, for remote soundings (retrieved by the method in use during 1976) to underestimate the true atmospheric variance. This characteristic of SATEM thicknesses is recognised in the statistical interpolation, to the extent that they are assigned a spatially correlated

observational error. But, assuming that the assigned error characteristics are correct, this simply means that the SATEMS are weighted so as to minimise the mean square analysis error. The final analysis is still a compromise between the observations and the first guess, and will therefore reflect any systematic oversmoothness implied by the observations.

5.4 The effect of leaving out rawinsonde observations

The results of the NOSONDE experiment paralleled those of the NOSAT experiment, to the extent that the differences between the CONTROL and NOSONDE analyses closely reflected the space and time variations in the observational data bases available to each cycle. But unlike the NOSAT cycle, in which the absence of satellite data resulted in a non-viable assimilation system over much of the southern hemisphere, the NOSONDE cycle remained stable, and the positions of the larger scale synoptic features agreed fairly well with the CONTROL cycle in both hemispheres (e.g. Fig.22). The largest differences, both quantitatively and synoptically, were over the continents of the northern hemisphere. Here the differences in intensity were often substantial, there were significant differences in thermal gradients, and the correspondence of smaller scale features was poor (e.g. Fig.23). At the other extreme, over much of the southern hemisphere (where there were few rawinsonde observations to leave out) the CONTROL and NOSONDE analyses agreed closely. However, in the Australia-New Zealand region, the differences between the two sets of analyses were similar to those observed in the northern hemisphere. The geographical variability in the quantitative effect of rawinsonde observations, relative to the quantitative effect of the different assimilation systems in the CONTROL and NMC cycles, is apparent from Figs.24a and 24b. In some respects these figures are "mirror images"

of Figs. 19a and 19b.

The systematic differences between the CONTROL and NOSONDE analyses were consistent with corresppnding results in 5.3. In the NOSONDE analyses, the wind speeds (Fig.25), and the variances of the thermal and geopotential fields (Figs. 26a and 26b), were generally lower than in the CONTROL. A spectral analysis of the variances indicated that the lesser available potential energy in the NOSONDE analyses occurred mainly in the middle and high wave numbers, but this was not the case with the geostrophic kinetic energy. An additional feature highlighted by this set of comparisons, was a strong positive bias in the NOSONDE 1000 - 500 mb thicknesses, over Europe and North America, relative to the CONTROL thicknesses.

5.5 The effect of using a better guess field at the start

The use of a climatological guess at the start, is open to criticism on the grounds that a practical, and better quality, guess field would be a short range forecast from a system already in operation. In order to test the sensitivity to the starting-up procedure, the NMC analysis at 00 GMT 4 February was interpolated to the ECMWF prediction model grid, and a six-hour forecast performed to provide the first guess for the analysis at 06 GMT. Thenceforth, procedures were identical with the CONTROL run. In this experiment (the WARMSTART run), the atmospheric state was therefore identical with NMC's at time zero; an objective was to determine how rapidly the WARMSTART analyses diverged from the NMC analyses, and (possibly) converged towards the CONTROL.

As indicated by Figs. 27a and 27b, within 36 hours (six cycles), the WARMSTART analyses had become very little different from the CONTROL analyses in the northern hemisphere. The result was qualitatively similar in the

southern hemisphere, but for the 1000 mb geopotential the process took longer than it did in the northern hemisphere. This appears to be consistent with the relative amounts of thickness and reference level data in each hemisphere. A synoptic benefit of a better starting guess was evident in the evolution of a depression to the southwest of Australia, the existence of which was independently confirmed by satellite imagery. Its development and movement during days 0 to 2 was captured better in the WARMSTART run, than it was in the CONTROL.

Despite the rather short lived effect upon the subsequent assimilated states, there were nevertheless significant differences between analyses using respectively a forecast and climatology for a first guess, even in relatively data-dense areas. The differences in these areas appear to be due primarily to the characteristically different shapes of the climatological and forecast error spatial correlation functions. For example, it was noted in TR6 that, using a climatological guess, the 1000 mb analysis at 00 GMT on 10 February failed to represent adequately a cut-off depression off southeast Greenland. So did the NMC analysis for the same time. However, the shortcoming was considerably ameliorated in the CONTROL run, using a forecast guess (Fig. 28). The generally more detailed character of the resultant analysis when using a forecast guess and forecast error covariances was also apparent.

5.6 The bias problem

The methods for horizontal and vertical interpolations and hydrostatic transformations between the coordinate systems and variables of the analysis scheme and the prediction model, have been mentioned in section 2.4. It is obviously desirable that such back-and-forth transformations should introduce as small a change as possible and should

not alter the fields in any systematic way. The procedures used in these experiments fell short of this objective, to the extent that a back-and-forth transformation introduced systematic changes, as well as random errors.

The typical systematic biases (Table 1), were two-fold. Firstly, the wind speedsat all levels were reduced. Secondly, the global mean geopotentials at 1000 mb and 850 mb were lowered by about 10 m and 5 m respectively, with much smaller mean changes at other levels (except 100 mb). The first effect is consistent with the tendency of cubic spline interpolations to truncate the peaks of the vertical wind profile. This should probably be ameliorated by the interpolation of increments instead of complete fields. The second bias is partly fictitious, since a geographical breakdown of the Table 1 statistics indicated that the larger biases at 1000 mb and 850 mb occurred over elevated topography. But a smaller bias of the same sign was evident over areas unaffected by pressure reduction methods.

In the statistical interpolation scheme, a basic assumption is that the population mean of the guess field error is zero. A bias in the guess field ensures a bias in the analysis of the same sign. Fig. 29 strongly suggests that this was the case in these experiments. The mean of the CONTROL 1000 mb geopotential guess field was consistently lower than that of the analysis, and the mean of the CONTROL analysis was itself consistently lower than that of the NMC 1000 mb analysis.

The bias introduced by grid transformations was at its most harmful in the tropics, where it was of the same order as the six-hour variability. In view of this, and the dynamical implications for the prediction model, probably not much meaningful information can be drawn from the tropics in these experiments.

At the time these biases were discovered, it had already been decided that the operational assimilation cycle would transform analysis increments rather than complete fields. Therefore, no time was spent investigating in detail the causes of the deficiencies in the existing programs. The experience nevertheless indicated the need to ensure that the new algorithms do not have similar defects.

6. Concluding remarks

The preceding experiments were the first attempt to evaluate the combined performance of the various components of the ECMWF data assimilation system. As only to be expected with a system in its infancy, a number of shortcomings were apparent. Work is already underway to rectify these.

Perhaps the greatest scope for improvement is in the specification of the observational and forecast error covariance parameters, which determine the relative contributions of the observations and the prediction to the assimilated state. The extent to which the current crude specifications may be refined and updated in practice, may well determine the cost-effectiveness of the multivariate statistical interpolation method over simpler but computationally less expensive methods.

The decision not to manually intervene in the assimilation cycles resulted in the rejection of some correct and synoptically important observations. While this loss detracted from the quality of individual analyses, the only experiment in which erroneous rejections may have contributed to progressive deterioration of analyses was the NOSAT experiment in the southern hemisphere, when a viable automatic system could hardly be expected in any case.

The other result of no manual intervention, namely the neglect of potentially useful information from satellite imagery interpretation, also had an adverse effect on the quality of southern hemisphere analyses, at least at the surface. But again, in the CONTROL run the effect was not a progressive one, and the system was stable enough to suggest that, with foreseen improvements in the surface observational data base, an automatic system may be viable in this hemisphere.
Nevertheless, the decision whether or not to intervene manually is primarily one of cost-effectiveness. If the aim is to produce consistently high quality analyses for historical and research purposes, a facility for manual intervention is probably essential.
Once it has been decided to use different vertical coordinate systems and different mass field parameters in the analysis and prediction, some error resulting from back-and-forth transformations is unavoidable. Nevertheless, a lesson of these experiment is that it is vital that such errors should not be systematically cumulative.

Despite the shortcomings enumerated above, the quality of the analyses produced even with this low resolution, physically unsophisticated system give much cause for optimism. In addition to the anticipated improvements in the observational data base, particularly in the tropics and the southern hemisphere, the computing power of the CRAY will permit at least a doubling of the global spatial resolution. It will also allow the introduction of physical parameterizations which have already yielded good results (Hollingsworth et al.,1978). The latter two developments can only result both in better guess fields from the prediction part of the cycle, and a greater ability to capitalise upon the resolution of available observational data.

Because of the known limitations of the assimilation system, the results of the observational data impact tests should be interpreted with caution. The following conclusions appear to be justified.

(i) The differences between the CONTROL and NOSAT analyses in the northern hemisphere were only of the same order as the differences between the CONTROL and NMC analyses, even in those areas with lesser amounts of non-satellite data. This result, together with intercomparisons of analyses from different centres reported by Desmarais et al. (1978, Ch.6), possibly says more about the "state of the art" of data assimilation, than it does about the potential benefit of satellite data. It appears that the effect of satellite data on northern hemisphere analyses is less than or close to the uncertainty in analysis, as indicated by the differences between analyses from different centres using all available data. It is therefore not surprising that the impact of satellite data on <u>forecasts</u> may be inconclusive and system dependent.

(ii) The systematic differences in the gross characteristics of the CONTROL, NOSAT and NOSONDE analyses underline the inherently different properties of satellite soundings and radiosonde observations, and the need to take account of these differences in an assimilation scheme.

(iii) The stable performance of the CONTROL run, as opposed to the NOSAT run, without manual intervention in the southern hemisphere, emphasises the vital rôle of satellite temperature soundings in this area. However, the uncertainties reflected by the differences between the CONTROL, NMC and Australian 1000 mb geopotential analyses in the southern hemisphere, indicate how essential it will be for the success of the FGGE to supplement the surface network by drifting buoys.

Acknowledgements

I gratefully acknowledge the willing advice and assistance from the experts on the various components of the assimilation cycle, namely D. Burridge, C. Clarke, R. Gibson, C. Little, A. Lorenc and C. Temperton. I thank A. Lorenc for his advice during the initial setting-up of the experiments, K. Arpe for the use of his diagnostic package and L. Bengtsson for his enthusiastic encouragement and guidance throughout.

Pressure level (mb)	1000	850	700	500	400	300	250	200	150	100
Mean square speed before $(m/s)^2$	62.6	116.9	153.7	329.9	491.8	696.2	741.7	730.5	641.0	396.6
Mean square speed after	47.3	105.0	147.4	317.7	478.5	664.2	703.4	704.9	634.4	389.3
Change in mean geopotential (m)	-10.0	-4.7	-1.3	+1.1	+0.6	-0.1	+1.6	+2.2	-0.2	+16.0
RMS difference in geopotential (m)	17.4	7.9	4.8	2.9	1.7	6.5	3.5	12.9	10.1	29.0
RMS difference in U-component (m/s)	2.8	1.2	0.8	0.8	1.4	1.4	2.5	1.4	3.0	5.0
RMS difference in V-component	2.8	1.1	1.0	1.1	1.5	1.8	2.6	1.7	3.1	4.5

Table 1 The effect of a back-and-forth transformation between the coordinate systems and variables of the analysis and prediction.

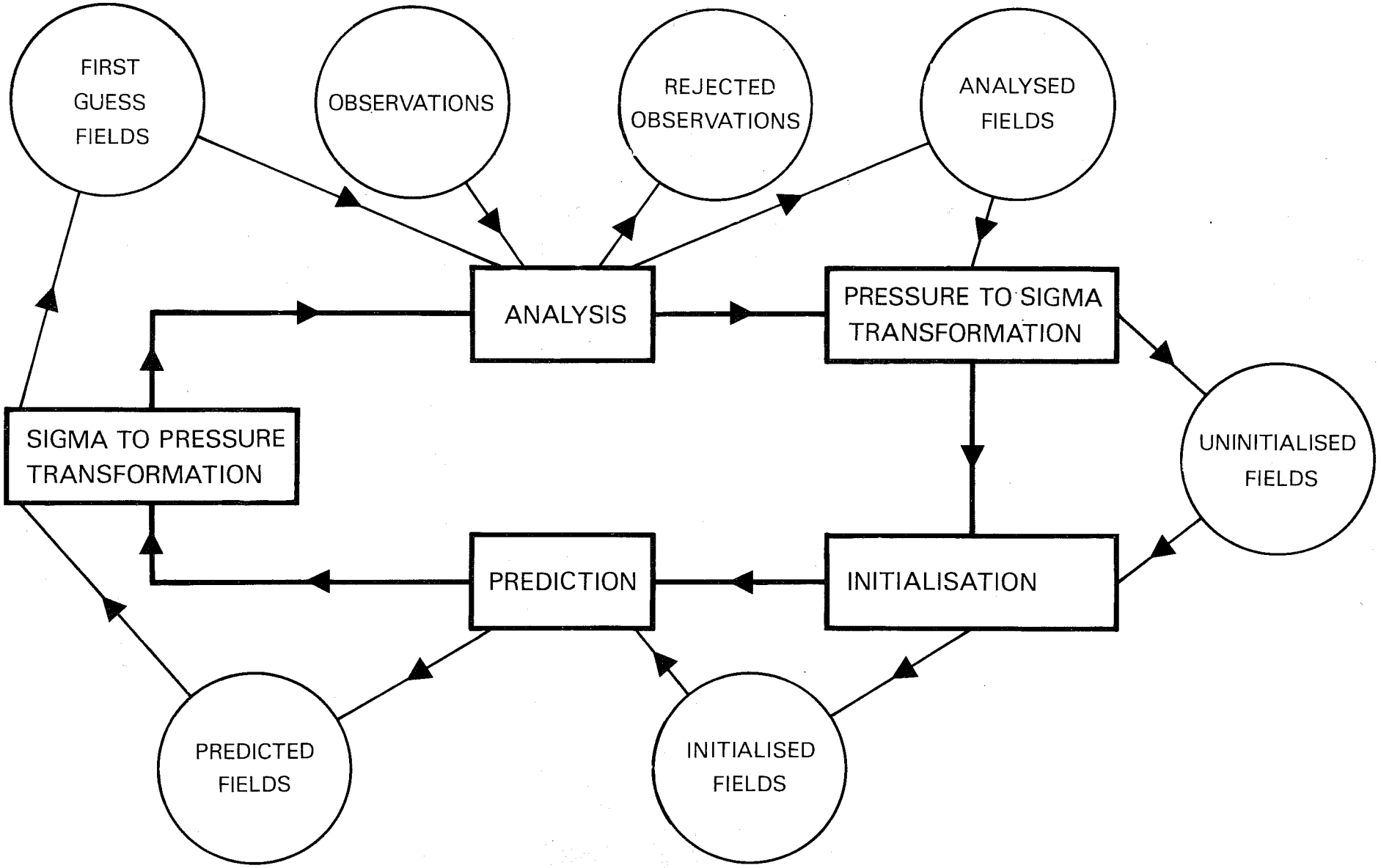


Fig. 1 Interrelation of the main functional components (square boxes) and data files (round boxes) in an assimilation cycle.

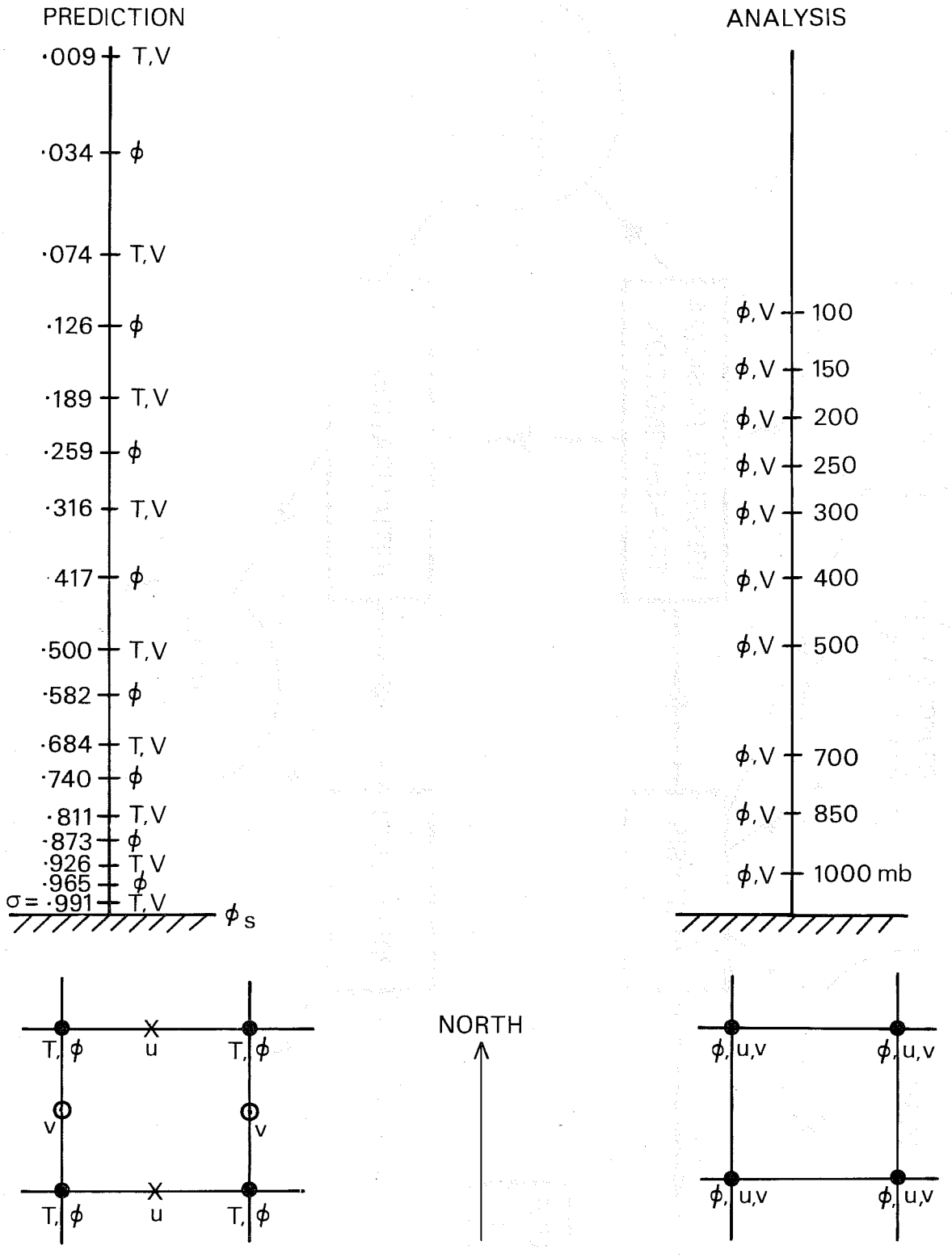


Fig. 2 The vertical and horizontal (latitude-longitude) grids and dispositions of variables in the prediction (left) and analysis (right) coordinate systems.

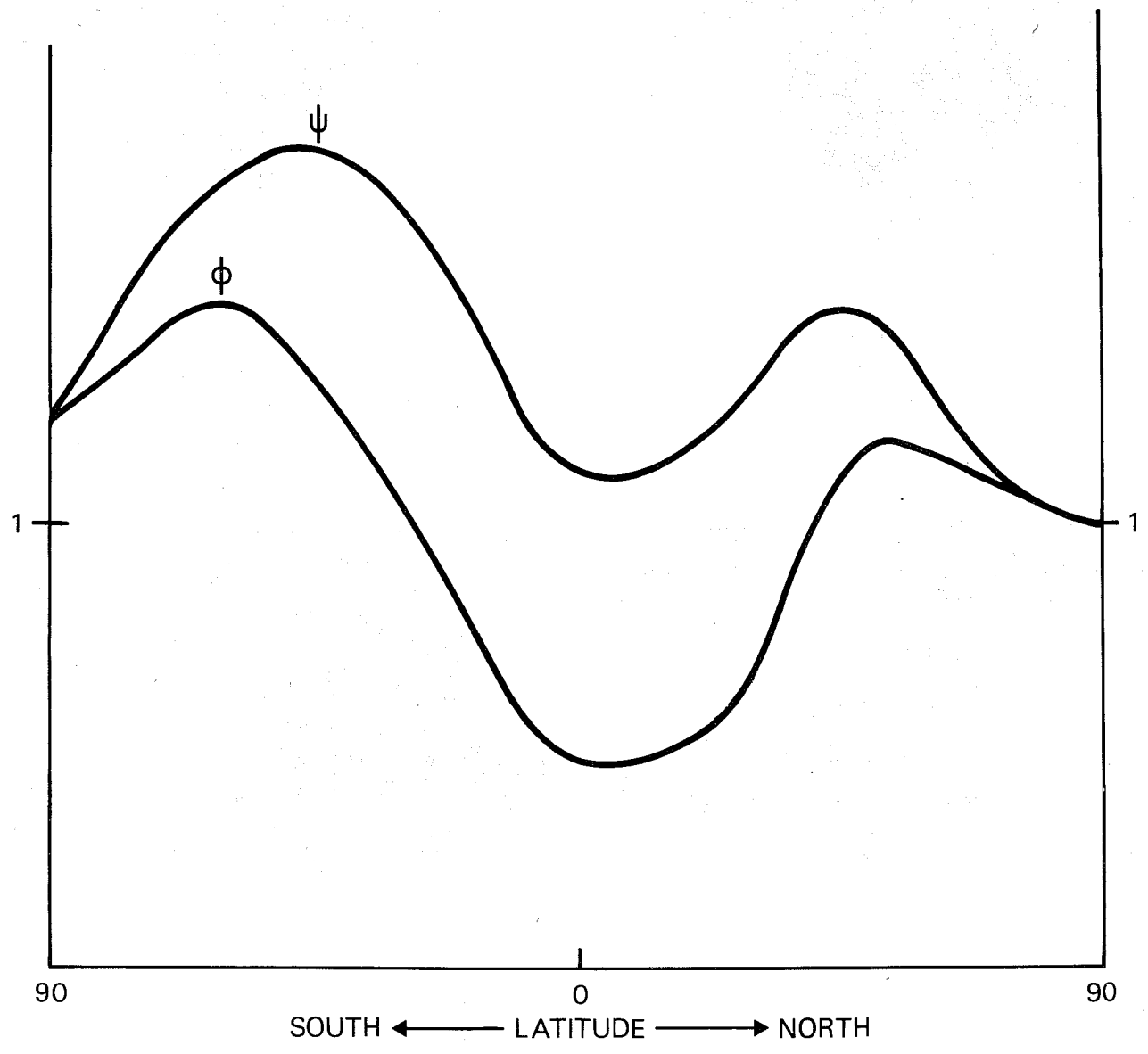
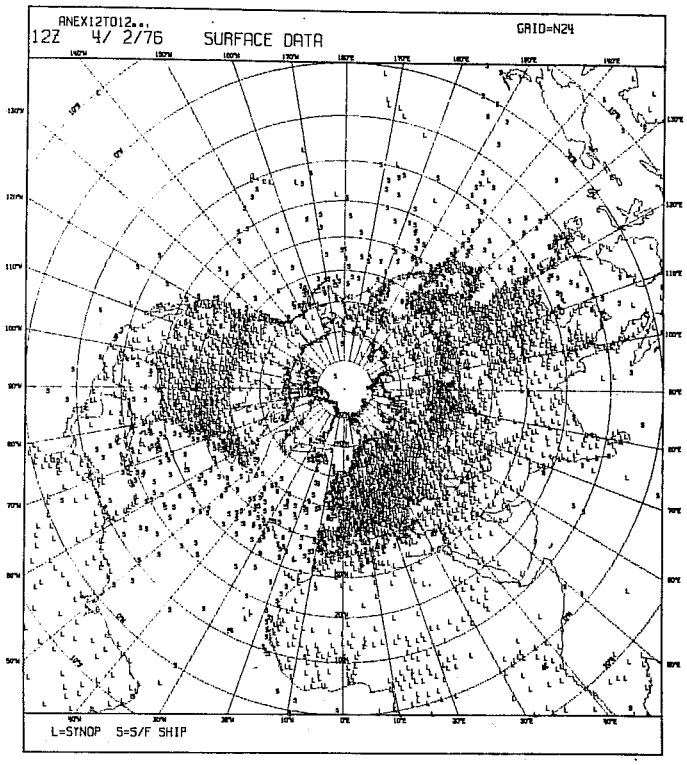
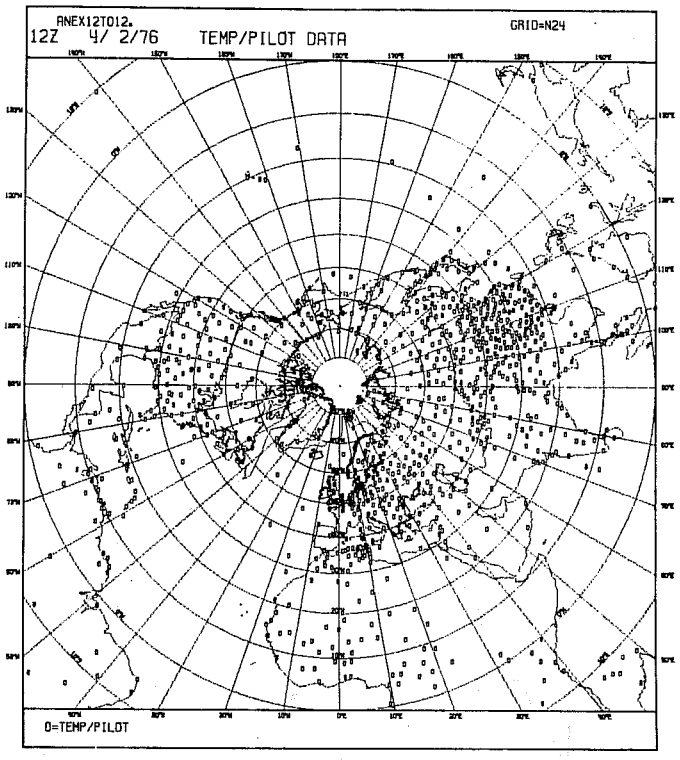
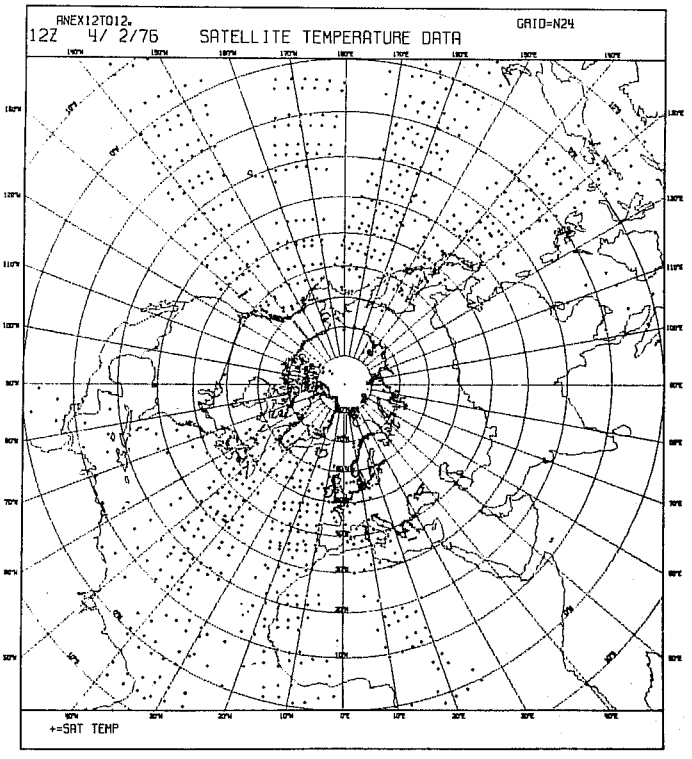


