COMPARISON OF MEDIUM RANGE FORECASTS MADE WITH TWO PARAMETERIZATION SCHEMES

BY

A. Hollingsworth, K. Arpe, M. Tiedtke, M. Capaldo,
H. Savijärvi, O. Åkesson, J.A. Woods

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

Page No.

Abstract

We consider the results of a series of forecasts on seven weather situations from February 1976 using two models which differ only in their physical parameterization processes. One set of parameterizations was developed at the Geophysical Fluid Dynamics Laboratory (GFDL) some years ago, the other at ECMWF more recently. The resolution of the models (N48, 15 levels) was fairly close to that which ECMWF may use in the first phase of operations beginning July 1979. The particular aim of the experiments was to study the importance of the differences in the parameterization schemes for the models, but a general view of features of the forecast results that might be available in the first phase of operations was also obtained.

Both models gave similar results in terms of forecast quality. When measured by the standard objective methods, the range of predictability seemed to be about 5 or 6 days; however a subjective synoptic evaluation could sometimes find useful indications as to the evolution beyond this time. A study of the systematic errors in the forecasts showed that these were mainly associated with a loss of energy in the largest waves; the evolution of these systematic errors appeared to be roughly linear in time. This is not to say that the systematic errors arise due to linear mechanisms. Regarding transient phenomena, the downstream intensification of baroclinic waves appeared sometimes to be predictable nine days in advance. A study of the energetics showed that the major part of the loss of energy in the long waves was due to a failure to maintain the stationary part of the long wave energy.

1. Introduction and summary

In the last decade there have been considerable efforts to determine the limit of useful predictability of large-scale atmospheric flow.

The work of many investigators (Miyakoda et al 1969, 1972, Druyan et al 1975, Baumhefner and Downey 1978) indicate that this limit is larger than four days and probably less than ten days. Many factors can affect the success of the forecasts, for example the quality and coverage of the data, the sophistication of the analysis system, the type of finite difference (or spectral) scheme, the resolution of the grid and the formulation of the physical parameterizations.

ECMWF was set up to try to attain the limits of forecast skill. Routine forecasting will begin at ECMWF in mid-1979. In the design of this forecasting system it was easy to make choices about some of the factors listed above and considerably more difficult to make choices about others.

Given the multiplicity of data sources and the necessity to analyse in mid-latitudes and in the tropics it was felt that a statistical interpolation analysis scheme and an intermittent data assimilation scheme were essential (Gandin 1963, Schlatter 1975, Rutherford 1976, Lorenc et al 1978). To make the forecast routinely in a reasonable time with a reasonable resolution, required the use of a semi implicit time scheme (Robert et al 1972, Burridge and Haseler 1979).

A staggered grid is economical when used with a semi implicit time scheme and so such a grid was chosen. The resolution is set by the available computer power and one should maintain a balance between horizontal and vertical resolution. The choice of parameterization package is not as clear cut as some of the other choices. We took the parameterization package of the 1969 GFDL model (Miyakoda et al 1969) as a bench mark against which to compare the performance of a new package which relaxed some of the restrictions of the GFDL package and treated some of the processes in a different manner (Tiedtke et al 1979).

The experiments reported here are the first extensive tests of the Centre's forecast model in a resolution close to that which will be used operationally. The experiments compare the performance of the two parameterization schemes in a set of seven ten-day forecasts from February 1976. The results show that with both parameterization schemes we can make useful forecasts in these winter cases for five or six days on average. Depending on the individual case the period of usefulness may be longer or shorter.

The Models

The main features of the model are that an enstrophy conserving finite difference scheme was used, and a semi-implicit time-stepping scheme. The basic grid was a latitude/longitude system with grid lengths $\Delta\lambda=\Delta\theta=1.875^{\circ}$. There were 15 vertical levels, distributed as in Table 2.1., so as to give sufficient resolution in the troposphere, near the lower boundary, and near the tropopause.