Fig. 3 The latitudinal profile of the relative forecast error for geopotential (ϕ) and stream function (ψ) used in the CONTROL run. See Larsen et al.(1977) for details







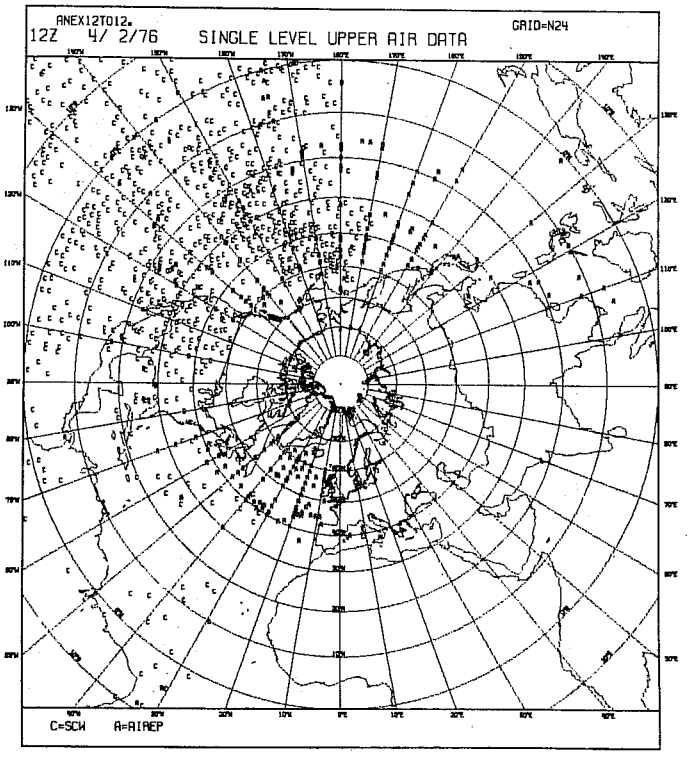
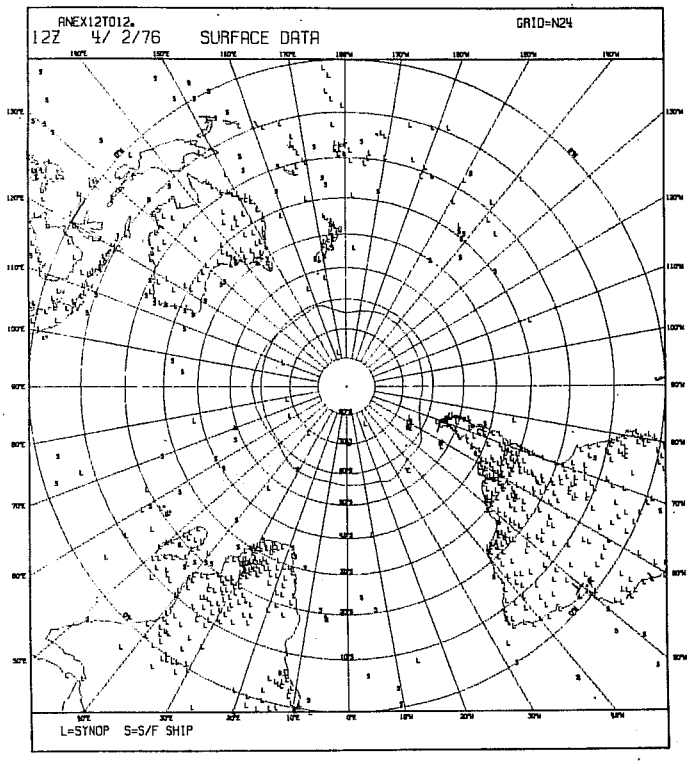
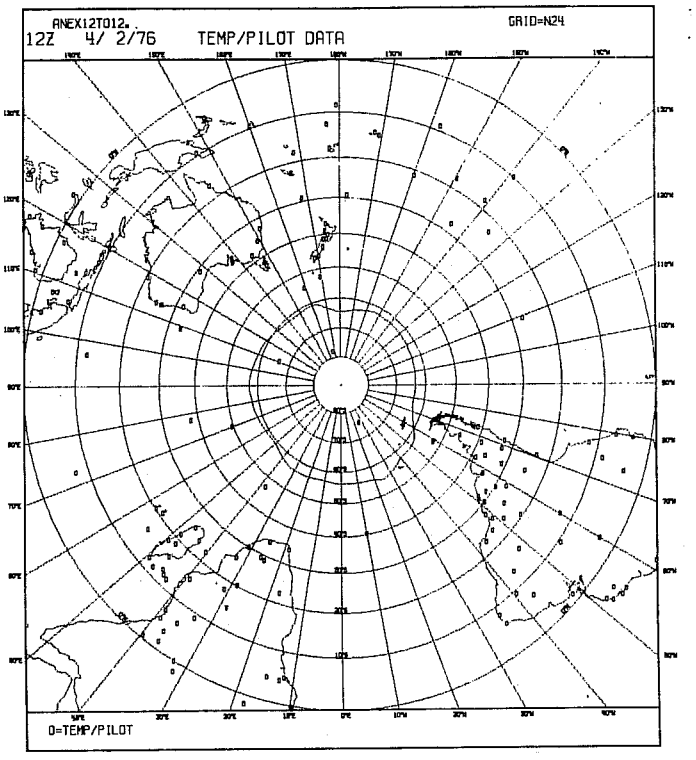
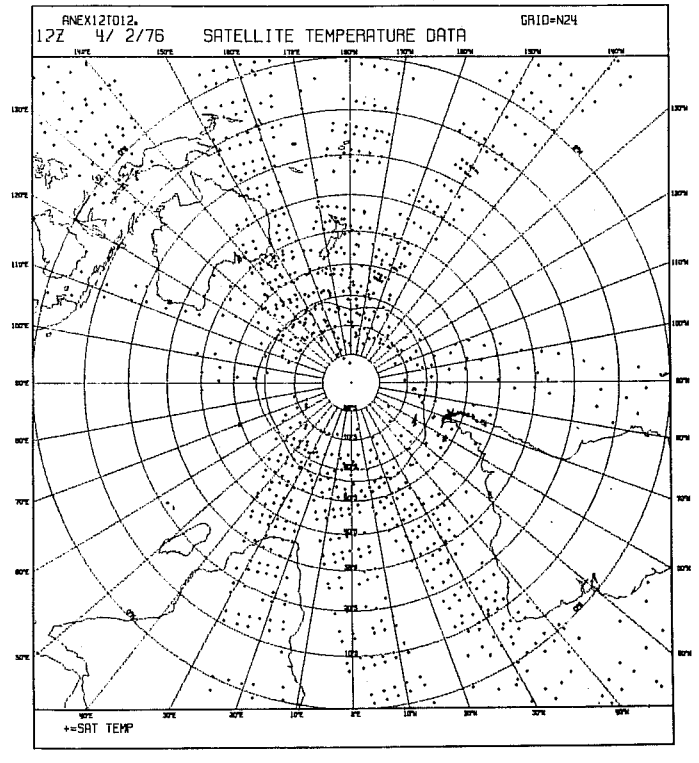


Fig. 4a Observational data base for the northern hemisphere at 12Z, 4 February 1976.







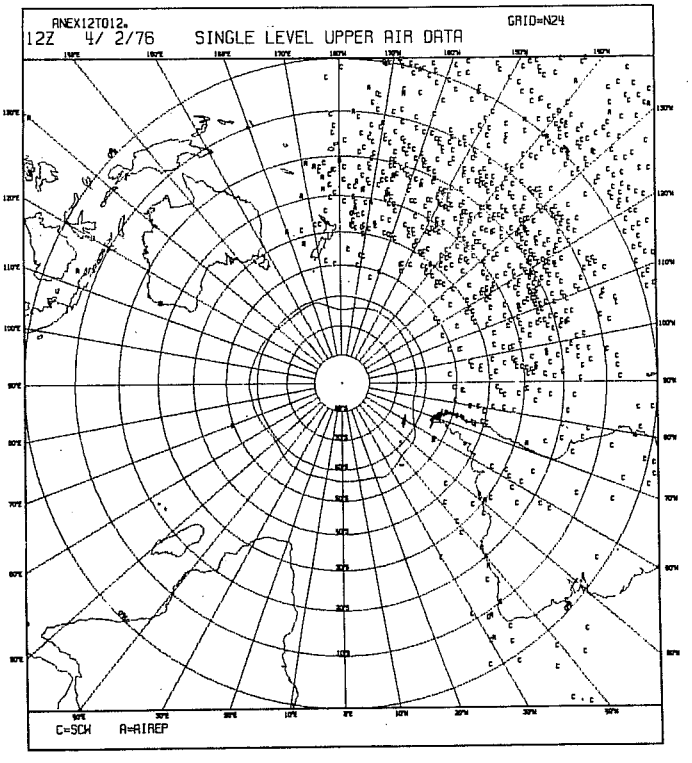
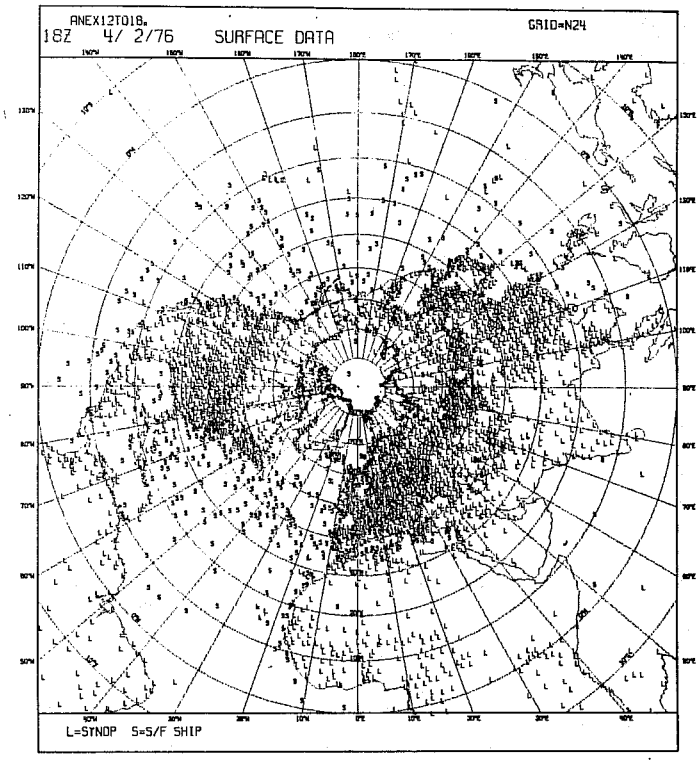
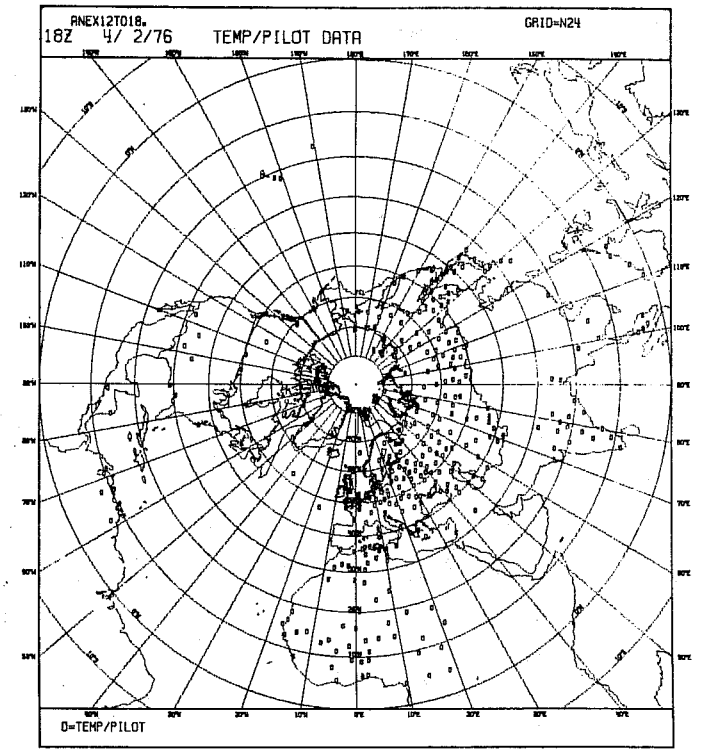
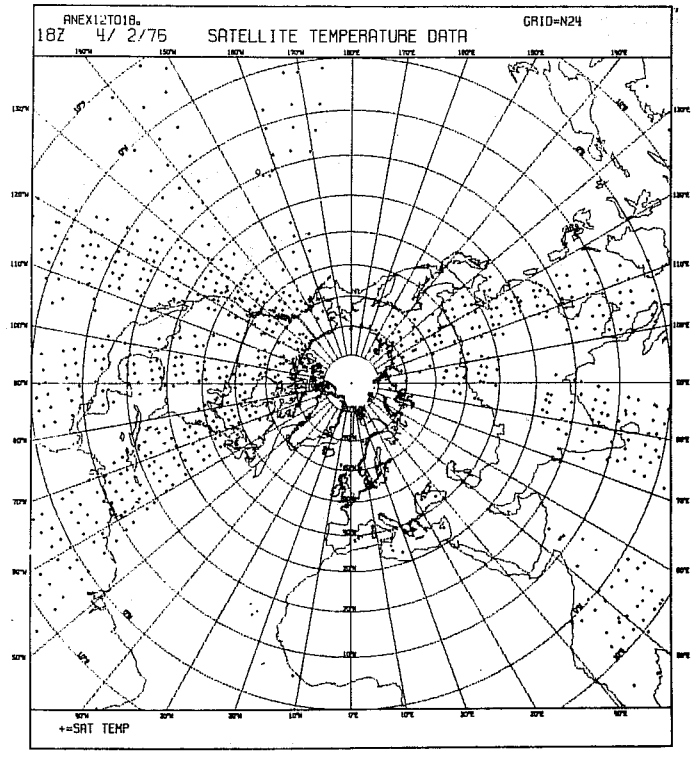


Fig. 4b As in Fig. 4a, for the southern hemisphere.







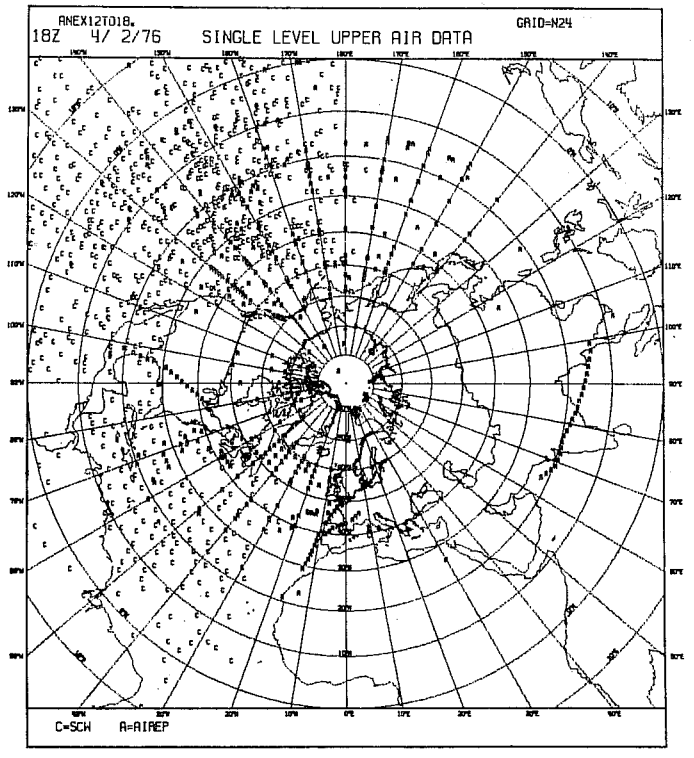
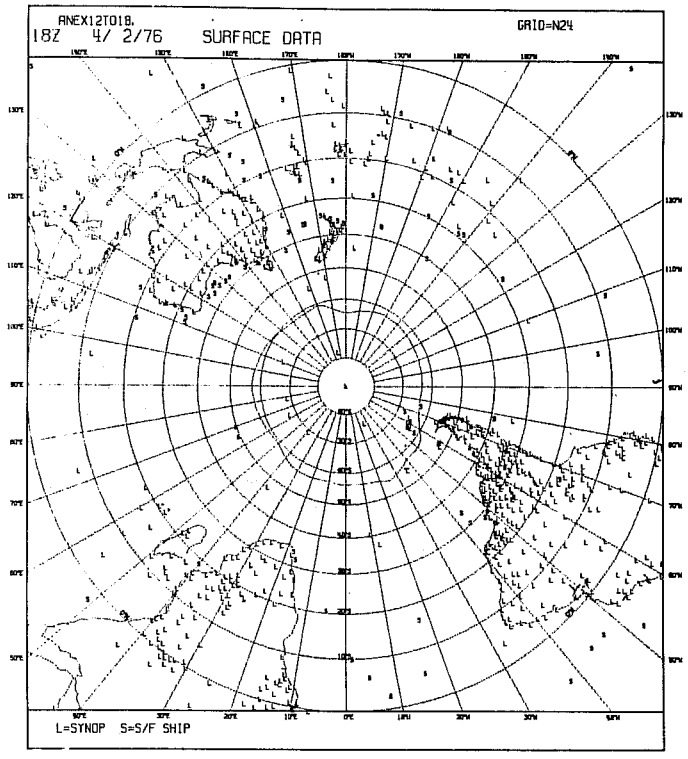
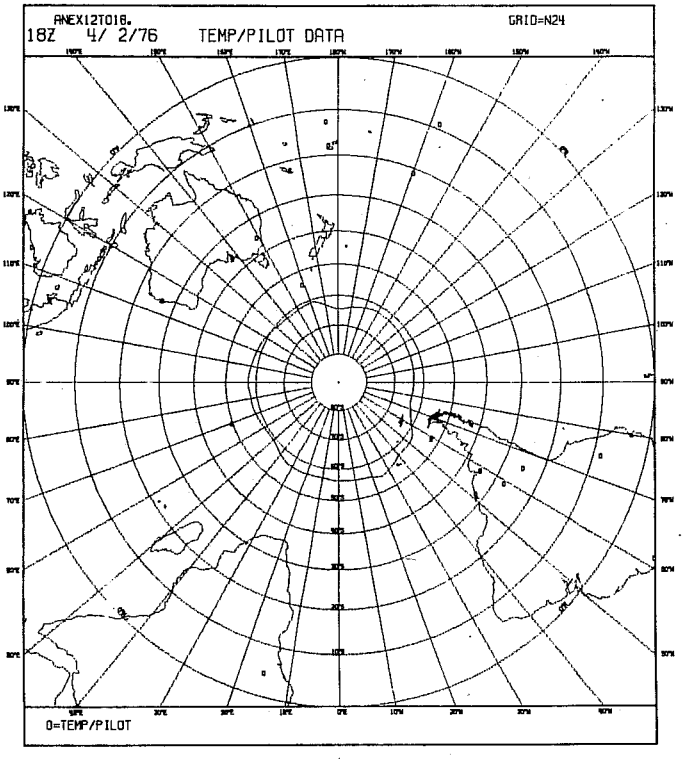
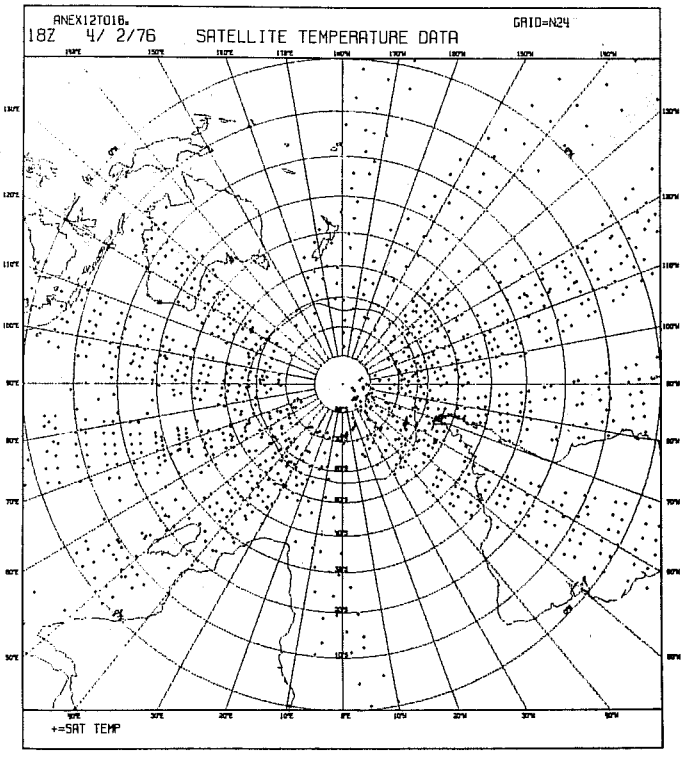


Fig. 5a Observational data base for the northern hemisphere at 18Z, 4 February 1976.







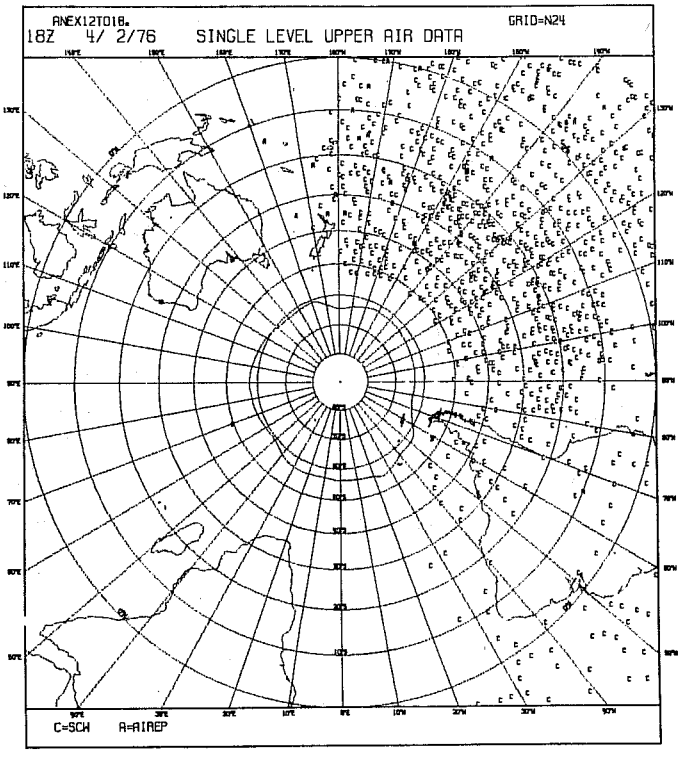
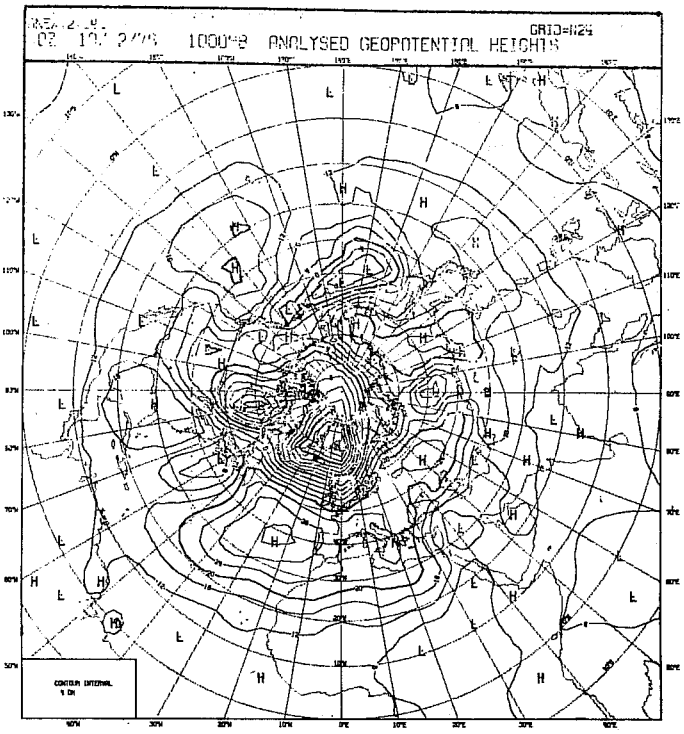
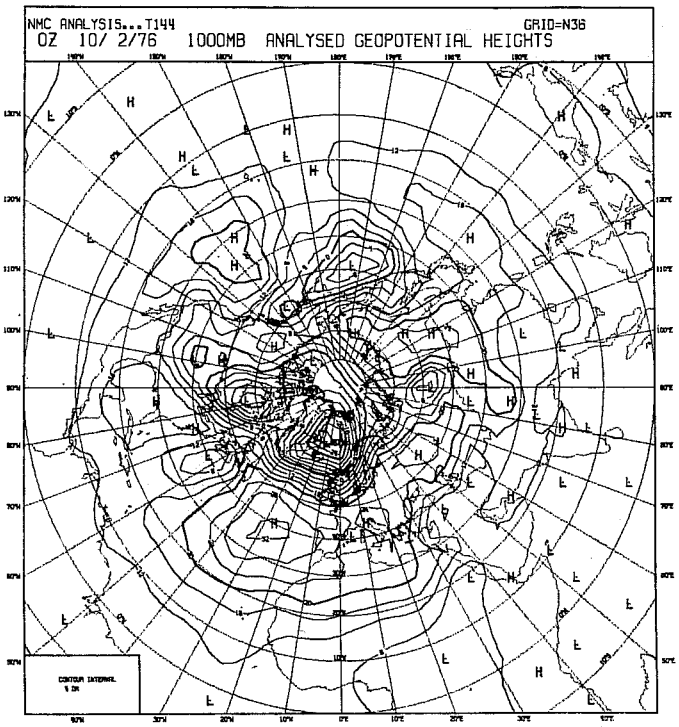


Fig. 5b As in Fig. 5a, for the southern hemisphere.





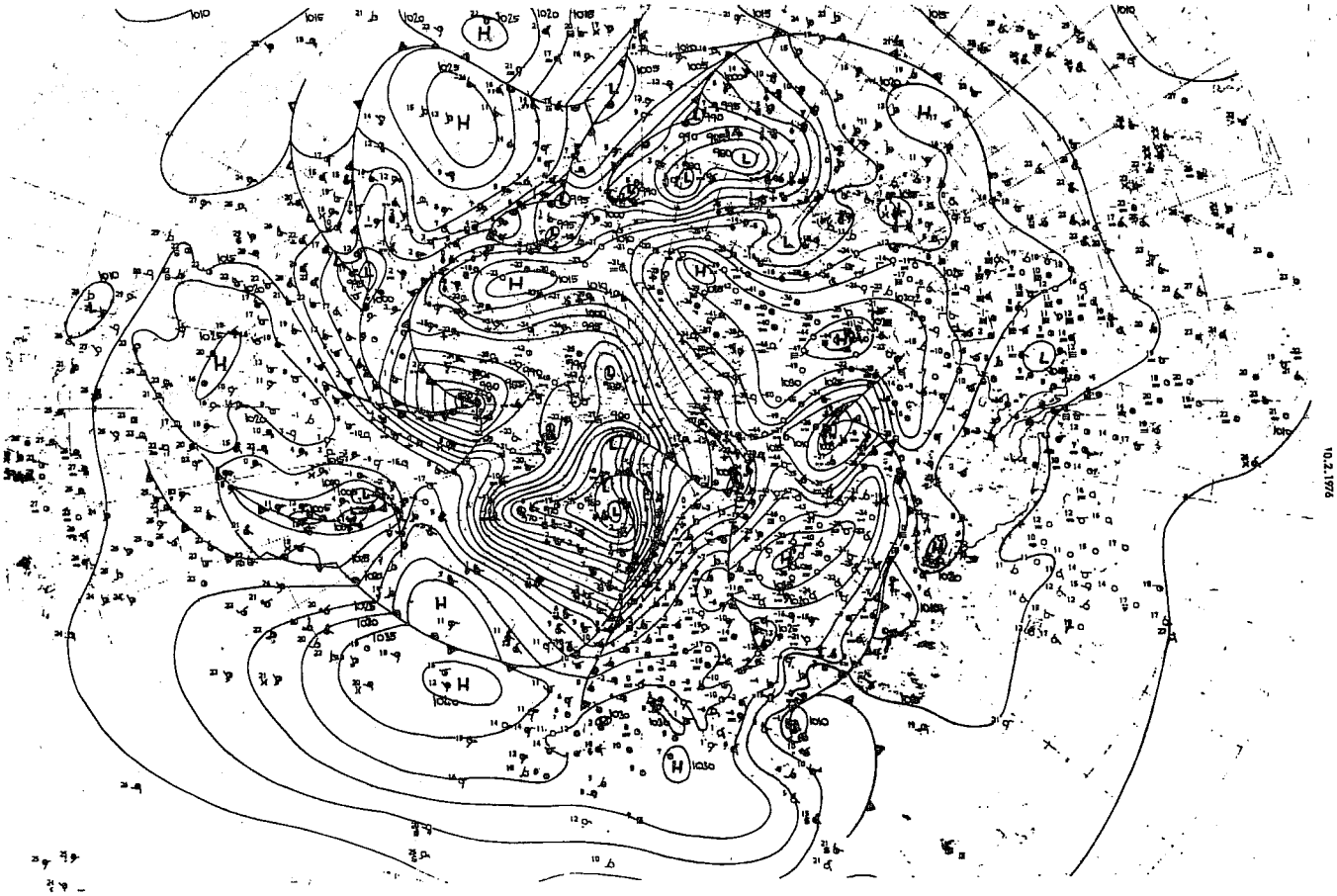
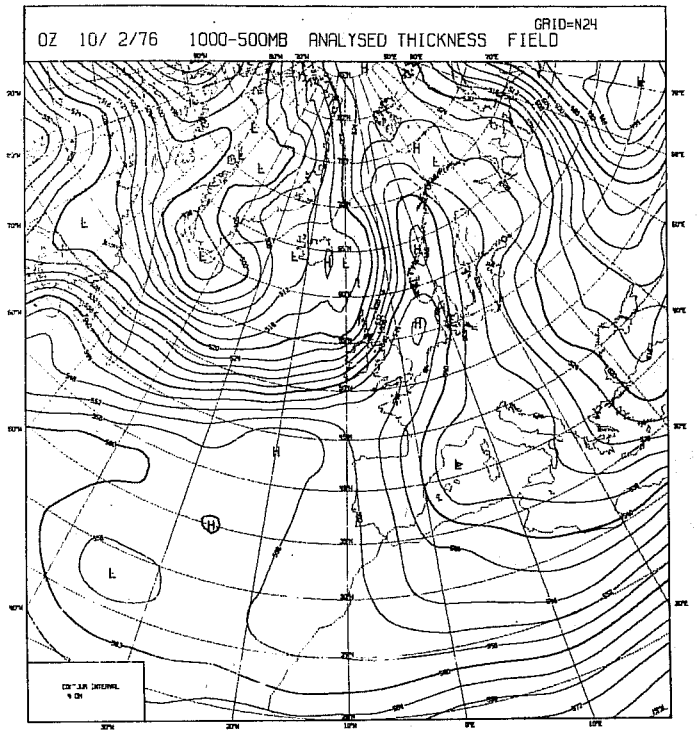
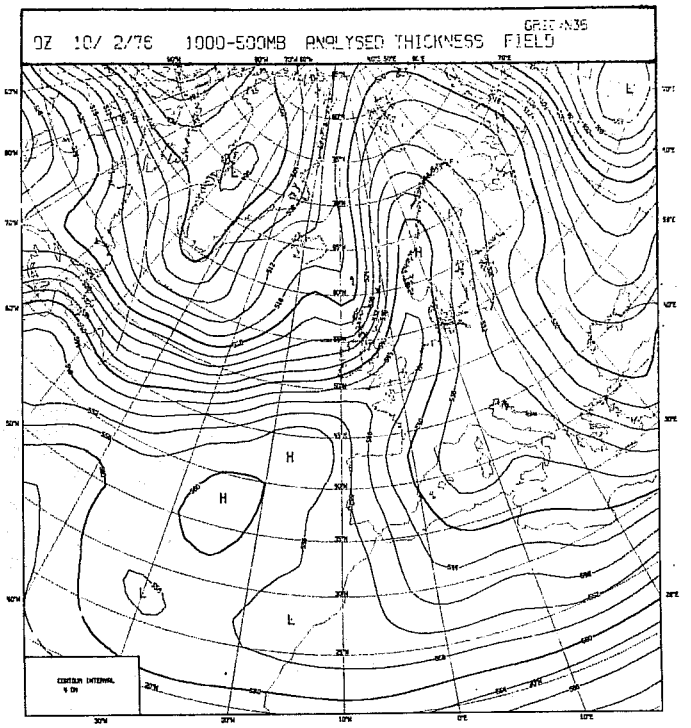


Fig. 6 CONTROL (top left), NMC (top right), and Deutscher Wetterdienst analyses of the northern hemisphere 1000 mb geopotential, or sea level pressure, for 00Z, 10 February 1976.





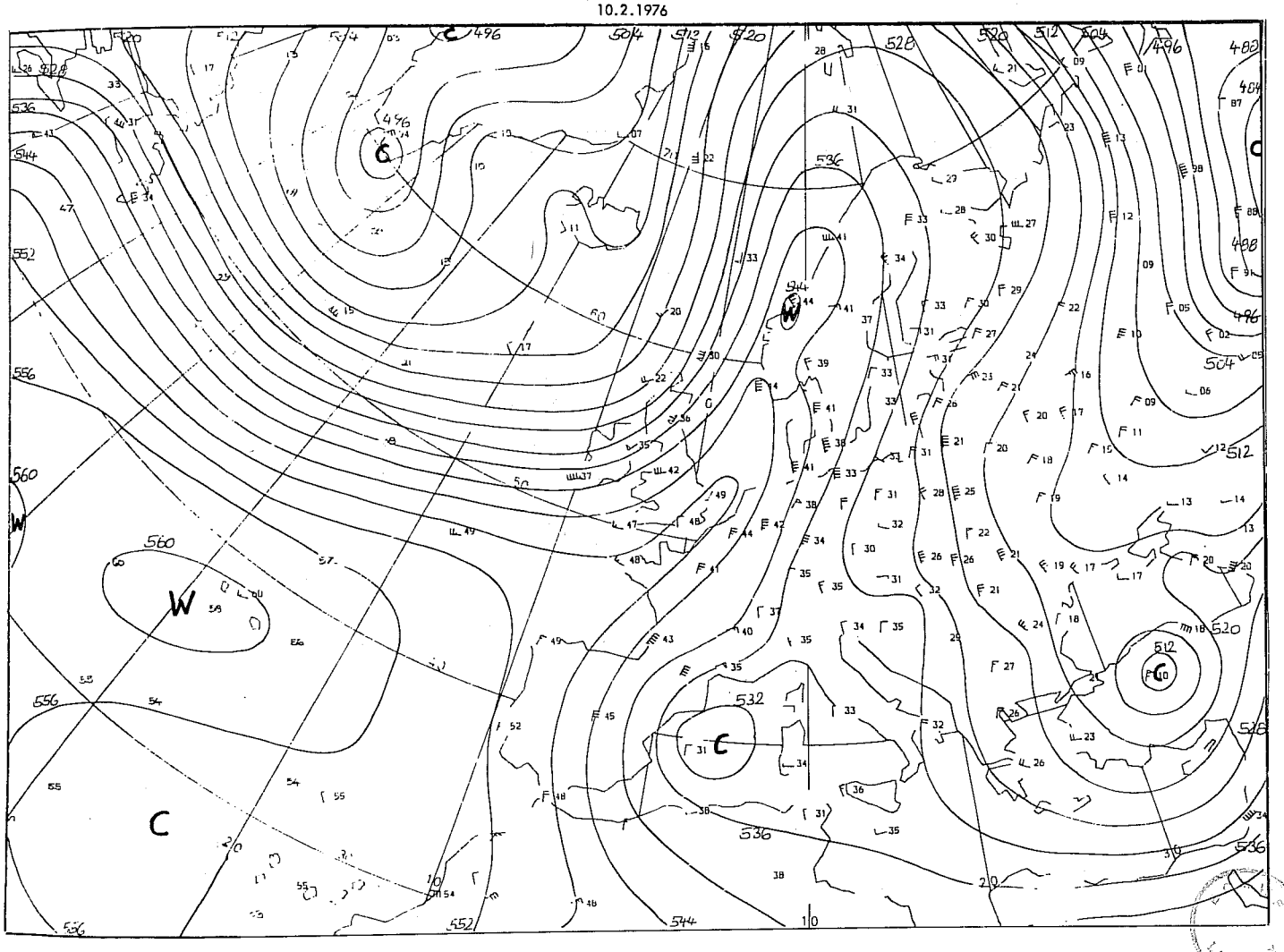


Fig. 7 As in Fig. 6, for the 1000-500 mb thickness, 00Z, 10 February 1976.

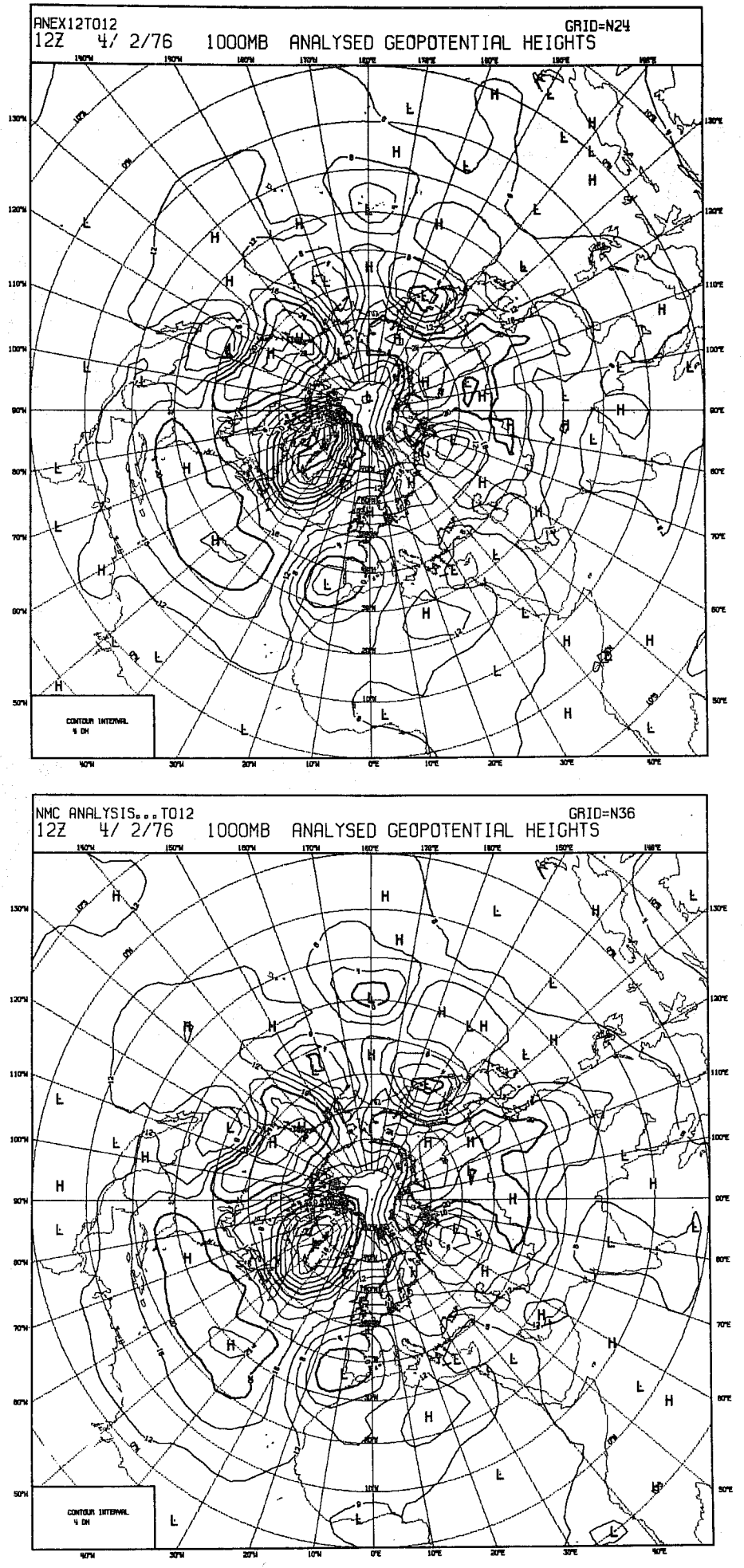
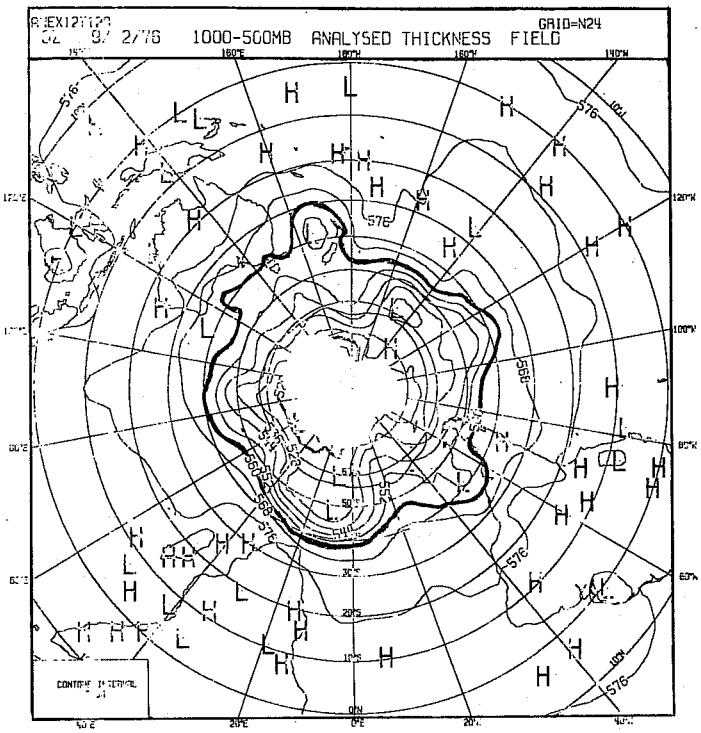
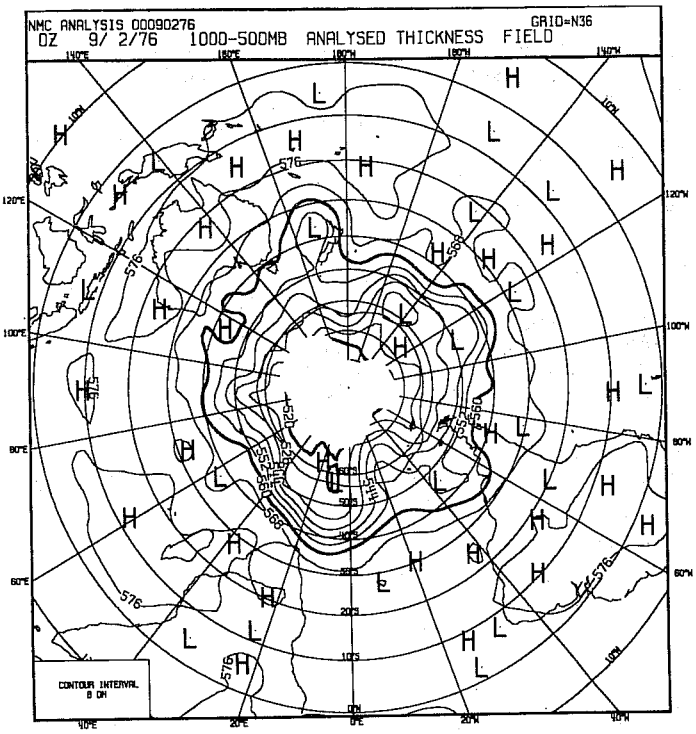


Fig. 8 CONTROL and NMC 1000 mb geopotential analyses for the northern hemisphere, 12Z, 4 February 1976.