Initial data for the experiments

The initial data were taken from the global analyses for February 1976 made by NMC Washington during the "Data Systems Test" (DST) Experiment. This was supplemented by 10 mb analyses for the Northern Hemisphere available from the United Kingdom Meteorological Office and climatological 10 mb data for the Southern Hemisphere. Sea surface temperatures were taken from February normals; (Alexander and Mobley 1974); the topography used in these experiments was the field originally prepared by Berkofsky and Bertoni (1955). The data were initialized using the technique of non-linear normal mode initialization.

Initial dates for the experiments were 3rd, 6th, 9th, 12th, 15th, 18th and 22 February 1976.

Synoptic evaluation

By and large the synoptic evaluation of the experiments indicated that five to six days was about the time for which the forecasts were directly useful over Europe and the Atlantic. However useful information as to the general synoptic trends was sometimes available beyond this. In particular, the phenomenon of downstream intensification of baroclinic waves was forecast successfully to 8 or 9 days, on occasion. Overall it was noted that the quality varied from one forecast to another, and, for any given forecast, from one region to another. There was little to choose between the forecast quality of the models.

Objective verification

The scores used for objective verification were root mean square (RMS) errors of height and temperature and anomaly correlation coefficient of height and temperature for the troposphere north of 20 $^{\circ}$N. The anomaly correlation coefficient is the correlation between forecast and observation after the monthly normal has been subtracted. This measure has quite different properties from the tendency correlation, where the initial field is subtracted. The tendency correlation is used for studying short range forecasts while the anomaly correlation is more suitable for the study of medium range forecasts. RMS errors of height and wind, tendency correlation and S_1 skill score for an area covering Europe were also calculated.

In the forecasts, the average anomaly correlation coefficient for the troposphere north of 20 $^{\circ}$N for the whole field and for wavenumbers 1 - 3 and 4 - 9 fell to 0.6 between five and six days. The rms errors for the total field and for the wavenumbers 1 - 3 each reached the level of the climatological variance at about the same time. At 500 mb the scores reached these levels also between five and six days, but they reached them about one day sooner at the 1000 mb level.

Systematic errors

By systematic errors is meant the difference between the average of the nth day of the forecasts and the corresponding analyses. The evolution of these errors can be studied as n increases. For sufficiently large ensembles of forecasts one can speak of the average of the forecasts as the "day n climatologies"; however the present ensemble of seven cases is too small to justify the use of such a term.

Even though the ensemble is small the averaged analyses for the seven cases showed a good deal of structural similarity to the monthly normal. Striking differences between forecast and analysis were the displacement of the Icelandic low into the Norwegian Sea and the over-intensification of the Aleutian low. The Siberian high was well maintained. Results for 500 mb had a similar character. The average difference fields show energy in the longest waves, mainly in wavenumber 2. The growth of errors appeared to be approximately linear in time. The causes of the errors are not clear. The linear behaviour, with time, suggests a shortcoming in some forcing effect such as the interaction with land-sea differences or topography. On the other hand the position of the maximum errors, on the eastern sides of the oceans, suggests that the over-development of surface lows, particularly over the oceans (which is a feature of the forecasts) may make an important contribution to the systematic errors. This in turn is probably associated with an incorrect distribution of latent heat release with height.

Energetics

For the purposes of discussion the average of the energy (kinetic or available) in the forecasts is referred to as the total energy, the energy in the ensemble mean state as the "S energy" (for "should be stationary") and the difference between the two as "T energy" (for transient).

As the forecasts progressed there was a steady growth in the energy of the zonal flow, a decline in the total long wave energy (waves 1 - 3) and a growth in the energy of the intermediate waves (waves 4 - 9). The spectrum between wavenumbers 10 - 20 appeared to follow a k^{-3} law.