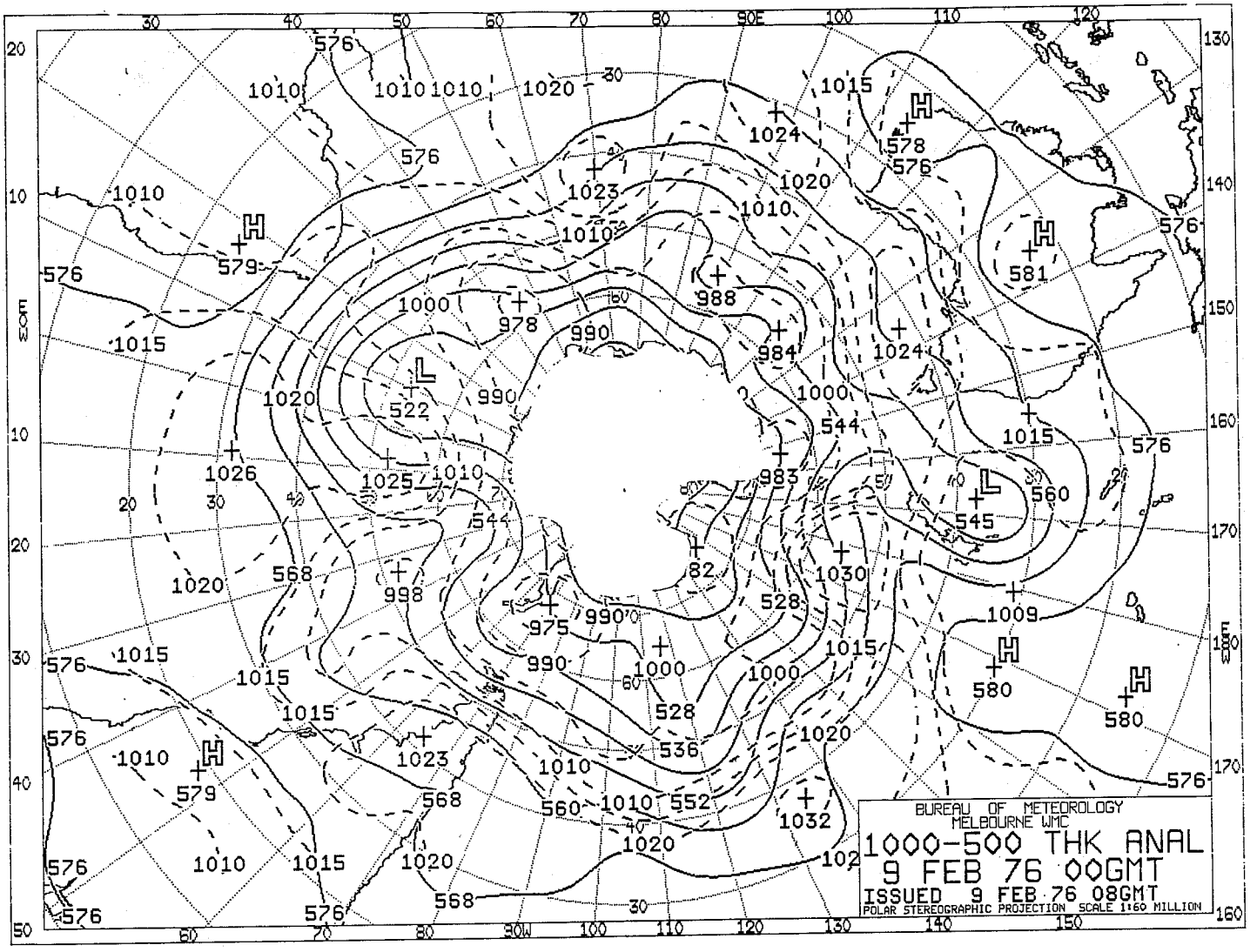


Fig. 9 CONTROL (top left), NMC (top right) and Australian (lower) 1000-500 mb thickness analyses for the southern hemisphere at 00Z, 9 February 1976. The Australian analysis also shows sea level pressure, dashed.

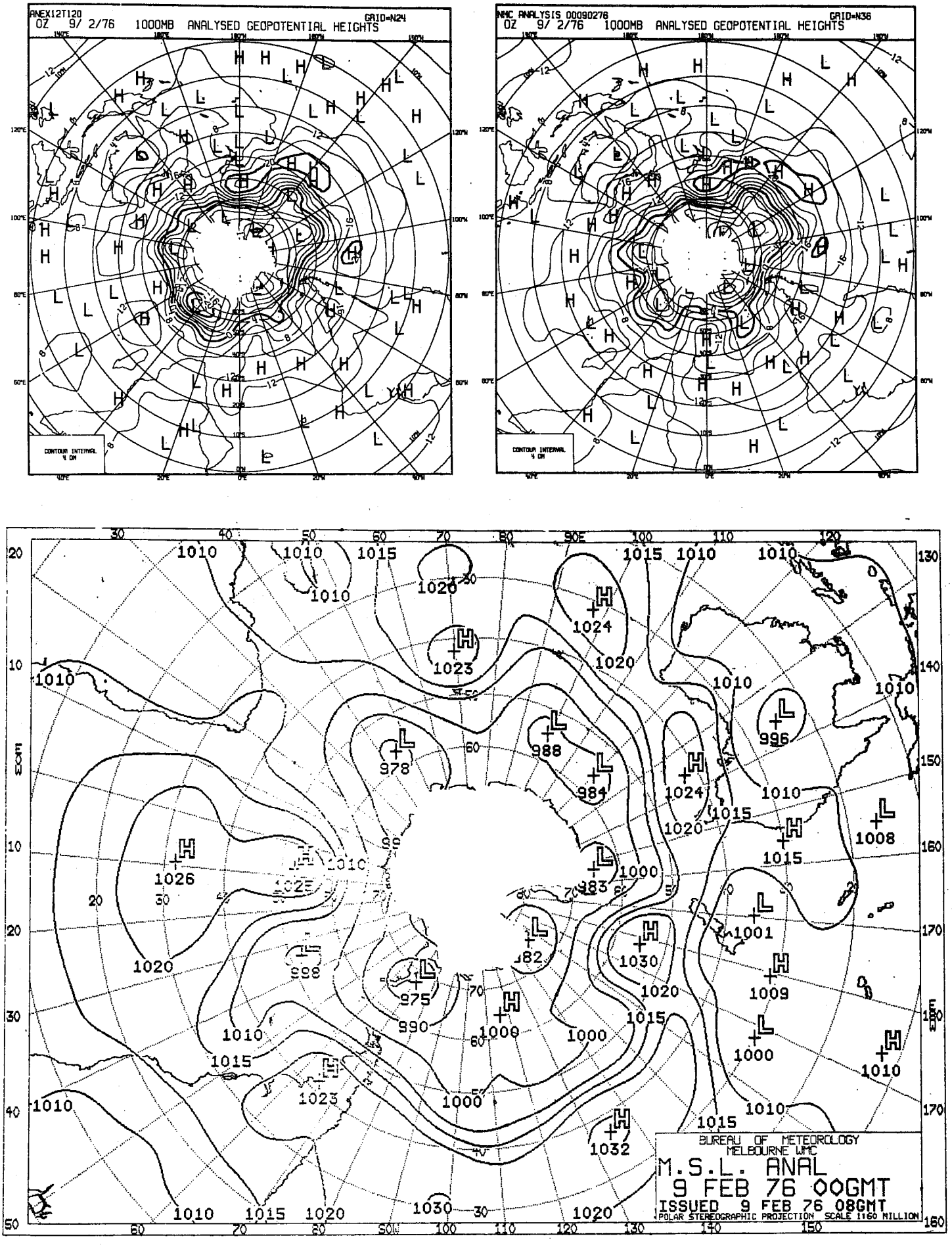


Fig. 10 As in Fig. 9, for the 1000 mb geopotential, or sea level pressure, at 00Z, 9 February 1976.

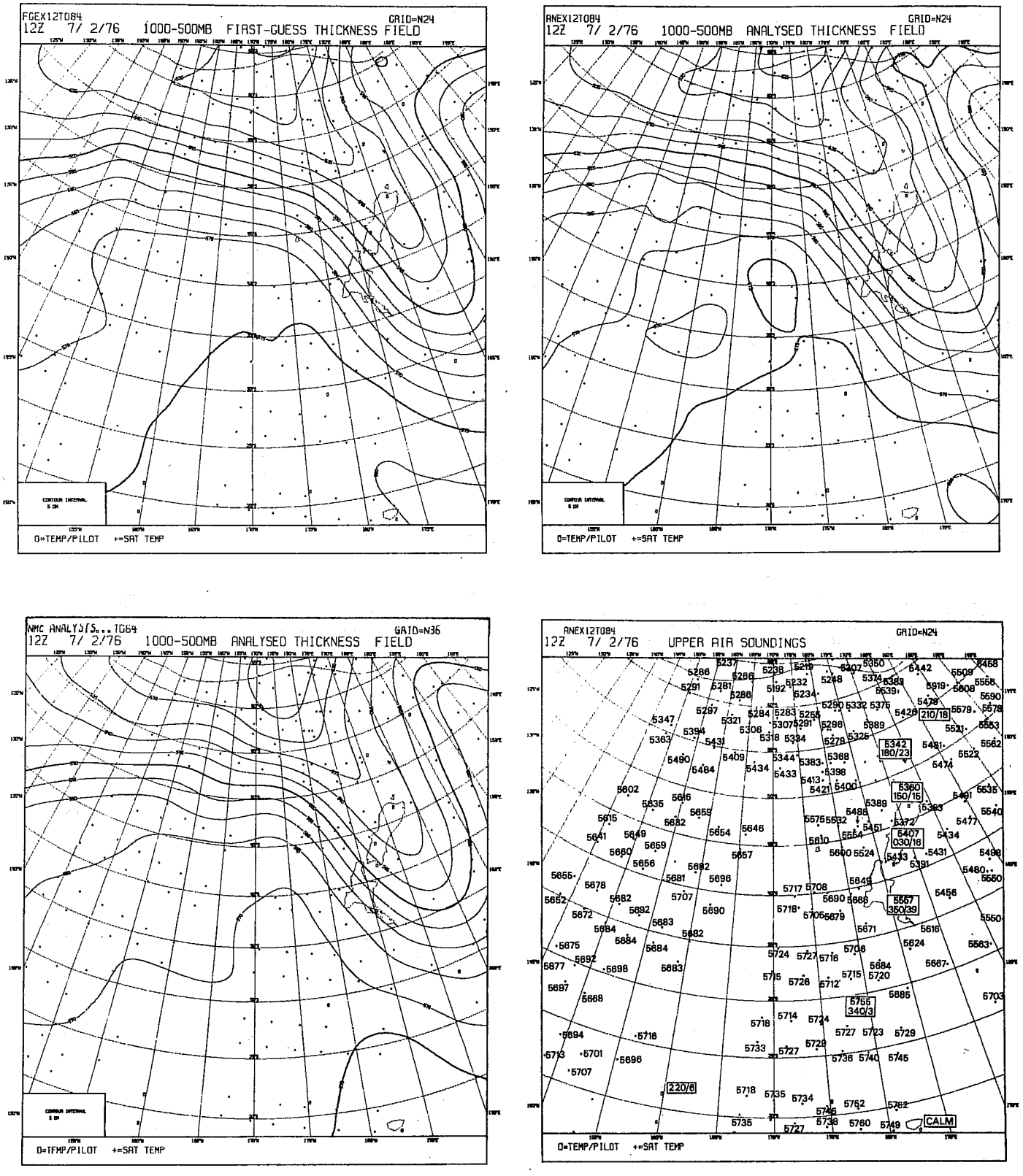


Fig. 11 CONTROL first guess (top left), CONTROL analysis (top right) and NMC analysis (bottom left) of 1000-500 mb thickness for 12Z, 7 February 1976. Observed thicknesses in dekametres are plotted at bottom right. Plots in boxes are rawinsonde thicknesses and shears.

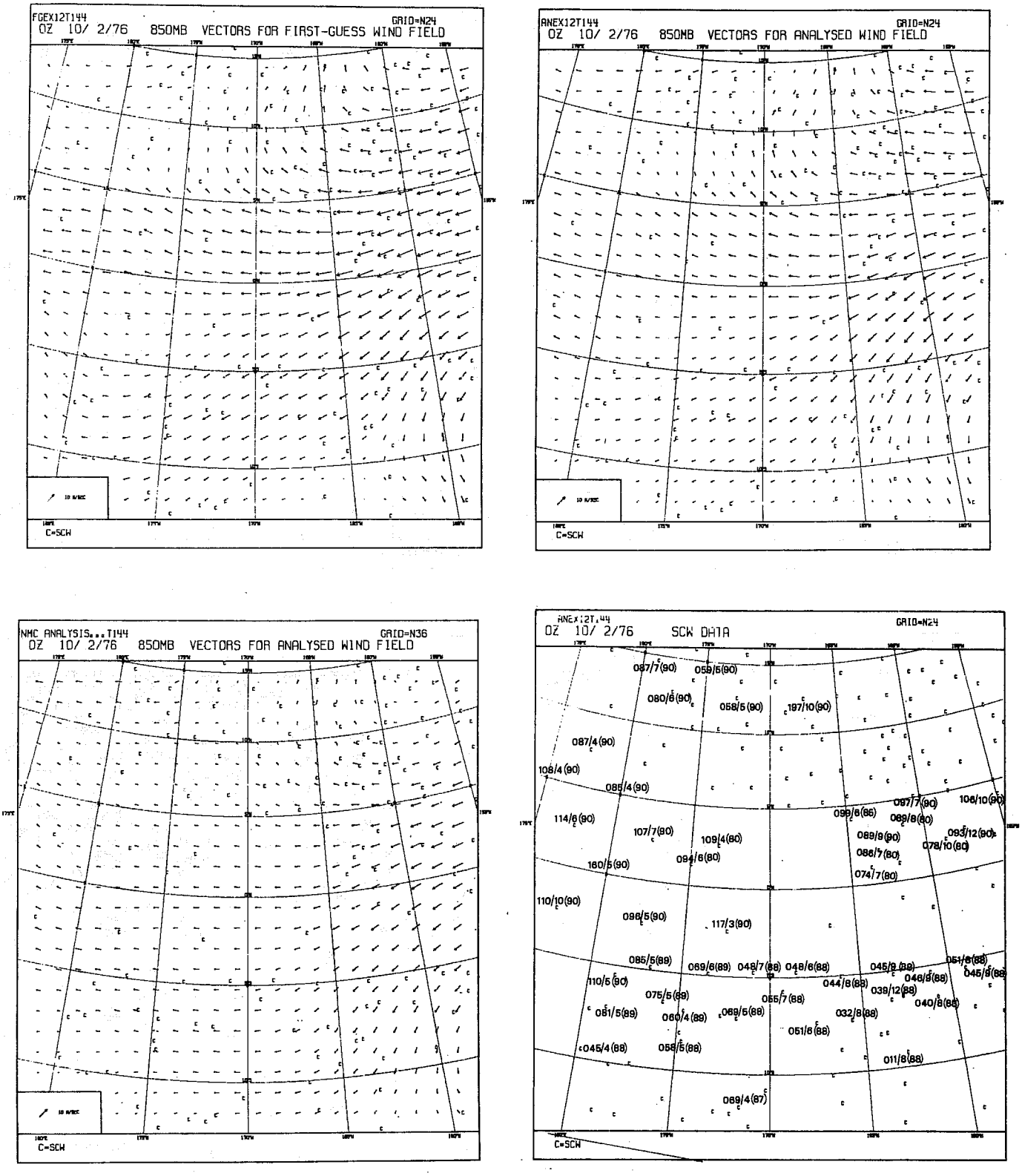
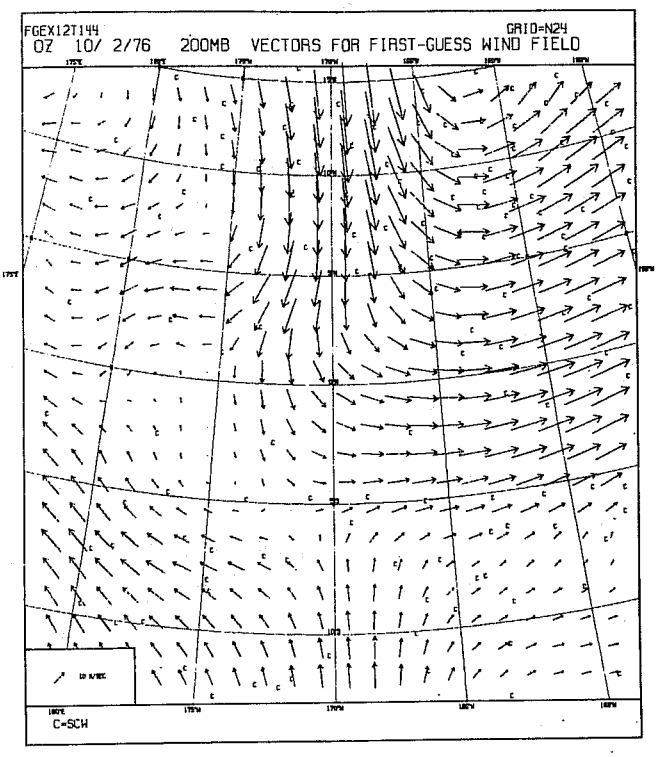
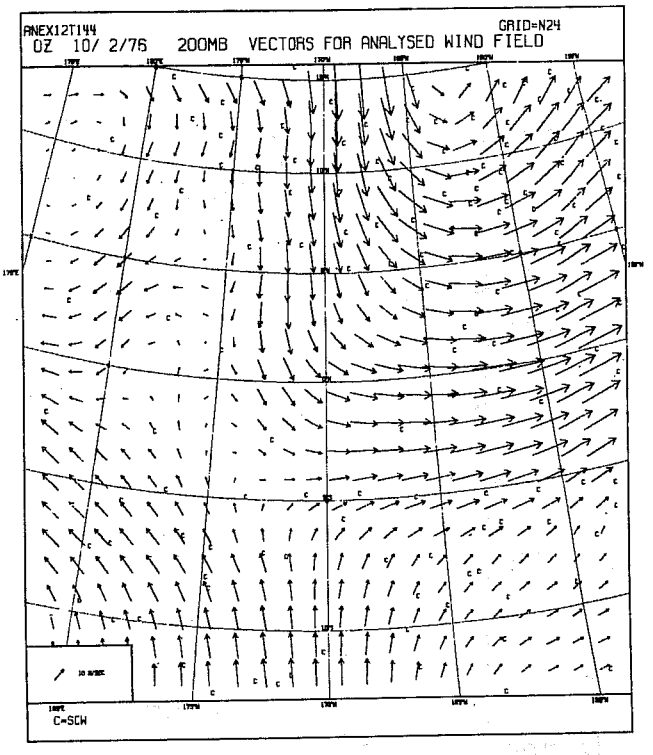
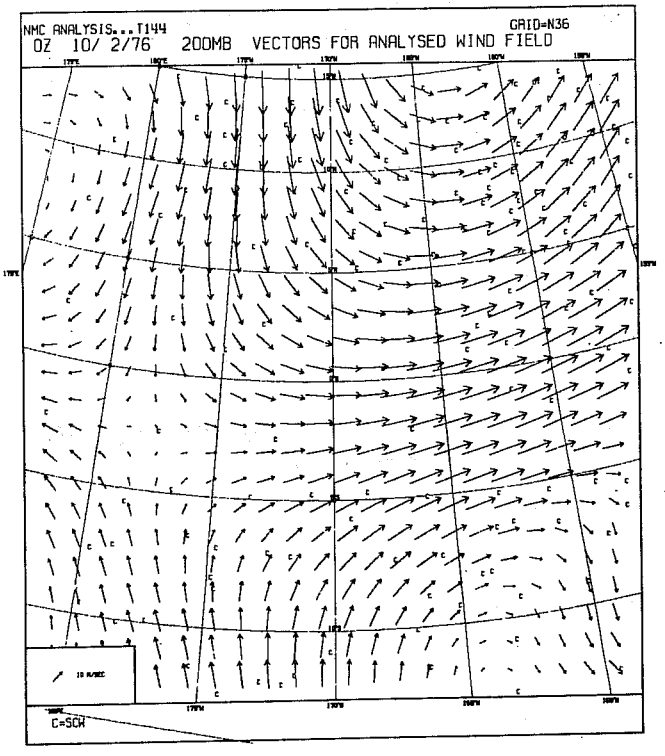


Fig. 12 CONTROL first guess (top left), CONTROL analysis (top right), and NMC analysis (bottom left) of 850 mb vector wind for 00Z, 10 February 1976. Observed winds are plotted at bottom right, with level (tens of mb) in parentheses.







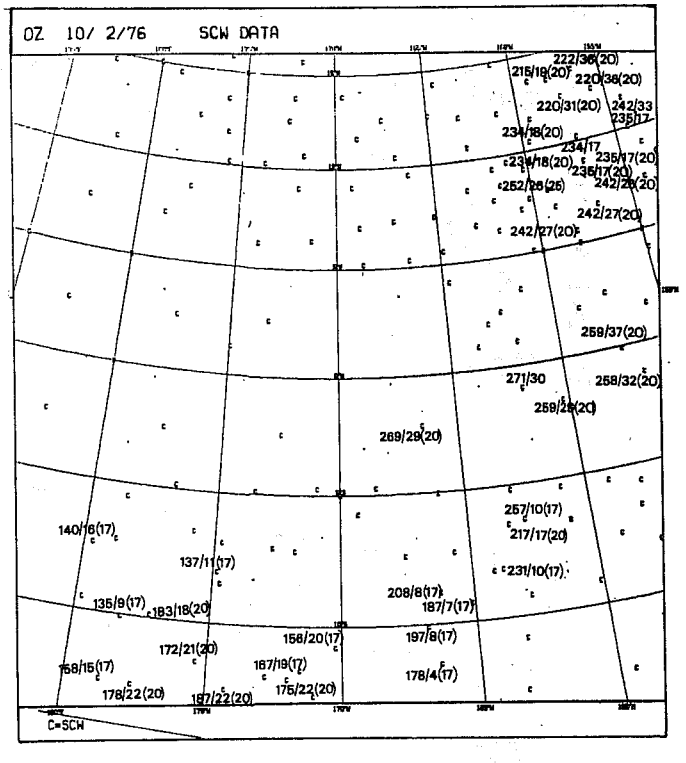


Fig. 13 As in Fig. 12, for the 200 mb vector wind, at 00Z, 10 February 1976.

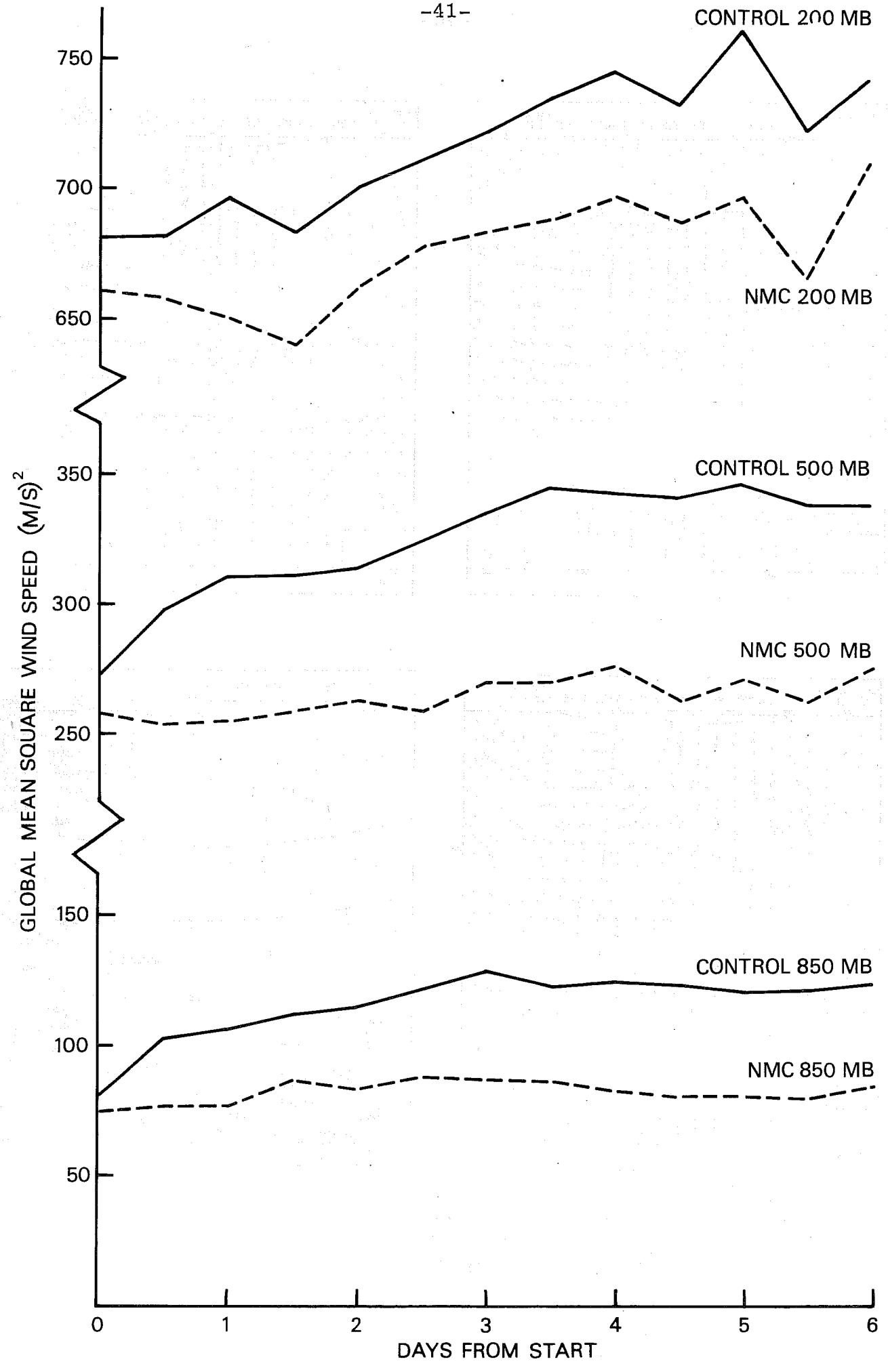


Fig. 14a. Global mean square wind speeds for CONTROL and NMC analyses

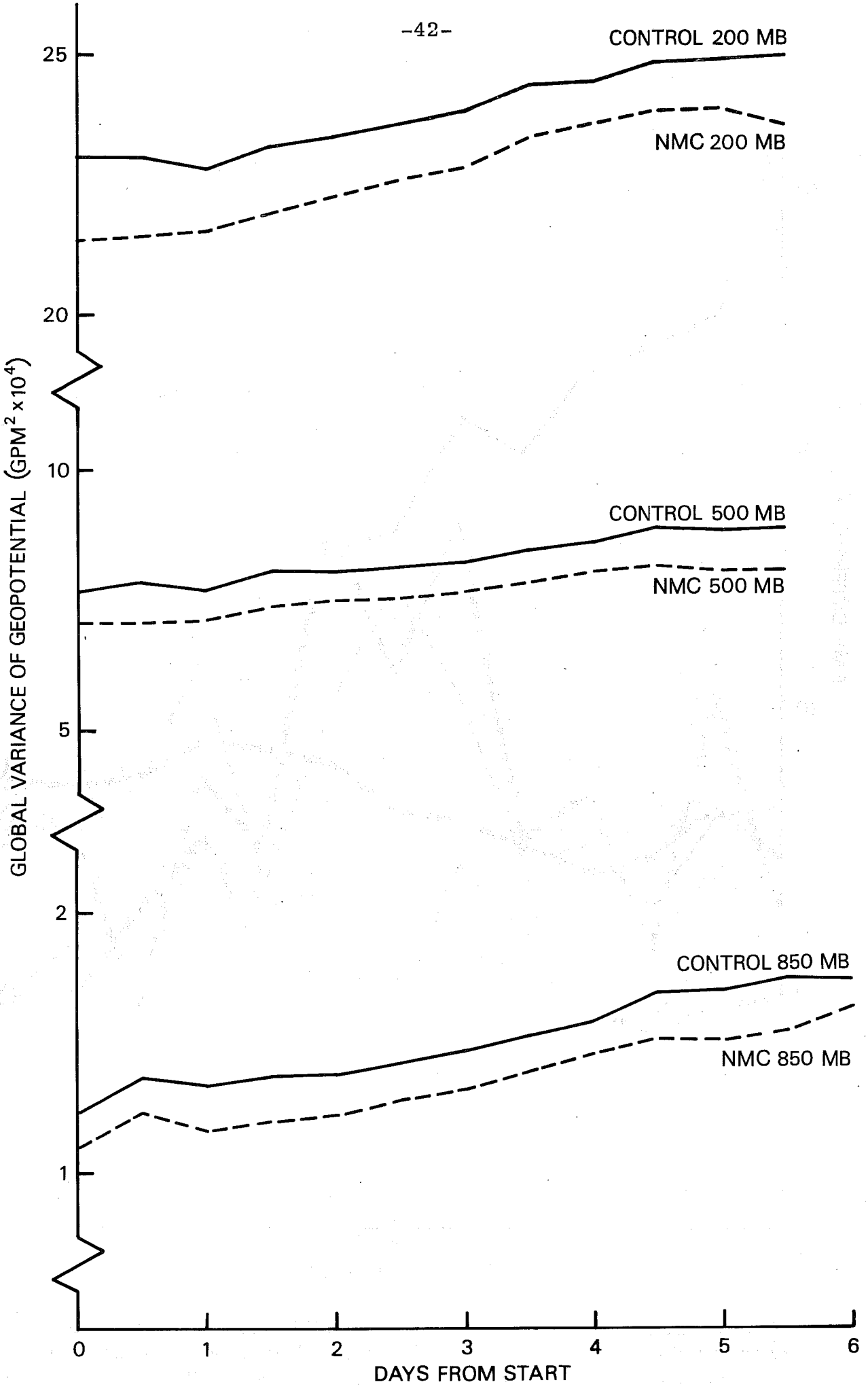


Fig. 14b Global variance of geopotential for CONTROL and NMC analyses.

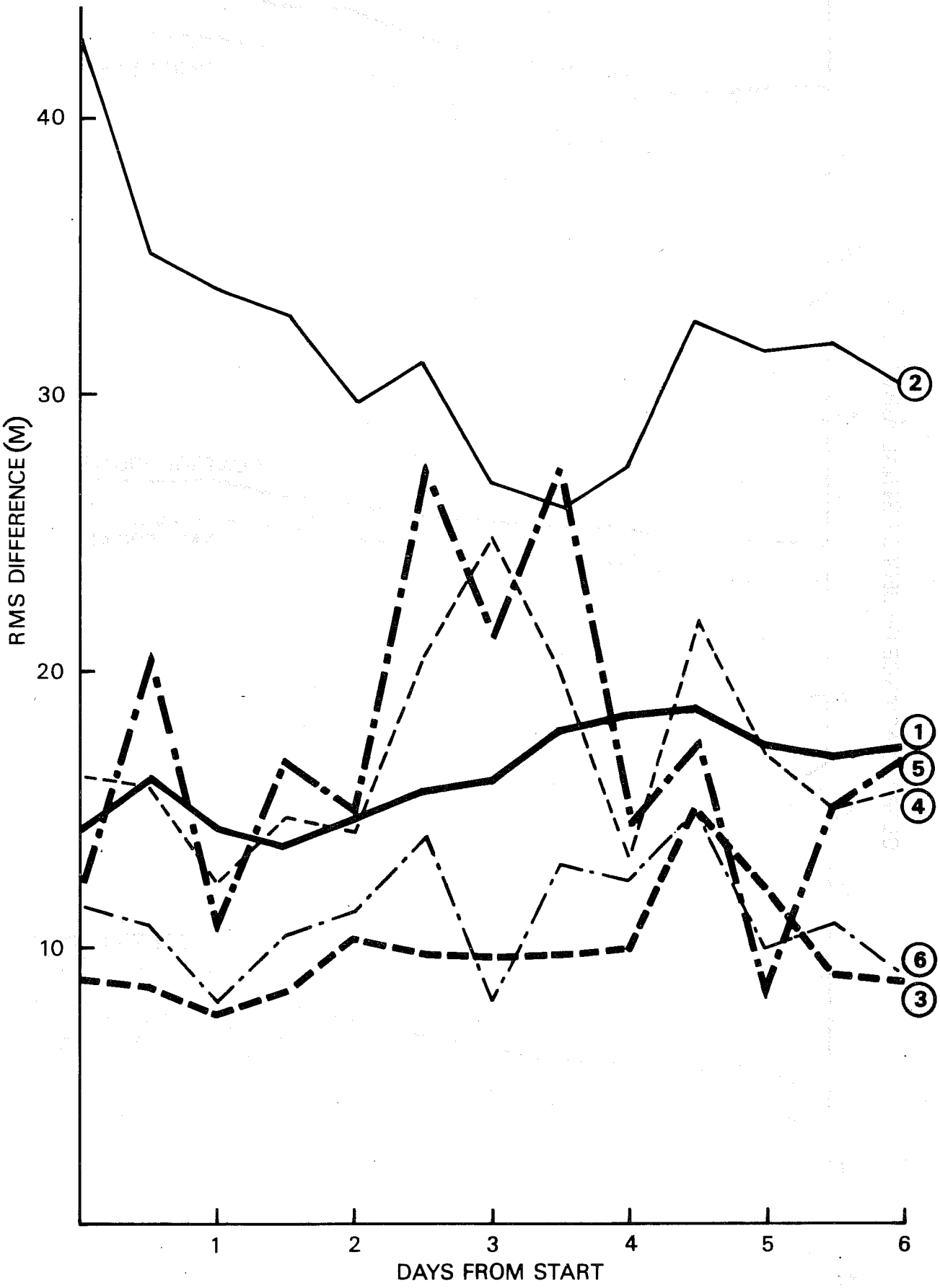


Fig. 15a RMS differences between CONTROL and NMC 1000 mb geopotential analyses for (1) northern hemisphere (2) southern hemisphere (0-70 south), (3) Europe (69-36 north, 2 west-47 east), (4) North Atlantic (69-36 north, 77-2 west), (5) North Pacific (47-28 north, 148 east-126 west), and (6) Australia-New Zealand (13-47 south,

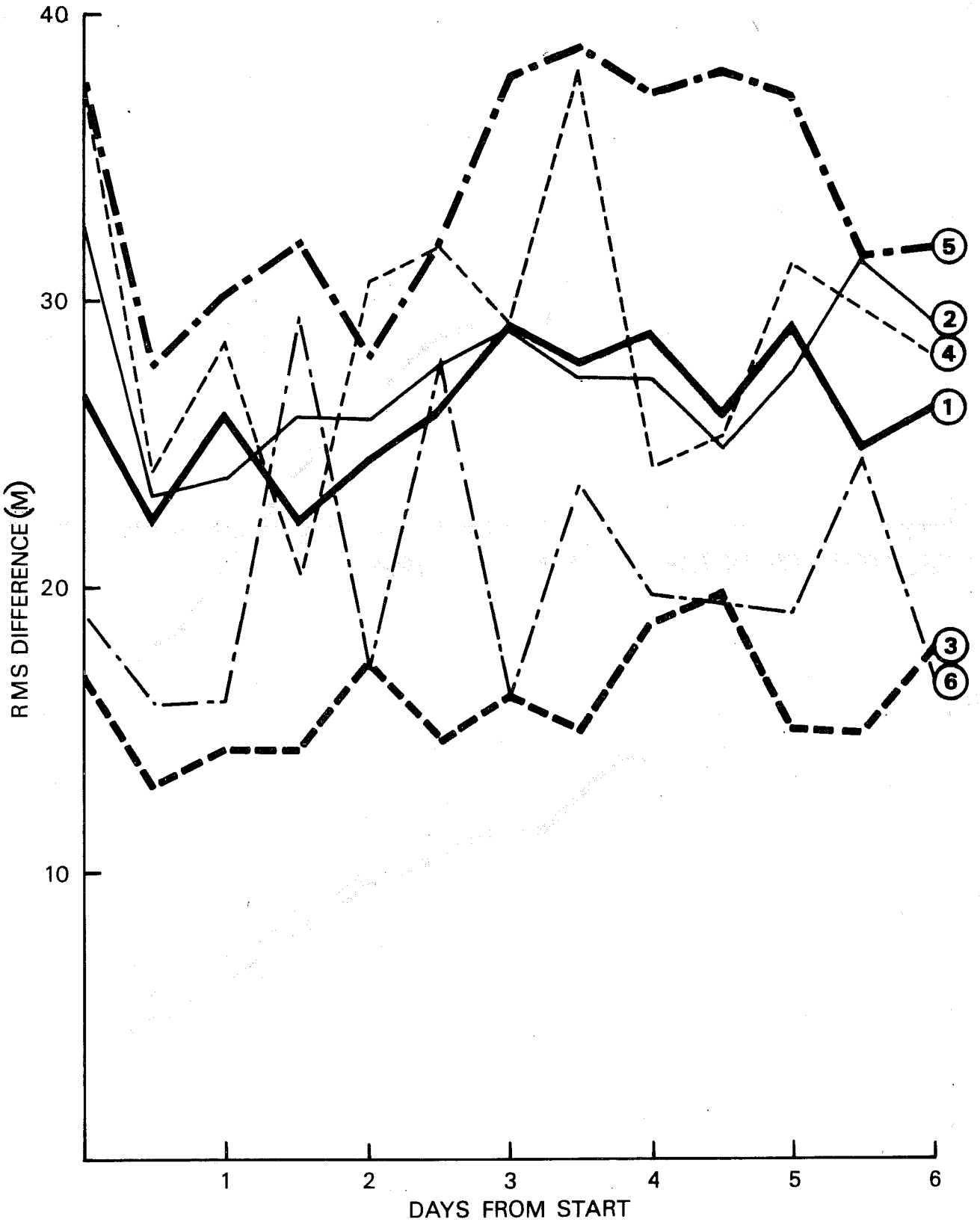


Fig. 15b As for Fig. 15a, for RMS differences between CONTROL and NMC 1000-500mb thickness analyses.

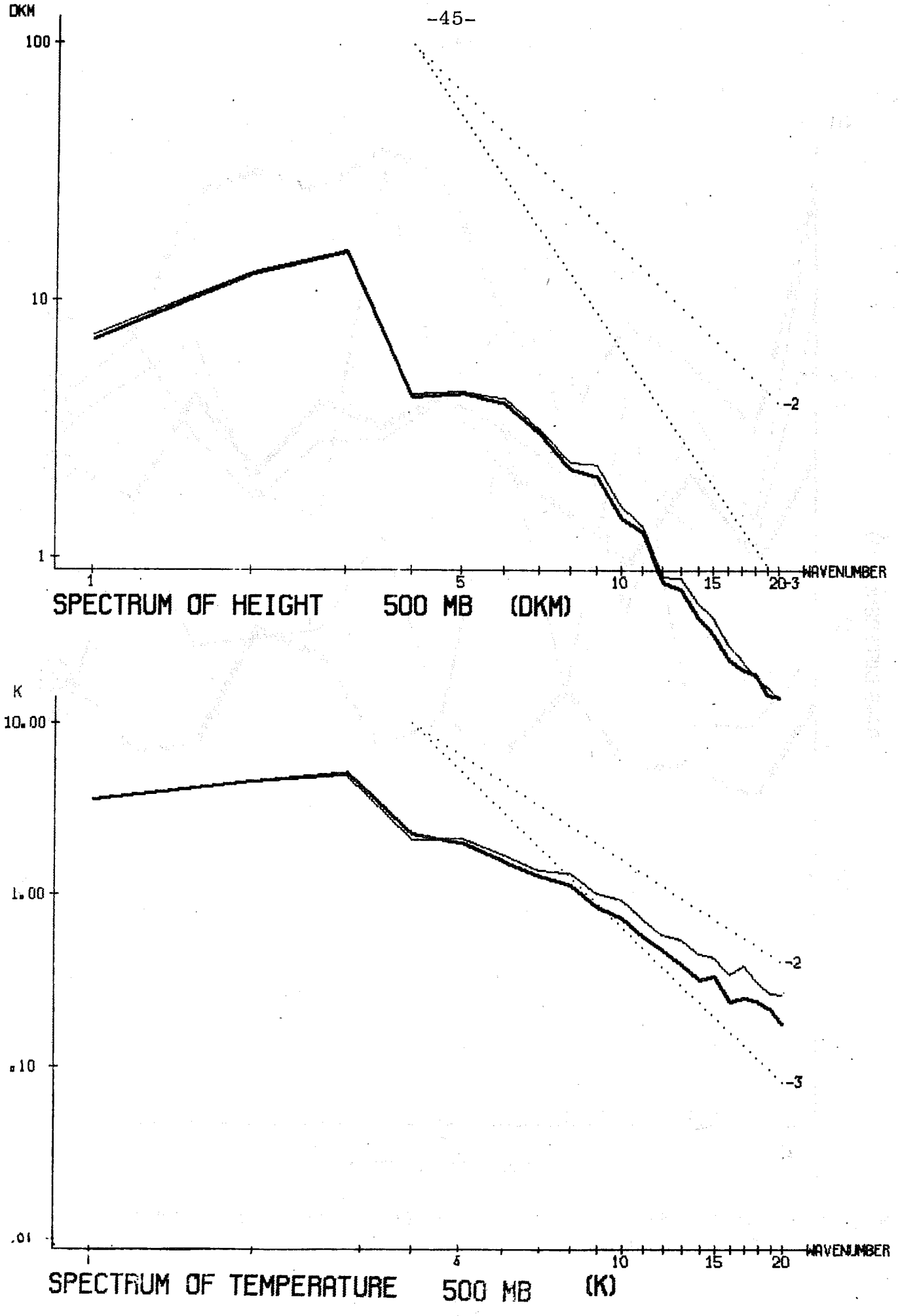


Fig. 16 Spectra of 500mb geopotential and temperature in the CONTROL (thin line) and NMC (thick line) analyses, between 40 and 60

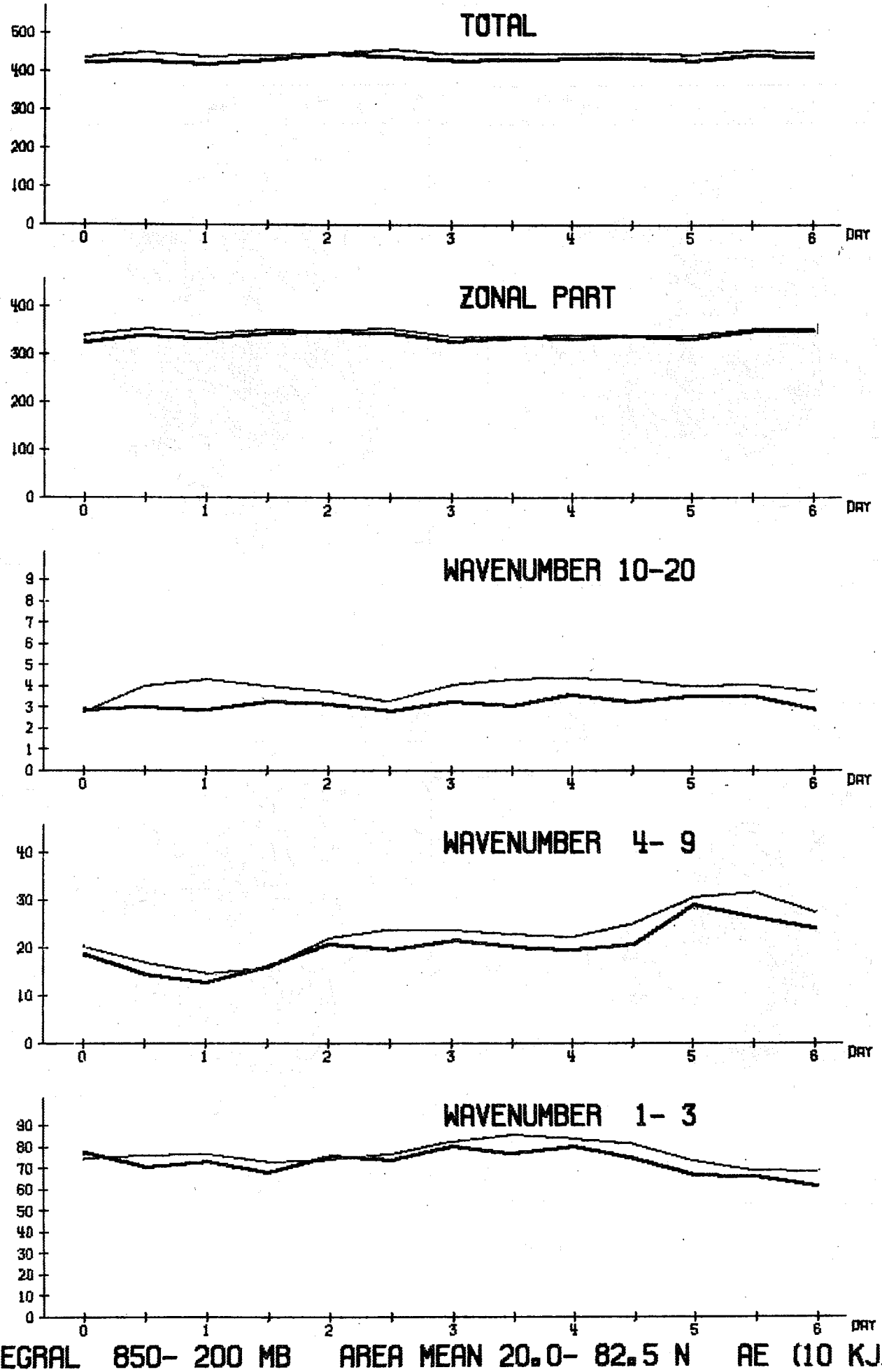


Fig. 17 Time variation in spectral bands of available potential energy in the CONTROL (thin line) and NMC (thick line) analyses. The energy was computed between 850 and 200 mb, from 20 to 82.5 degrees N

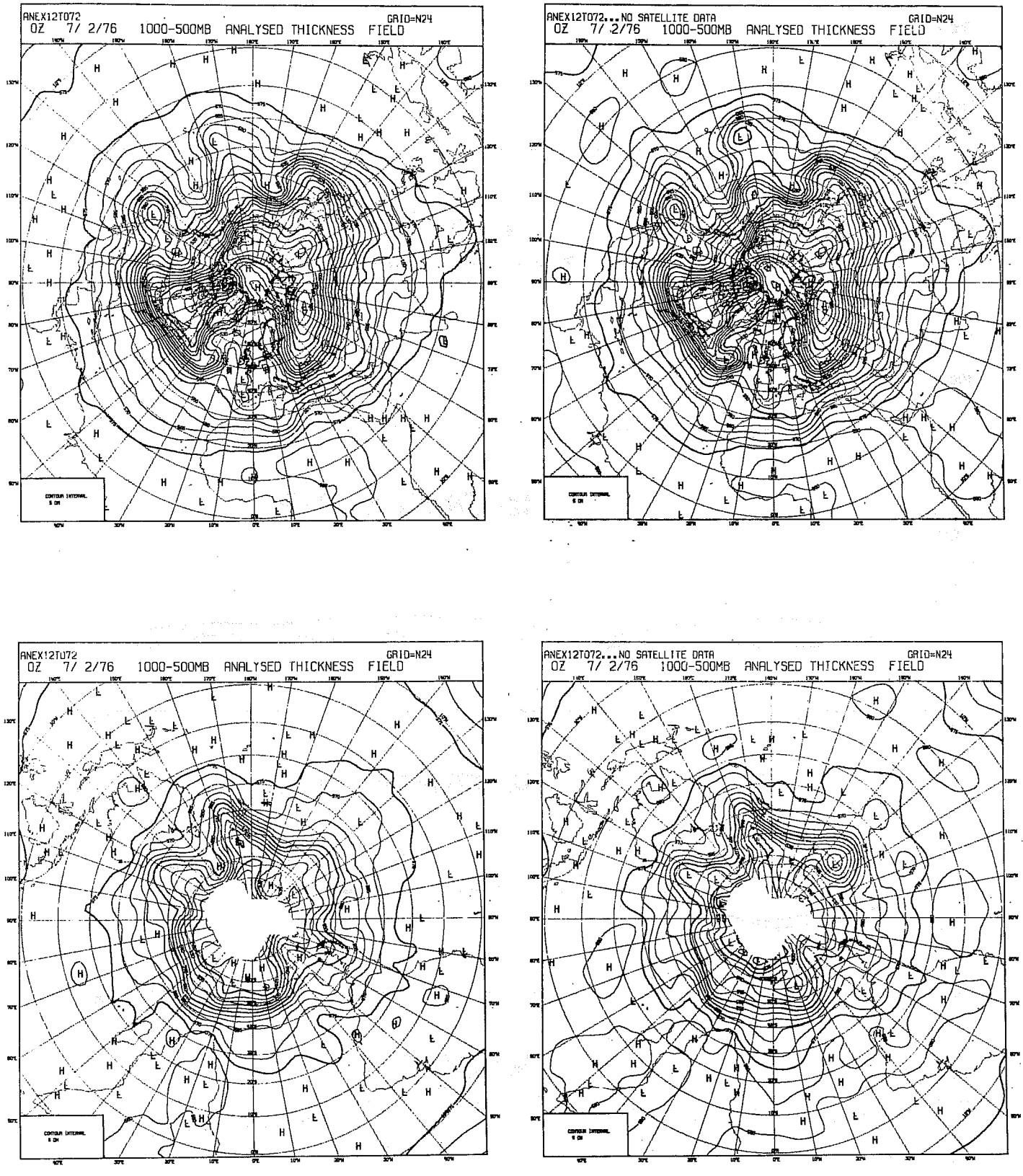


Fig. 18 CONTROL (left) and NOSAT (right) 1000-500 mb thickness analyses for 00Z, 7 February 1976.