The drop in the total long wave energy seems to be due to a fall in the S energy. Indeed the S part of the flow is a major contributor to the available potential energy and to the energy conversion terms in this wave band. The level of energy in the baroclinic wave band became rather too high in the course of the forecasts; this was particularly marked at low levels, where the disturbances tended to overdevelop.

Differences between the models

The differences between the models grew slowly in the first five days and more rapidly thereafter. If one model were used as a predictor for the other it would be quite good up to five or six days but it would be rather poor on average by the tenth forecast day.

2. Finite difference scheme and resolution

The finite difference scheme and the arrangement of variables in the horizontal have been described in Burridge and Haseler (1977). The horizontal arrangement of variables is known as the "C" grid (Arakawa and Lamb, 1977). The horizontal scheme was the enstrophy conserving scheme of Sadourny (1975). We had earlier used an enstrophy-energy conserving scheme also due to Sadourny and also described in Burridge and Haseler (1977) but this proved to have a serious deficiency (Hollingsworth and Kållberg, 1979). Both schemes were based on the advective form of the equations. The horizontal grid was a regular latitude-longitude grid with a grid length of 1.875° in each direction (N48). The vertical coordinate was the sigma coordinate of Phillips (1957). All variables were carried at the main levels except ϕ and $\dot{\sigma}$ which were calculated at the half-levels. ϕ was then averaged to calculate a value for the pressure gradient calculation. The vertical differencing is a little different from that described in Burridge and Haseler (1977). The hydrostatic equation was written in the form

$$\phi_{k+\frac{1}{2}} - \phi_{k-\frac{1}{2}} = -RT_{vk}(\ln\sigma_{k+\frac{1}{2}} - \ln\sigma_{k-\frac{1}{2}}),\ k = 2....NLEV$$

except for the top level where we had

$$\phi_{1\frac{1}{2}} - \phi_1 = -RT_{v1}(\ln\sigma_{1\frac{1}{2}} - \ln\sigma_1)$$

where T_v is virtual temperature.

Energy conservation was maintained by requiring that the term $\frac{1}{\sigma}$ in the expression from the energy conversion term in the thermodynamic equation be written in the form

$$\left(\frac{1}{\sigma}\right)_k \to \frac{\ln\sigma_{k+\frac{1}{2}} - \ln\sigma_{k-\frac{1}{2}}}{\sigma_{k+\frac{1}{2}} - \sigma_{k-\frac{1}{2}}}$$

The spacing of the levels in the vertical is non uniform.

With an N48 resolution we had enough computer capacity to treat a 15 level model. We wanted to have a good resolution of the boundary layer with a view, eventually, to exploring the importance of the diurnal cycle. We wanted a resolution of ~1.5 km in mid troposphere in view of our resolution of 150-200 km in mid latitudes. We wanted good resolution near the tropopause and finally we felt that we could not produce operational analyses for heights much greater than 25 mb. We finally decided then on the spacing shown in Table 2.1. The values are in fact specified by the polynomial

$$\sigma_k = (3s + 7s^3 - 6s^4)/4 \qquad s = (2k.1)/2K\,,\ k = 1,2..K, K=15$$

This function was determined by the requirements that it be linear near s=0, quadratic near s=1 and should have the values 0, ½, 1 for arguments 0, ½, 1. The practical advantage of a functional representation is that we have a smooth variation of the layer thicknesses. A functional representation is also of advantage in theoretical studies. The levels used in the GFDL 9-level model are included in Table 2.1 for comparison together with the standard atmosphere heights, height interval and sigma thickness. In terms of pressure the resolution is roughly 50 mb in the stratosphere, 70 mb near the tropopause, 90 mb near 500 mb and there are four levels in the boundary layer (below 840 mb).