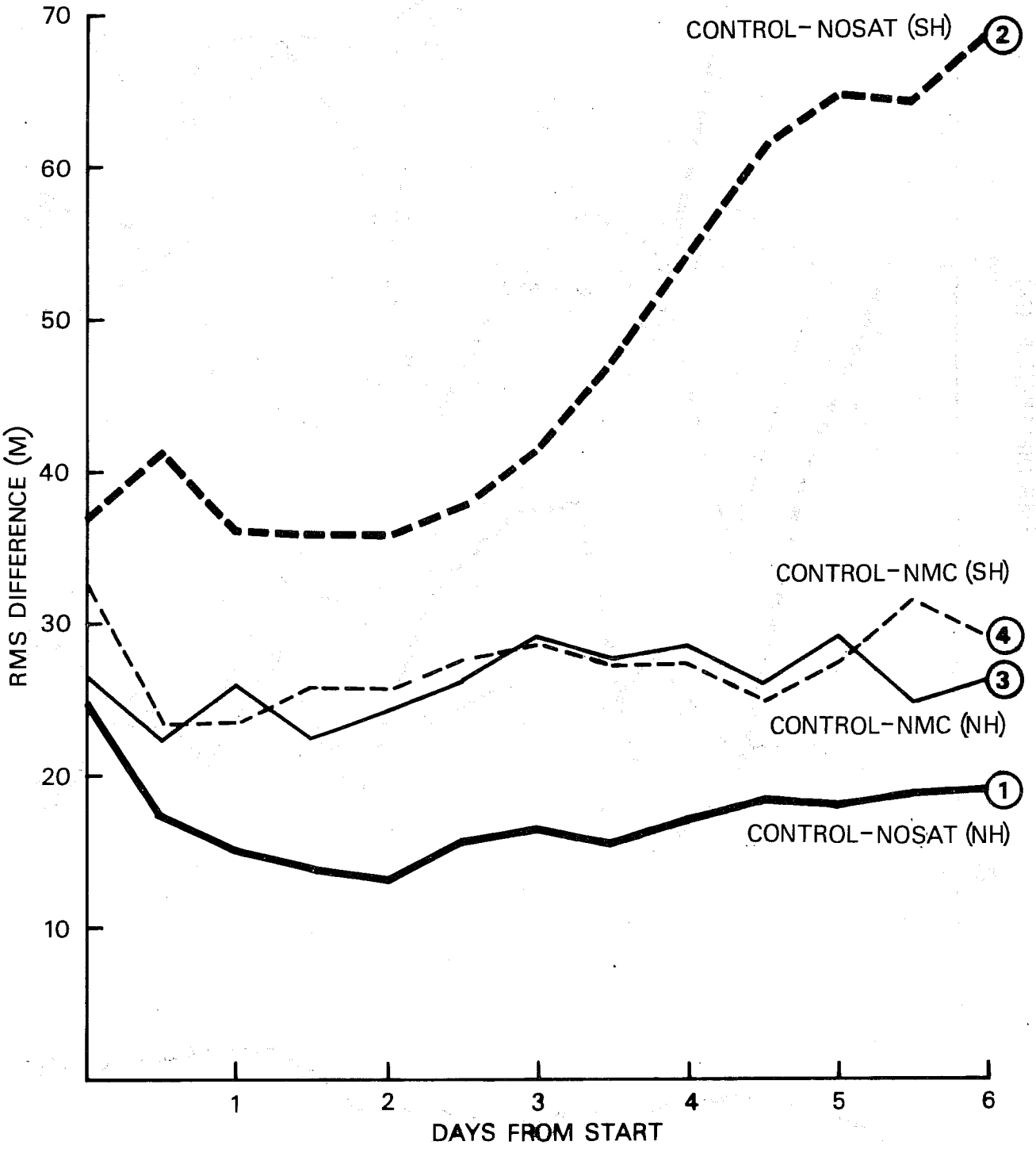


Fig. 19a RMS differences between CONTROL and NOSAT 1000-500 mb thickness analyses over (1) northern hemisphere and (2) southern hemisphere. Shown also are corresponding differences between CONTROL and NMC - (3) northern and (4) southern hemisphere.

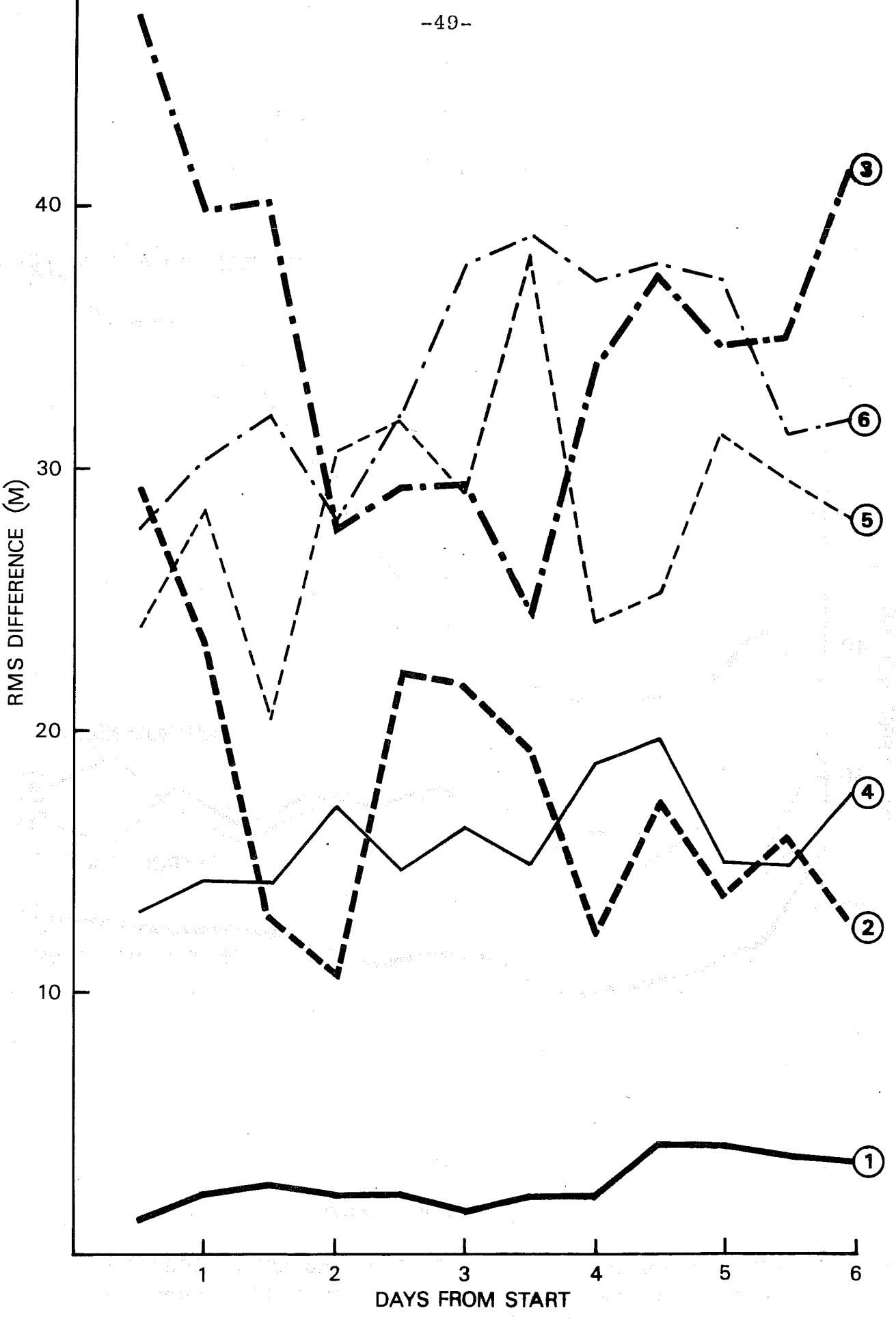


Fig. 19b RMS differences between CONTROL and NOSAT 1000-500 mb thickness analyses over (1) Europe, (2) North Atlantic, and (3) North Pacific. Shown also are the corresponding differences between CONTROL and NMC (4) Europe, (5) North Atlantic, and (6) North Pacific.

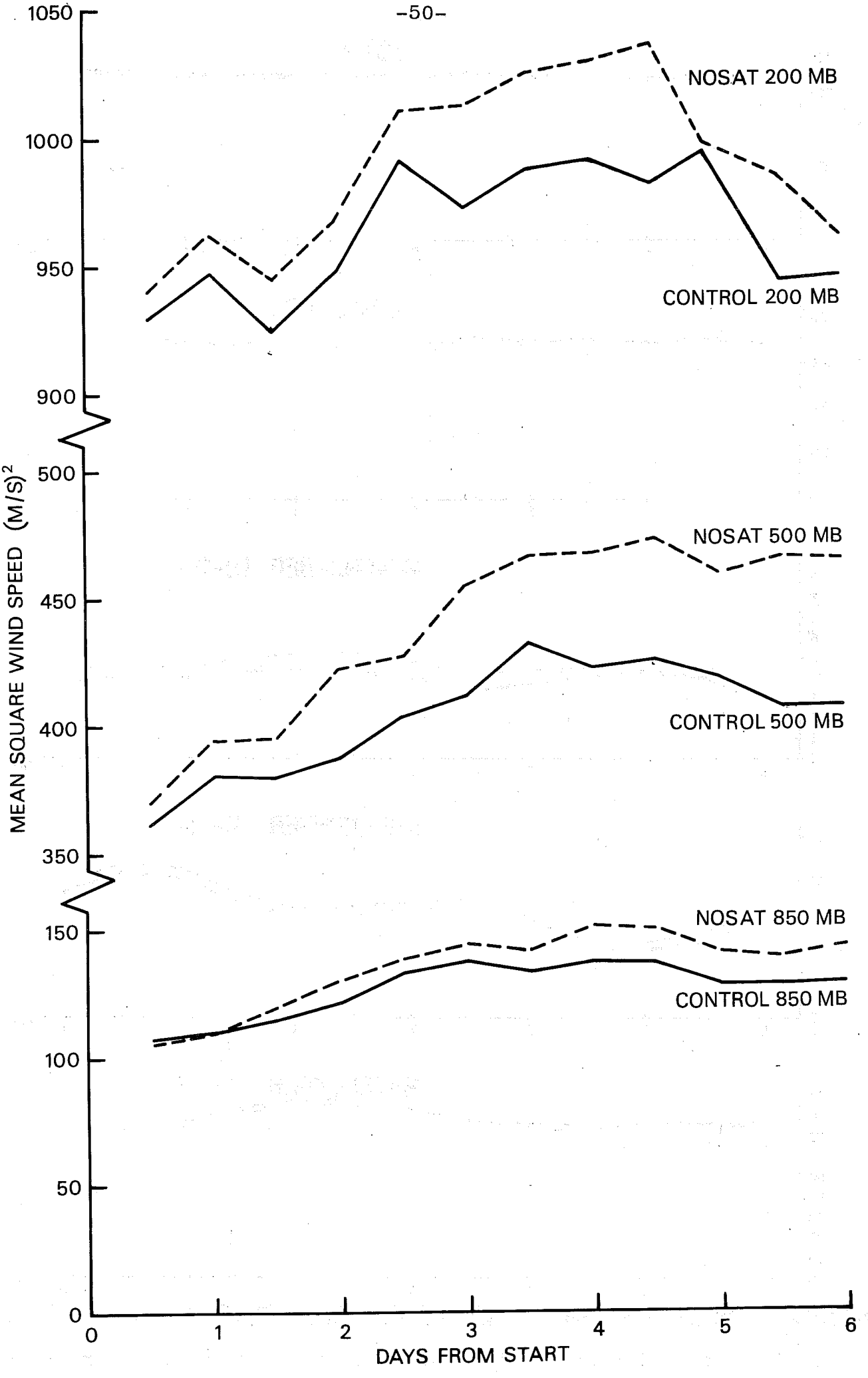


Fig. 20 Northern hemisphere mean square wind speeds for CONTROL and

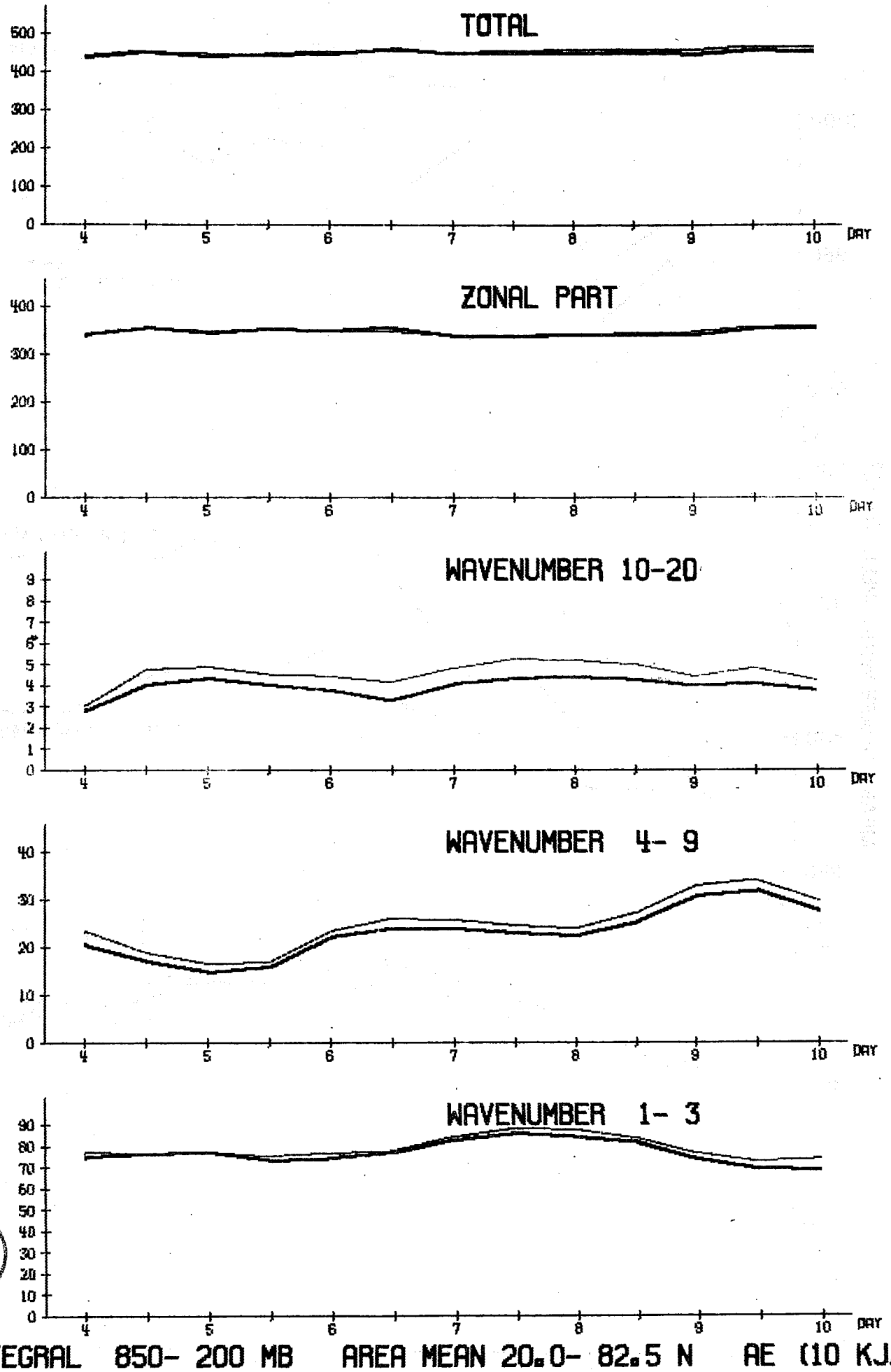


Fig.21a Time variation in spectral bands of the available potential energy in the CONTROL (thick line) and NOSAT (thin line) analyses. The energy was computed between 850 and 200 mb, from 20 to 85 degrees N

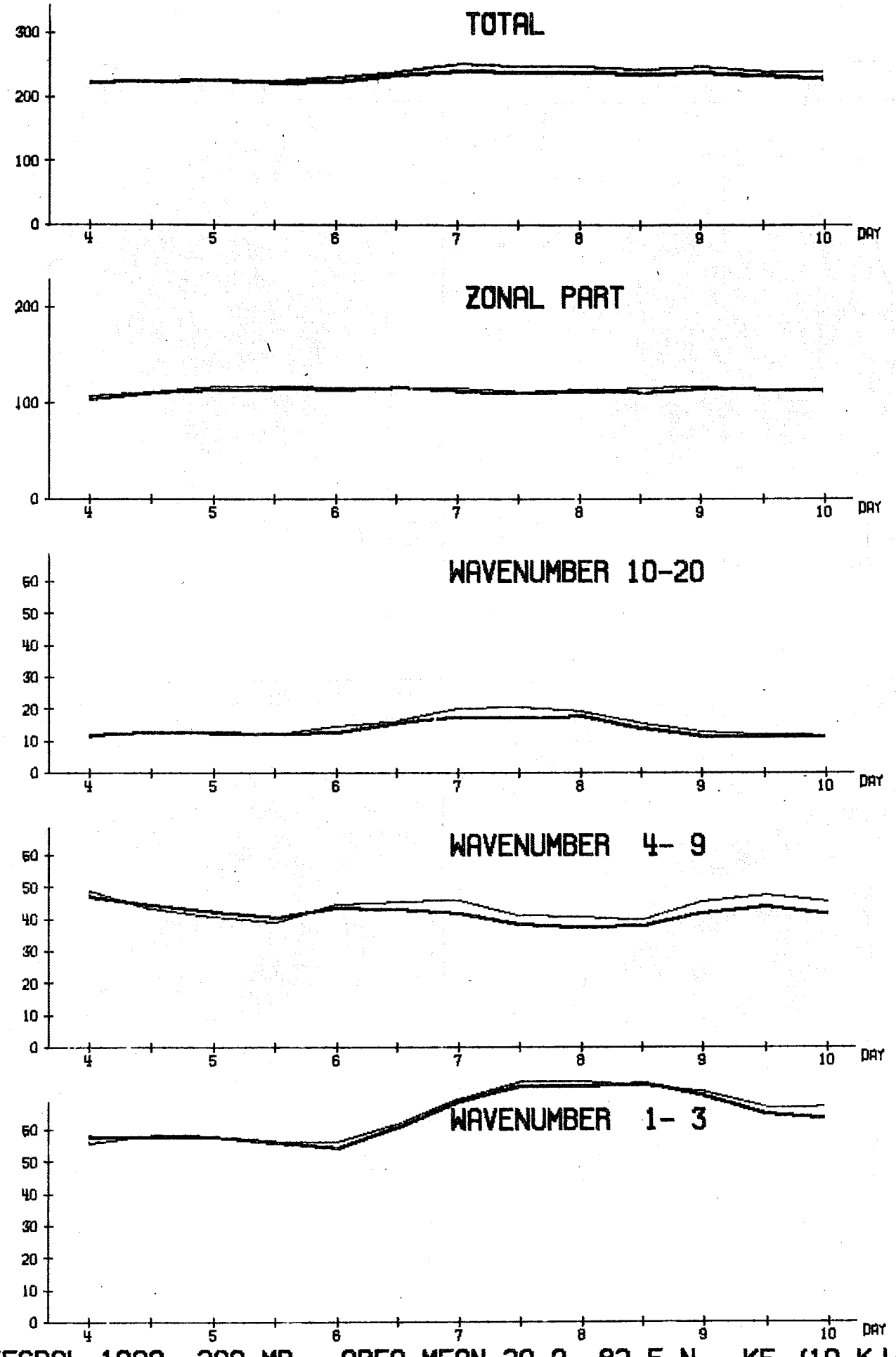


Fig.21b As in Fig. 21a, for the kinetic energy (computed from the geostrophic wind)

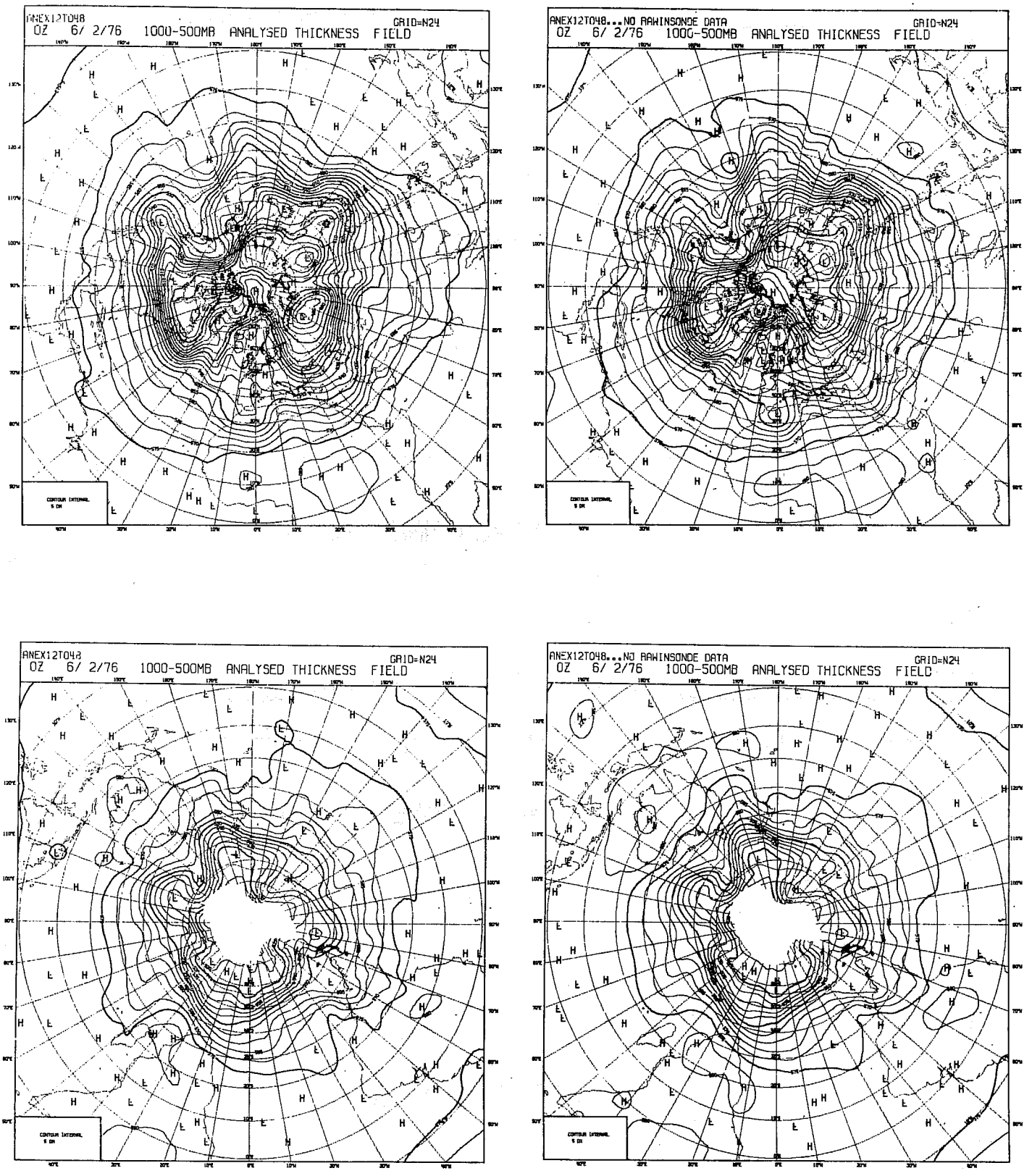


Fig. 22 CONTROL (left) and NOSONDE (right) 1000-500 mb thickness analyses for 00Z, 6 February 1976.