TABLE 2.1

Values of σ and height (m) at main model levels together with height and σ thicknesses. Also shown are the σ values of GFDL 9 level model.

k	σ_k	z_k	$z_k - z_{k+1}$	$\sigma_{k+\frac{1}{2}} - \sigma_{k-\frac{1}{2}}$	GFDL σ levels
1	0.025	24900	7100	0.050	0.009
2	0.077	17800	3400	0.053	0.074
3	0.132	14400	2400	0.058	
4	0.193	12000	1950	0.064	0.189
5	0.260	10050	1700	0.071	
6	0.334	8350	1500	0.077	0.316
7	0.414	6850	1300	0.083	
8	0.500	5330	1200	0.087	0.500
9	0.588	4350	1150	0.090	
10	0.678	3200	850	0.088	0.684
11	0.765	2350	800	0.084	
12	0.845	1450	750	0.075	0.811
13	0.914	700	430	0.061	0.926
14	0.967	270	240	0.042	
15	0.996	32		0.016	0.991

The time scheme was a semi implicit leap frog scheme with a time step of 15 min (Burridge, Haseler 1979). The CFL criterion in polar regions was satisfied by setting to zero all Fourier amplitudes in the tendencies of p_*,T,u,v,q whose wavelength was less than the smallest resolvable zonal wavelength at 30^o. We refer to this procedure as total tendency chopping. A time filter (Asselin, 1971) was used, with coefficient 0.05 to prevent the splitting of the solution into two families in time. In view of the long time step this is a modest value for the coefficient.

3. Parameterization of sub-grid scale processes

Two parameterization schemes were used:

1. The GFD scheme designed by Smagorinsky et al (1965) and by Manabe et al (1965). This scheme has been extensively used at GFDL for many studies of the general circulation of the atmosphere and for many forecasting experiments. Apart from the radiation scheme the parameterizations were recoded for efficient implementation on our system.

2. The ECM-scheme recently developed at ECMWF (A full description can be found in Tiedtke et al (1979)).

The main features of the schemes are given in Table 3.1. The following processes are treated in both schemes:

- Radiative exchanges
- Surface fluxes of momentum, of sensible heat and of moisture.
- Turbulent vertical fluxes of momentum, of sensible heat and of moisture.
- Condensation processes due to large-scale processes and due to convection.
- Detailed treatment of surface conditions including prediction of soil moisture and of snow amount.

The GFD-scheme and the ECM-scheme differ mainly in the following points:

Large-scale condensation:

GFD: Large-scale condensation occurs whenever the relative humidity exceeds 80%.

ECM: A height dependent criterion for condensation is used, $r_c = 0.80 + 0.20\ e^{-7.0(1-\sigma)}$ so that $r_c = 100\%$ at $\sigma = 1$ ($r_c \sim 90\%$ at $\sigma = 0.9$ and $r_c = 80\%$ at $\sigma = 0$). Evaporation of rain in non saturated subcloud layers is also treated.

Surface fluxes

GFD: Surface fluxes calculated as for neutral stratification.

ECM: Surface fluxes dependent on stratification according to Monin-Obukhov similarity theory.

Turbulent vertical fluxes above surface layer

GFD: Momentum and moisture fluxes occur only in the lower troposphere ($z < 2500$ m). Fluxes calculated by means of mixing length theory for neutral stratification.

Sensible heat fluxes considered only for unstable stratified conditions (by dry convective adjustment scheme).

ECM: Turbulent fluxes of momentum, of moisture and of sensible heat parameterized by vertical diffusion, where the diffusion coefficients depend on thermal stratification and on windshear.

Moist convection

GFD: Moist convective adjustment scheme (Manabe et al (1965)).

ECM: Kuo-convection scheme (Kuo, 1974).

Radiation

GFD: The absorbers, H_20, $C0_2$, 0_3, and the clouds are prescribed as functions of latitude and height.

ECM: Moisture and clouds are taken from the model, thus providing a radiation-cloud feedback. $C0_2$ is specified with a constant mixing ratio and 0_3 is specified as a function of all three space variables.

Surface temperature

GFD: T_s over land diagnosed from energy balance equation.