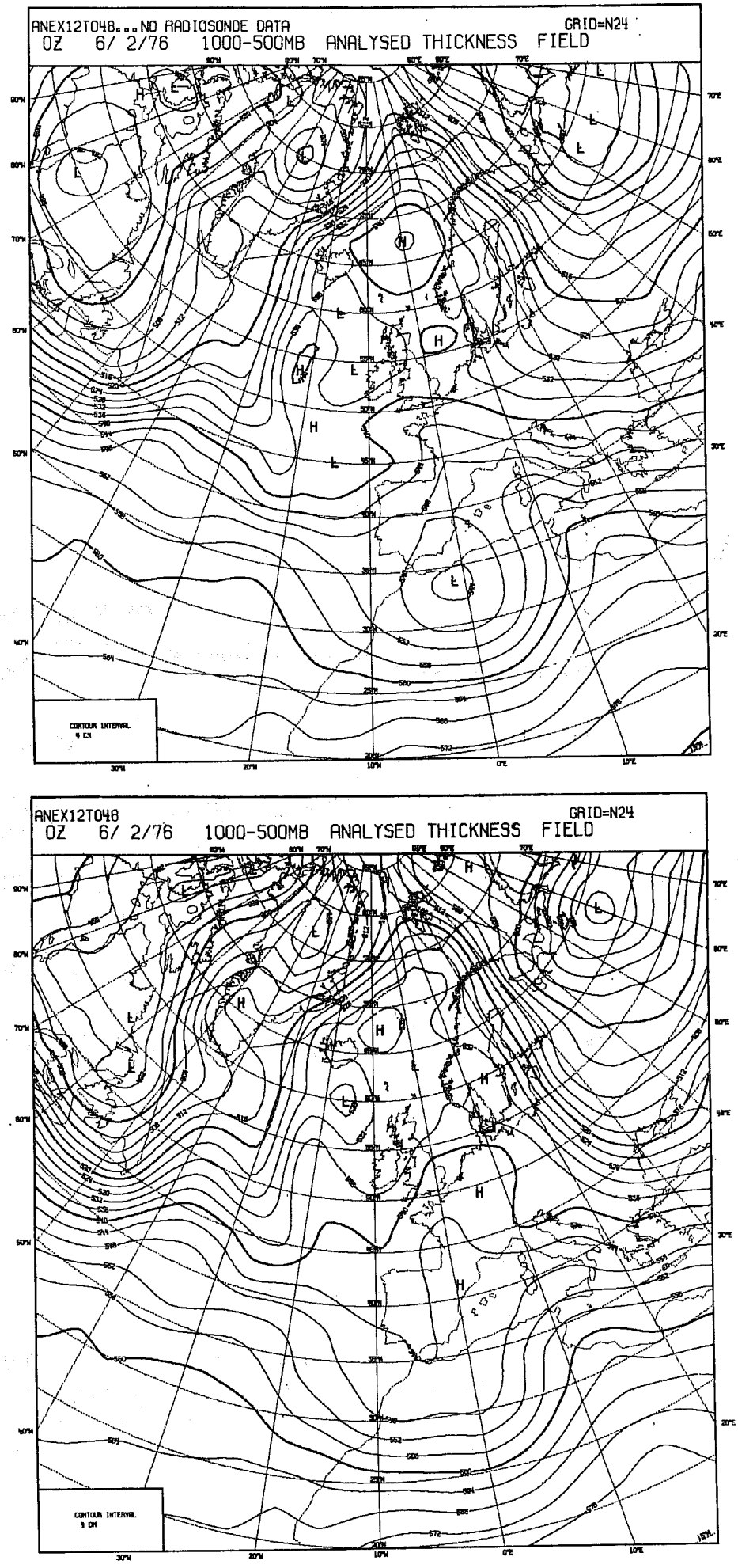


Fig. 23. CONTROL (top) and NOSONDE (bottom) 1000-500 mb thickness analyses for 00Z, 6 February 1976.

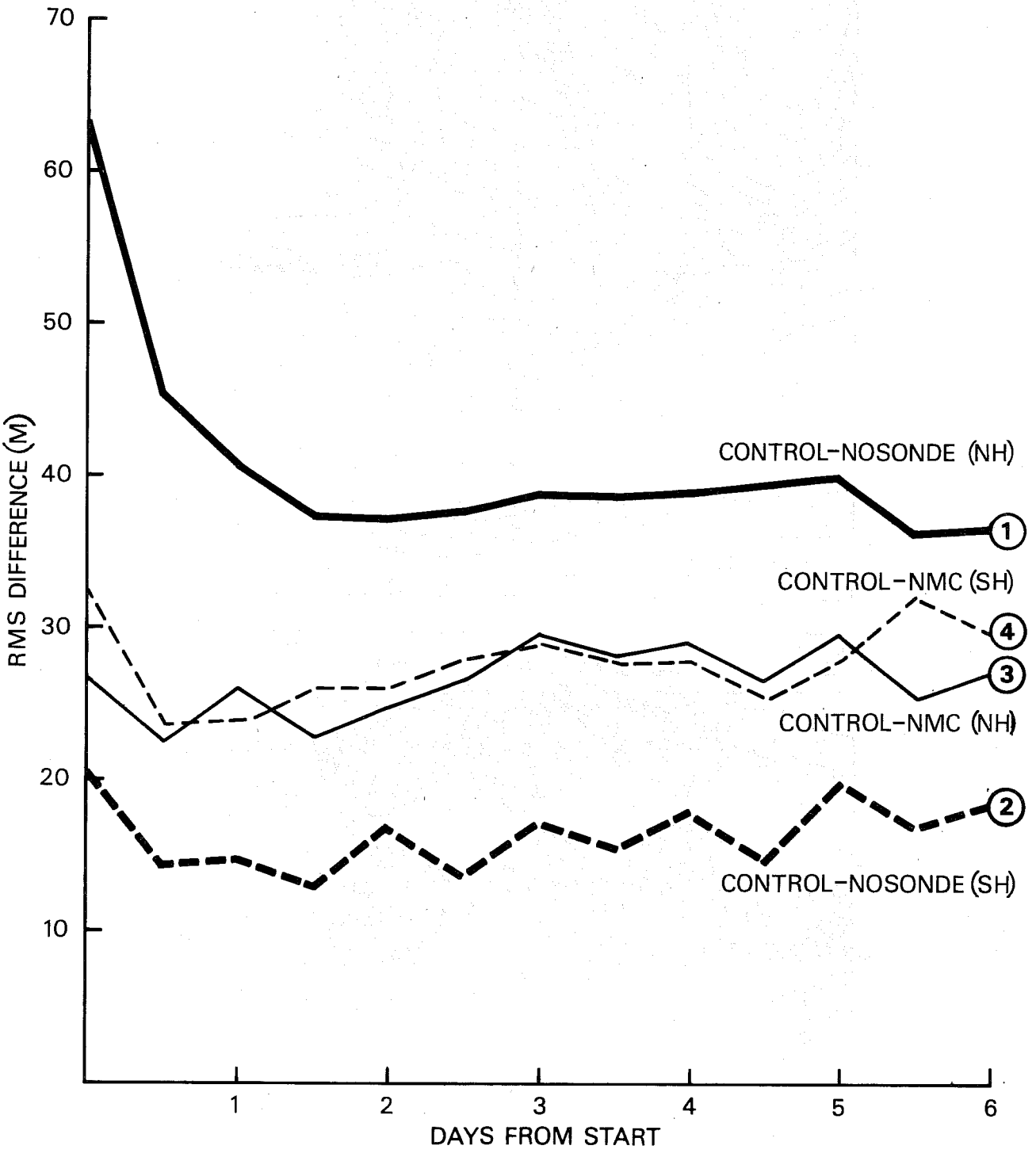


Fig. 24a RMS differences between CONTROL and NOSONDE 1000-500 mb thickness analyses over (1) northern hemisphere and (2) southern hemisphere. Shown also are the corresponding differences between CONTROL and NMC - (3) northern and (4) southern hemisphere.

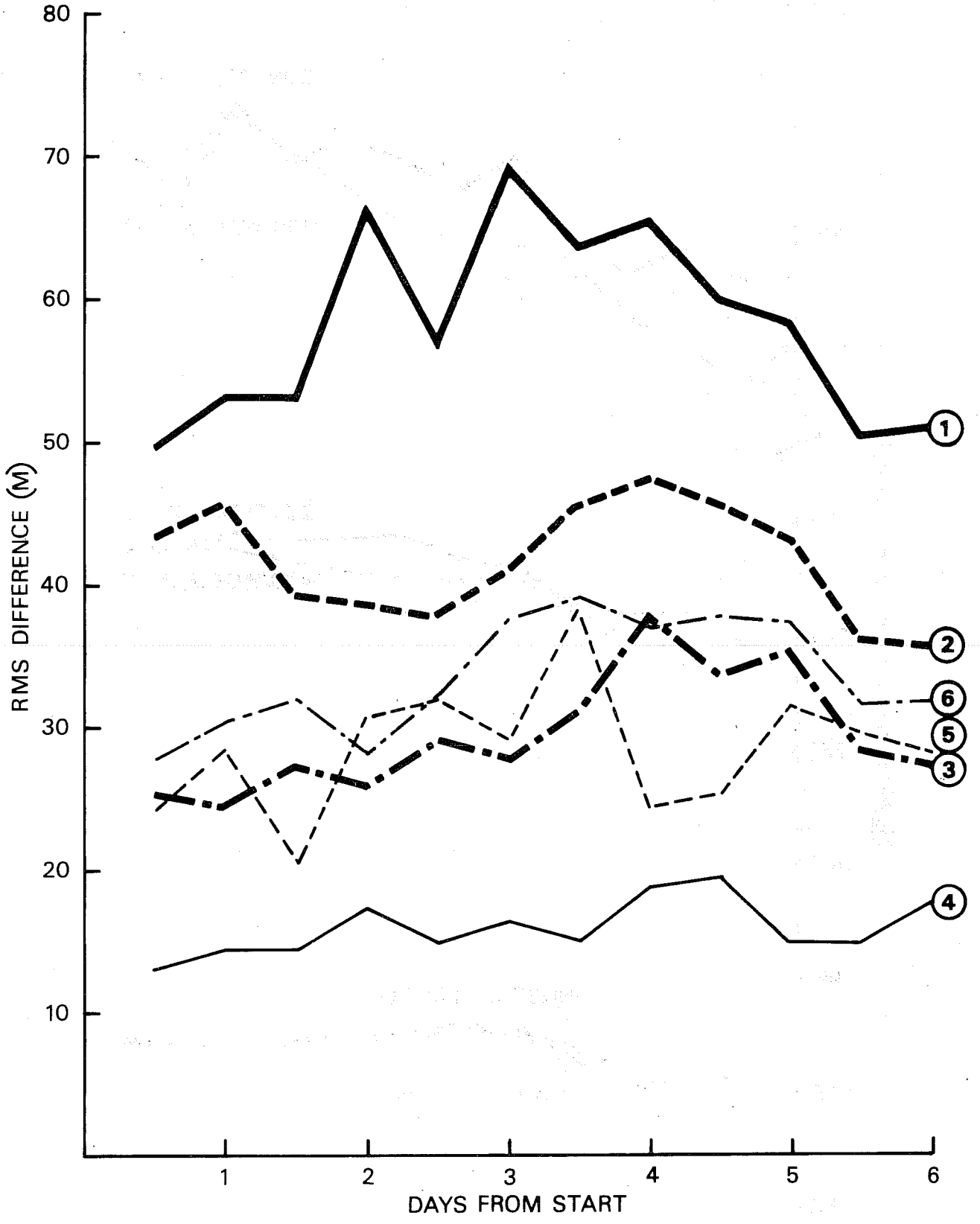


Fig. 24b RMS differences between CONTROL and NOSONDE 1000-500 mb thickness analyses over (1) Europe (2) North Atlantic and (3) North Pacific. Shown also are the corresponding differences between CONTROL and NMC - (4) Europe, (5) North Atlantic, and (6) North Pacific.

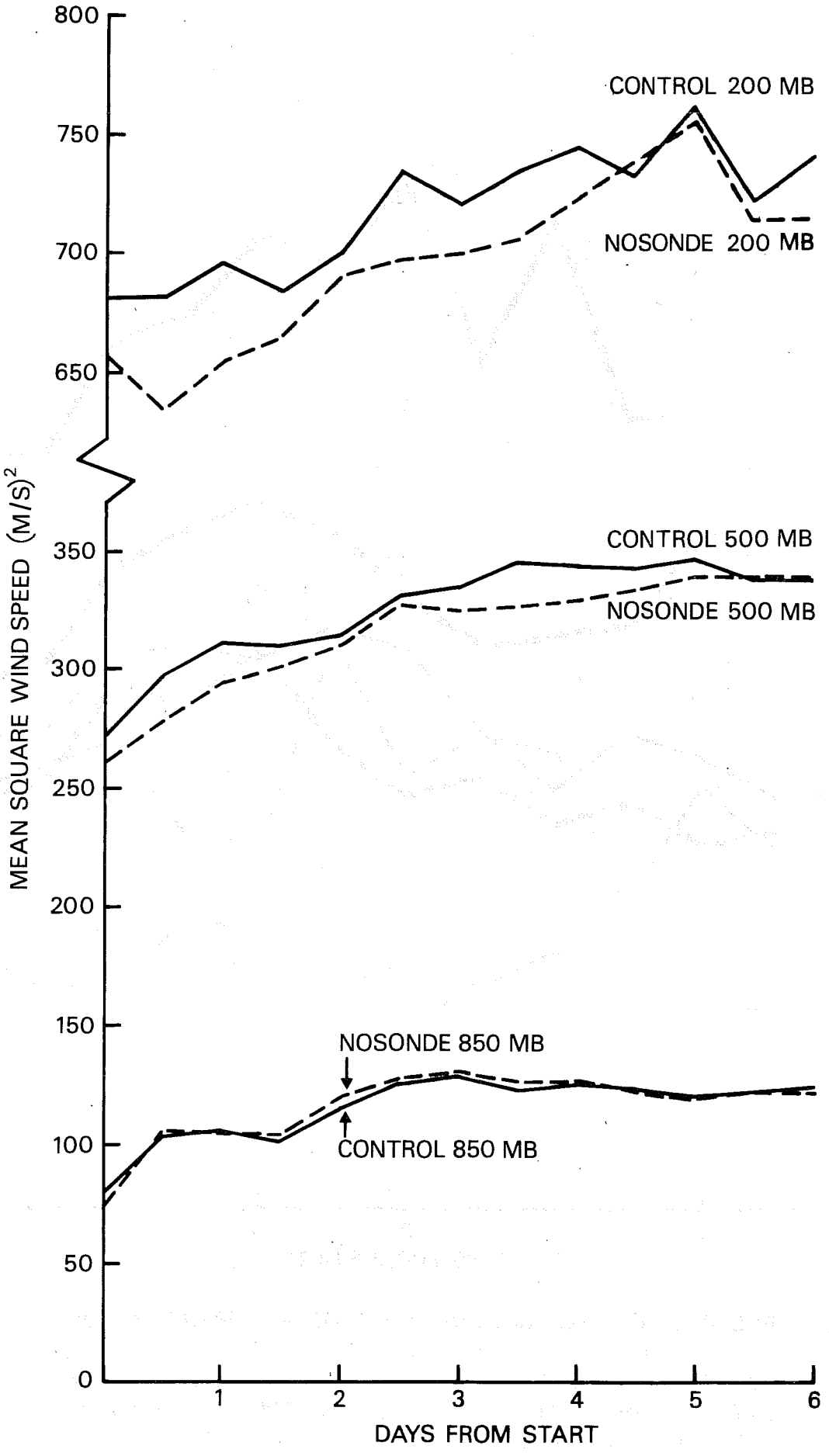
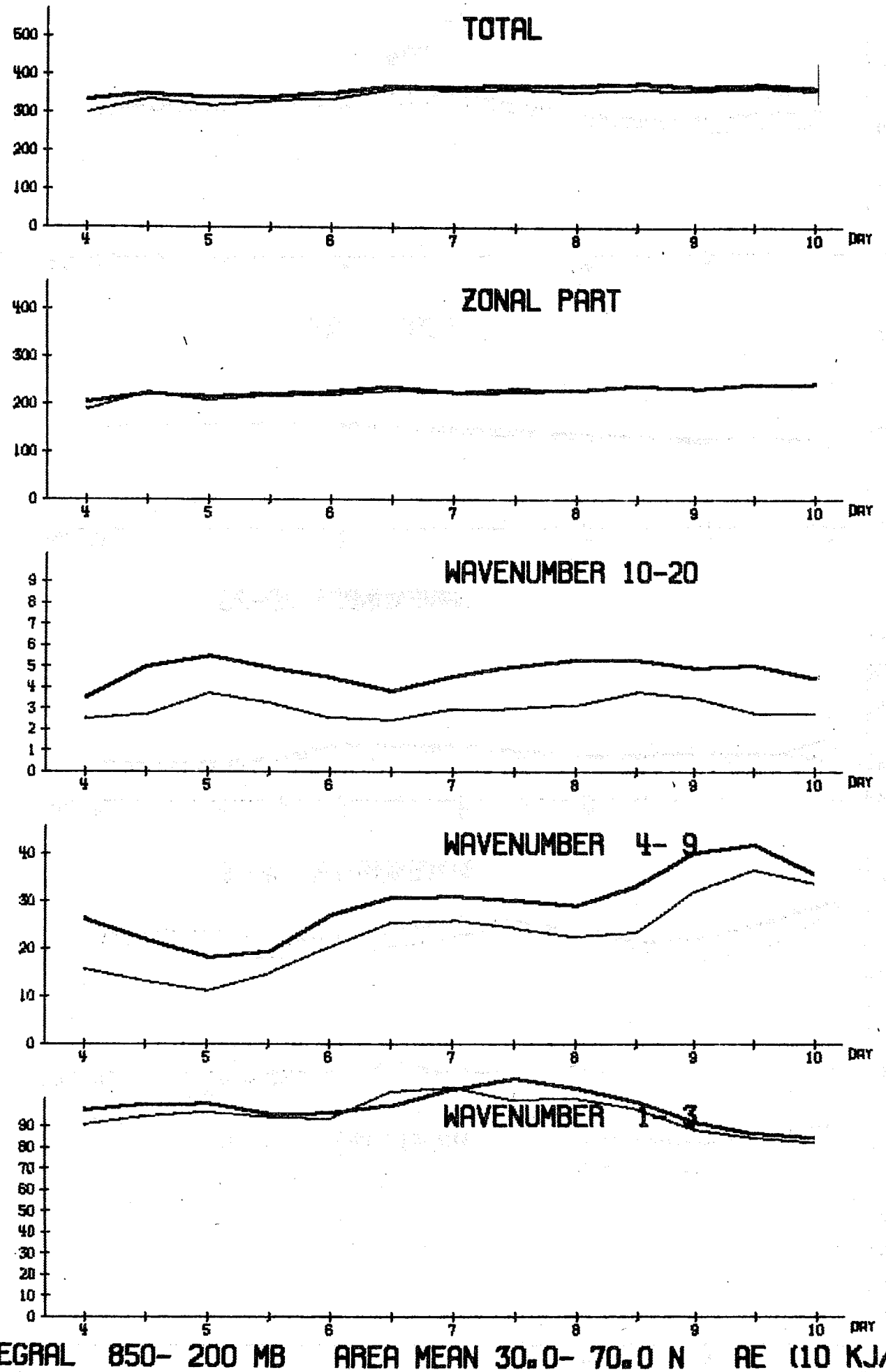


Fig. 25 Global mean square wind speeds for CONTROL and NOSONDE analyses



g.26a Time variation in spectral bands of the available potential energy in the CONTROL (thick line) and NOSONDE (thin line) analyses. The energy was computed between 850 and 200 mb, from 30 to 70 degrees N.