ECM: T_s predicted over land.

Horizontal turbulent fluxes

The formulation for the horizontal turbulent fluxes is described in Hollingsworth and Geleyn,(1979). The formulation is nonlinear and involves a fourth order spatial operator. The nondimensional coefficient was taken to be 0.01 for runs with both parameterization schemes.

TABLE 3.1

Survey of the parameterization schemes

Process	GFD	ECM
Large-scale condensation	No clouds Condensation, if relative humidity exceeds 80%	No clouds Condensation, if relative humidity exceeds $U=0.8+0.2e^{-7.0(1-\sigma)}$ Evaporation of rain
Convection		
Dry convection	Dry convective adjustment	Mixing by vertical diffusion of sensible heat, moisture and momentum
Moist convection	Moist adiabatic adjustment	Deep convection by Kuo convection scheme
Turbulent motion		
Horizontal diffusion	Non-linear, fourth order in space	Non-linear, fourth order in space
Vertical surface fluxes	$F_X \sim C_d(\lvert V_n\rvert)(X_h - X_s)$ $C_d = \text{const.}$	$F_X \sim C_d(\lvert V_h\rvert)(X_h - X_s)$ $C_d = C_d\left(R_i, \frac{h}{z_o}\right)$ (R_i = Richardson number)
Vertical fluxes above surface layer	Vertical diffusion assuming $K=K(z)$ with $K=0$ for $z > 2500$ m	Vertical diffusion assuming $K=K(R_i,\ell)$ and $\ell=\ell(z)$
Radiation	Absorber: Zonal means of H_2O, CO_2, O_3, clouds (specified at 3 levels)	Feedback between moisture (clouds) and radiation.
Surface values		
T_s Ocean	Prescribed	As in GFD
T_s Land	Diagnosed (energy balance equation)	Predicted
Soil moisture	Predicted	As in GFD
Snow	Predicted	As in GFD

4. The data for the experiments

4.1. Initial data sources

The main body of data used was the global data from the Data Systems Test carried out by NMC Washington for February 1976.

We selected seven cases, viz. 3,6,9,12,15,18,22 Feb all at 00Z. From the NMC DST tapes we had (with a resolution of N36) winds and geopotentials at 1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50 mb together with relative humidity at 1000, 950, 700, 500, 400, 300 mb; and temperature at 1000 mb.

We also needed 10 mb data as our top model level is $\sigma = 0.025$ and so approaches 10 mb over high topography.

From the British Meteorological Office we had the 00Z 10 mb geopotential analysis on their octagon grid (resolution ~ 300 km) and from NCAR, with a resolution of N18, we had zonal averages (for December/January/February) of velocity and geopotential at 10 mb derived from height level data. The top model level data were interpolated from the octagon data and the DST data; in the southern hemisphere the interpolation was between the zonal mean climatology and the DST data.

From GFDL we had, on their modified Kurihara grid, surface geopotential, soil moisture, albedo and snow taken from their data set for March 1 1965.

Finally from RAND (Alexander and Mobley, 1974) we had February mean sea surface temperatures with a resolution of N90.

4.2. Initial data interpolation

These data were bi-linearly interpolated in the horizontal to our latitude longitude grid. Bi-linear interpolation was also used to stagger the components of horizontal wind. From the octagon 10 mb height, in the northern hemisphere, velocities were calculated geostrophically and then extrapolated to the zonal mean at the equator.

In order to interpolate vertically to sigma-levels we had first to determine surface pressure p_*. When $p_* < 1000$ mb this was done from the pressure level data by linear interpolation of ln(p) in height between the heights of the nearest standard pressure levels. When $p_* > 1000$ mb the surface pressure was found by

$$p_* = \exp\left\{\frac{Z_{1000}}{\beta} + \ln 1000\right\} \text{ where } \beta = \frac{R}{g}\left\{\frac{T_{1000} + T_*}{2}\right\}$$

and $T_* = T_{1000} - \gamma z_{1000}$ where γ is -6.5°K/km.