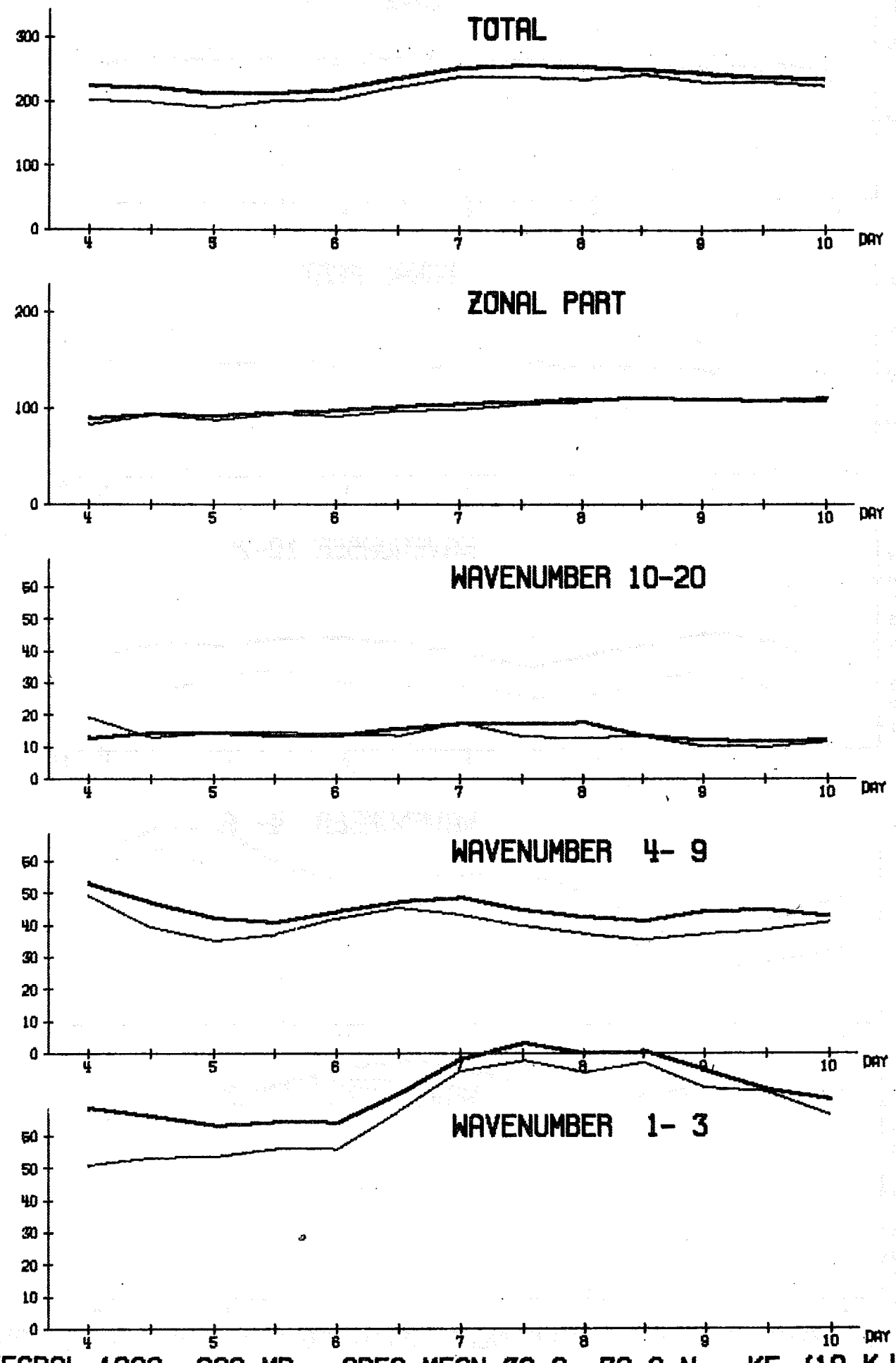


Fig. 26b As in Fig. 26a, for the kinetic energy (computed from

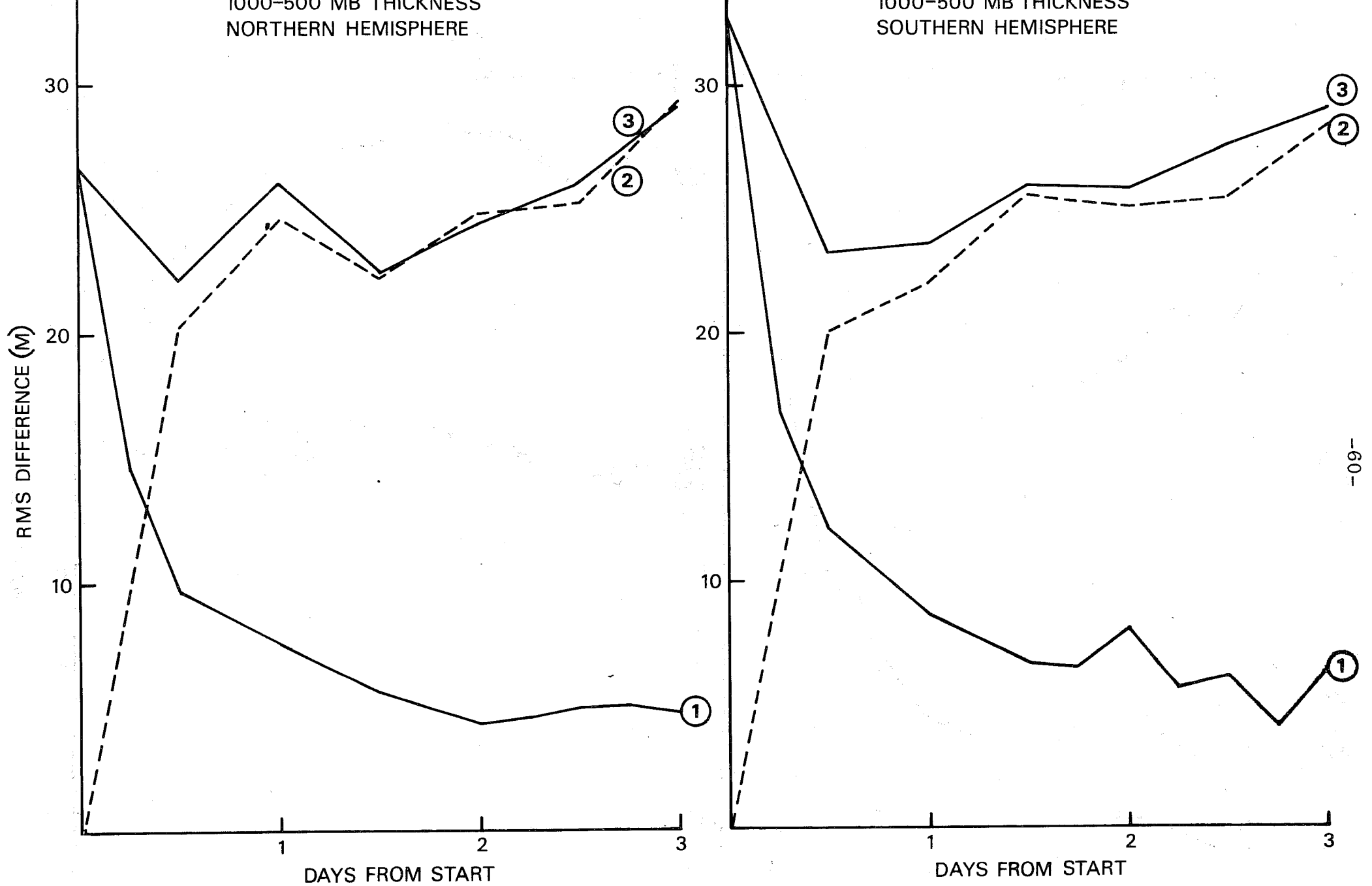


Fig. 27a 1000-500 mb thickness analysis difference between (1) CONTROL and WARMSTART, (2) NMC and WARMSTART, and (3) CONTROL and NMC. Northern hemisphere (left), southern hemisphere (right).

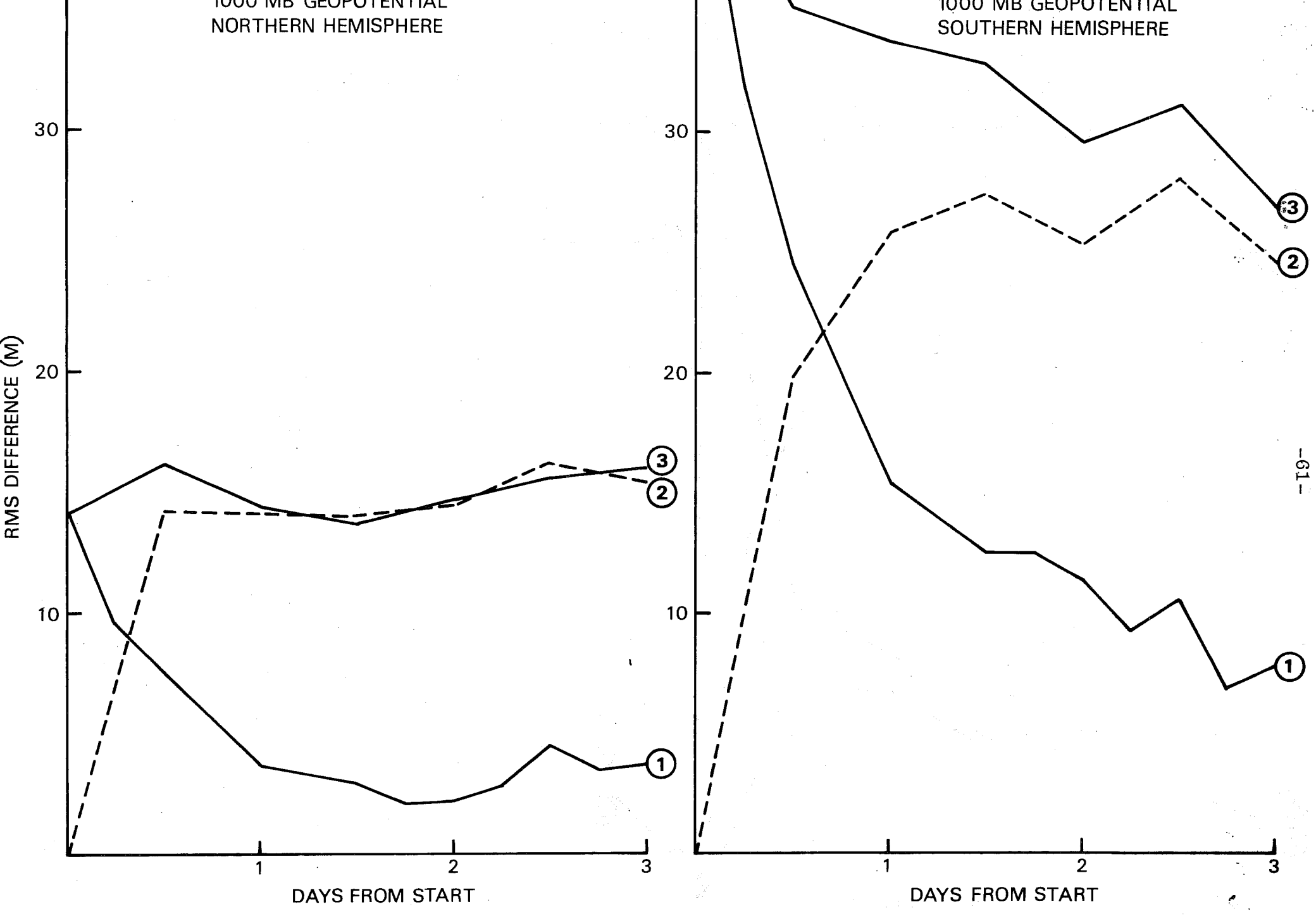
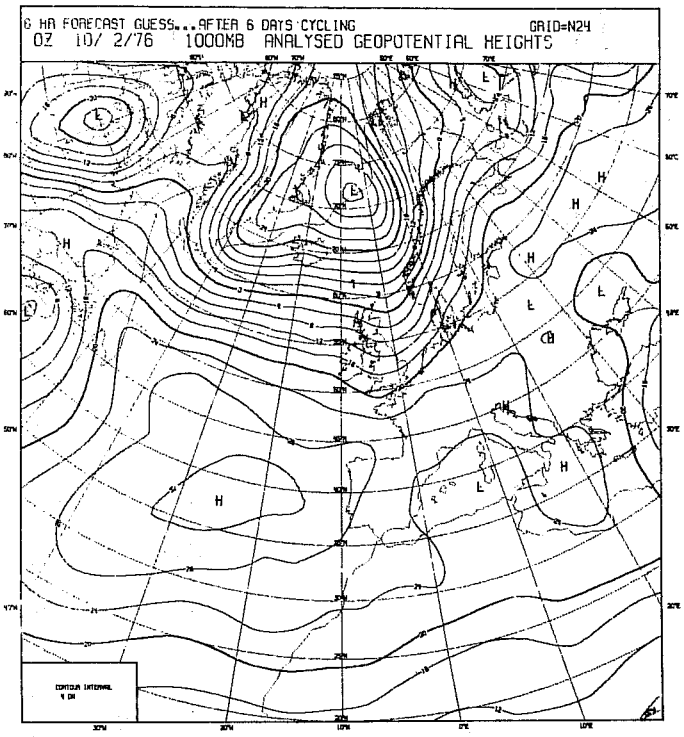
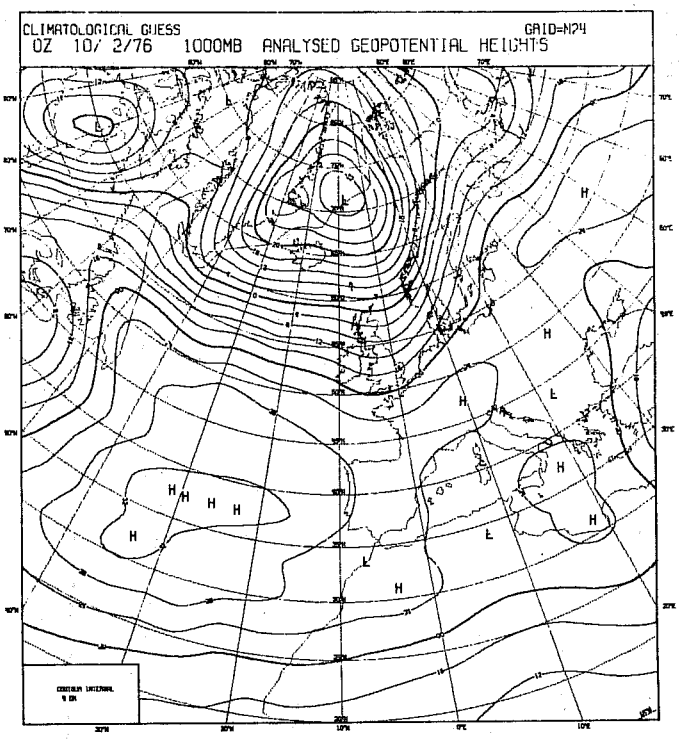
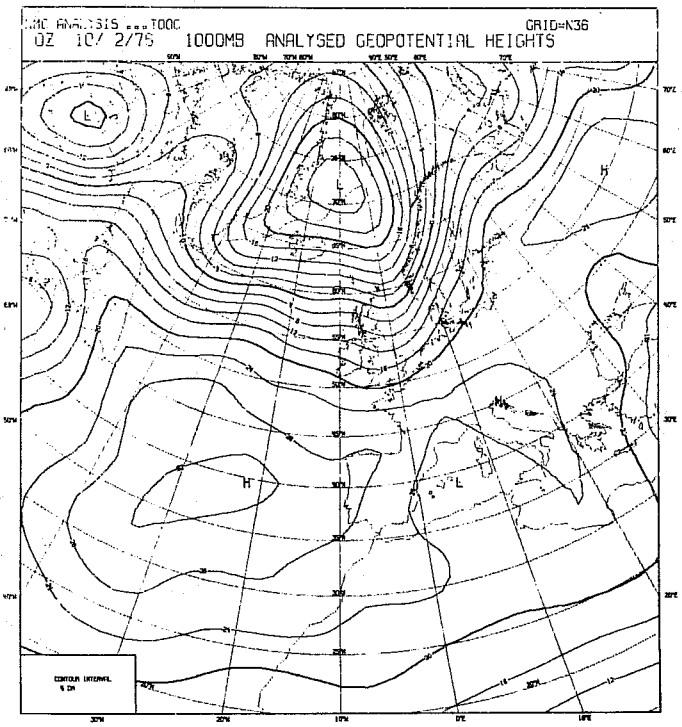
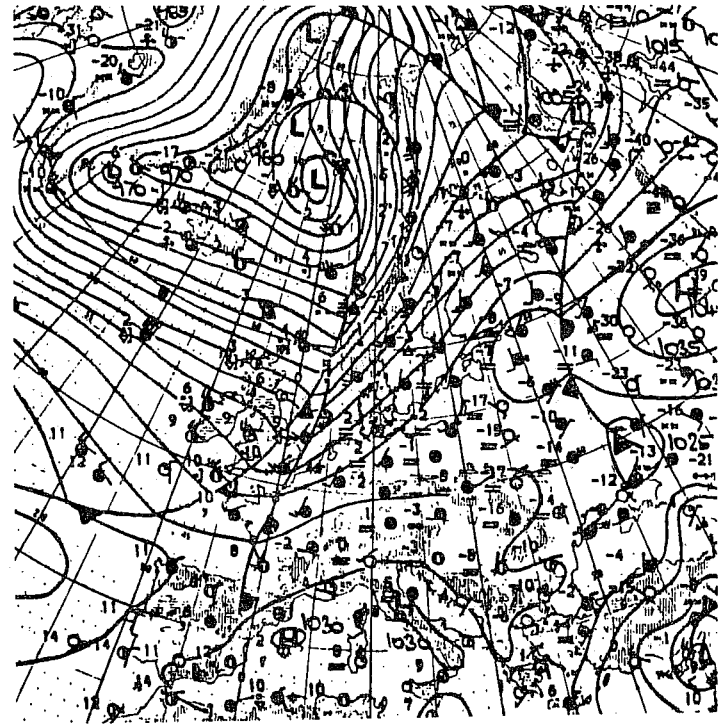


Fig. 27b. As in Fig. 27a, for 1000 mb geopotential









ig. 28 1000 mb geopotential analyses for 00Z, 10 February 1976. CONTROL (top left), TR6 climatological guess (top right), NMC (lower left), and Deutscher Wetterdienst hand analysis (lower right).

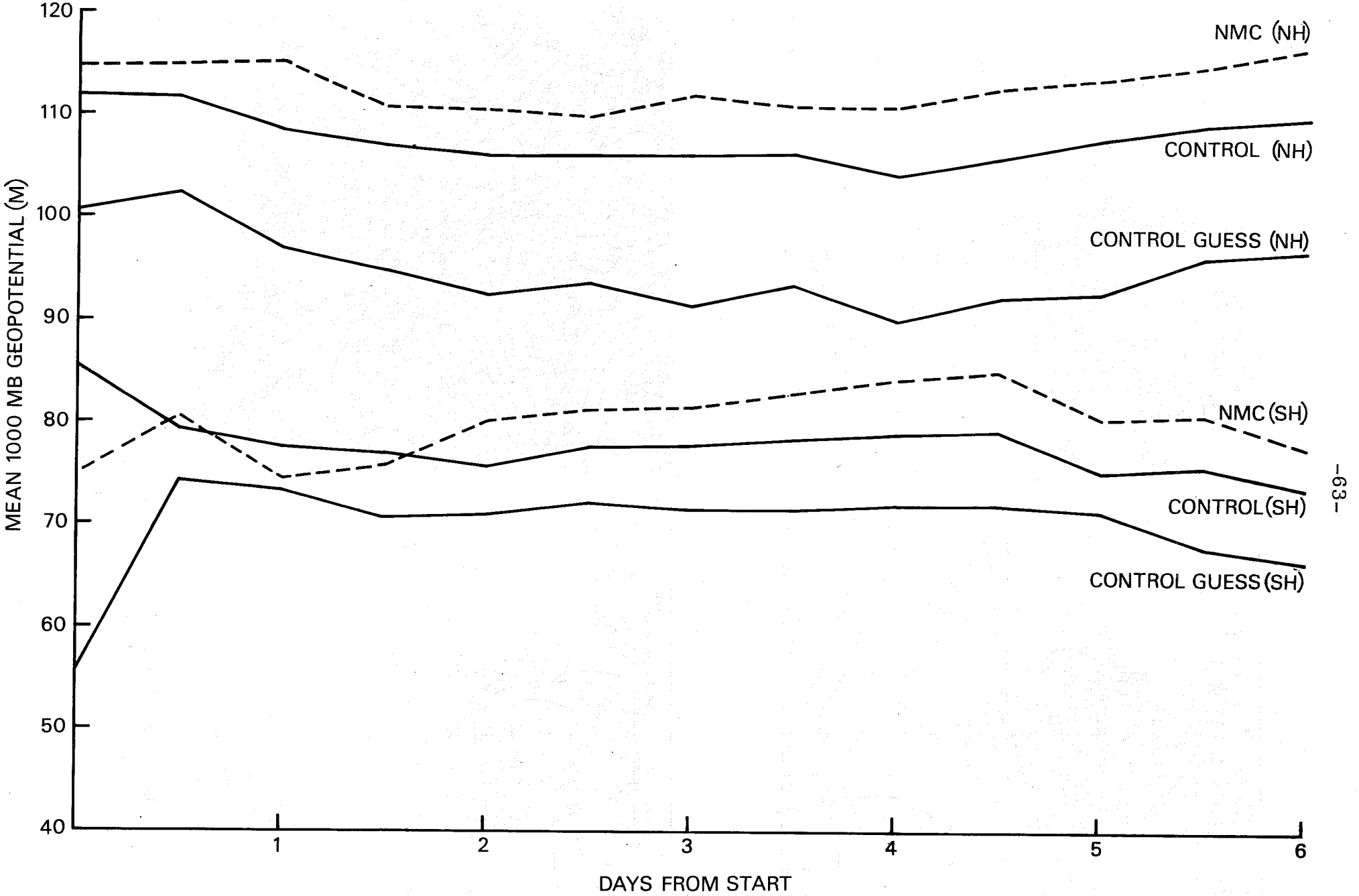


Fig. 29 Mean 1000 mb geopotentials in the CONTROL first guess, CONTROL analysis, and the NMC analysis for the northern and southern hemispheres.

References

Arpe, K. Bengtsson, L. Hollingsworth, A. and Janjic, Z.	1976	A case study of a ten day prediction. ECMWF Technical Report No.1
Burridge,D.M. and Haseler, J.	1977	A model for medium range weather forecasts - adiabatic formulation. ECMWF Technical Report No.4
Desmarais, A., Tracton, S. McPherson,R. and van Haaren,R.	1978	The NMC report on the Data Systems Test. (NASA contract S-70252-AG). US Department of Commerce, NOAA,NWS.
Hollingsworth, A. et al	1978	Comparison of medium range forecasts made with two parameterization schemes. ECMWF Technical Report No. 13 (in preparation)
Larsen,G.,Little,C. Lorenc, A. and Rutherford, I.	1977	Analysis error calculations for the FGGE. ECMWF Internal Report No. 11.
Lorenc, A. Rutherford, I. and Larsen, G.	1977	The ECMWF analysis and data assimilation scheme - analysis of mass and wind fields. ECMWF Technical Report No.6
Phillips, N.	1976	The impact of synoptic observing and analysis systems on flow pattern forecasts. Bull. Amer.Met.Soc., 57 (10), 1225-1240.
Rutherford, I.D.	1977	Interfaces between the prediction model and the analysis program in the data assimilation system; interpolation questions. ECMWF Working Paper 1/12/E/RD2/046/1977.
Temperton, C. and Williamson, D.	1978	Normal mode initialisation for a multi-level grid-point model. ECMWF Technical Report No. 11 (in preparation).

EUROPEAN CENTRE FOR MEDIUM RANGE WEATHER FORECASTS

Research Department (RD)
Technical Report No. 12

No. 1 A Case Study of a Ten Day Prediction

No. 2 The Effect of Arithmetic Precision on some Meteorological Integrations

No. 3 Mixed-Radix Fast Fourier Transforms without Reordering

No. 4 A Model for Medium-Range Weather Forecasting -Adiabatic Formulation-

No. 5 A Study of some Parameterizations of Sub-Grid Processes in a Baroclinic Wave in a Two-Dimensional Model

No. 6 The ECMWF Analysis and Data Assimilation Scheme - Analysis of Mass and Wind Fields

No. 7 A Ten Day High Resolution Non-Adiabatic Spectral Integration : A Comparative Study

No. 8 On the Asymptotic Behaviour of simple Stochastic - Dynamic Systems

No. 9 On Balance Requirements as initial conditions

No. 10 ECMWF Model - Parameterization of Sub-Grid Scale Processes

No. 11 Normal mode initialization for a multi-level grid point model

No. 12 Data Assimilation Experiments