Knowing p_* the pressures corresponding to the model's sigma-levels are determined. Geopotential was then interpolated linearly in ln(p) to the sigma half-levels (except $\sigma = 0$) and to the highest main sigma level. From the geopotentials, using the models hydrostatic equation, virtual temperatures at the main levels were calculated. Finally temperatures were calculated by an iteration using the virtual temperature and the relative humidity.

Velocity components were interpolated linearly in ln(p) to the main σ-levels. Where extrapolation below 1000 mb was necessary the following procedure was followed. Having calculated the pressure at 1000 m by linear interpolation in ln(p), u and v at 1000 m were also calculated by linear interpolation in ln(p) and finally $\left\{{u \atop v}\right\}_{\text{lowest level}} = k \left\{{u \atop v}\right\}_{1000\text{ m}}$ where k = 0.5 over land, 0.667 over sea.

The stratospheric humidity mixing ratio (q) was treated in the following way. Below 300 mb the relative humidity r was interpolated linearly in ln(p) to the sigma levels. Having found T by the procedure indicated q was defined by the approximate formula

$$q = \frac{\varepsilon r e_s(T)}{p}$$

For levels above 300 mb the mixing ratio was computed from relative humidity extrapolated from below 300 mb according to the formula

$$r(p) = r(p_t)\,\frac{p}{p_t}$$

where p_t is the nearest sigma level below 300 mb. This extrapolation of r was adopted to prevent q being precisely zero at levels above 300 mb which would have caused problems for the radiation scheme. It does in some circumstances permit an increase of q with height but this was not important.

4.3. Initialization

All the data sets were initialized using the non-linear normal mode initialization scheme described in Temperton and Williamson (1979).

4.4. Processing of forecast data

Geopotential (ϕ) and p-vertical velocity (ω) are not normal history variables of the model. At twelve hourly intervals they were therefore calculated from the other variables on σ-levels using the models finite difference scheme. The variables were interpolated linearly in ln(p) to pressure coordinates. The main problem was the extrapolation of data below surface in mountainous areas. The "Shuell method" (Shuman and Newell 1975) was used for extrapolating the temperature and height fields. The wind-field and the relative humidity field were assumed constant from the lowest model level downwards.

Because of limitations on the amount of data transfer only data of a small area covering Europe and the Atlantic north of 10 ^{O}N was saved at full resolution in the original grid. To cover the whole globe a second data set was created with a coarser longitudinal resolution, namely a representation by zonal Fourier series with wavenumbers up to 20. This data set has a vertical resolution identical to that of the available verification data while the smaller European/Atlantic area had a few more levels especially in the boundary layer.

Most of the presentations and verifications used in this report are based on the Fourier truncated global data set.

All calculations of e_s both in data preparation and in the model integrations were made with the Magnus formula (cf Murray 1967) viz

$$e_s(T) = 6.1078 \exp \left[\frac{a(T-273.16)}{T-b}\right]$$

where $a = 17.2693882$, $b = 35.86$.

5. Synoptic evaluation of the experiments

5.1 Method

A synoptic assessment is made of selected experiments, three of which have been studied, each with two different physics, Geophysical Fluid Dynamics Laboratory physics (referred to as "GFD forecasts") and ECMWF physics(referred to as "ECM forecasts"), based on February 1976 analyses. By comparing the predicted patterns and values of geopotential and temperature with the NMC analyses, these experiments have been subjectively assessed from a forecaster's point of view, in order to estimate the potential usefulness of the Centre's products to a forecaster in an operational environment. The experiments were considered from the point of view of a forecaster who might have implicitly followed the guidance given to him by the experiments and an estimate was thus made of the synoptic value of the forecasts. The assessments are mostly concentrated on the North Atlantic, European and Mediterranean regions. Some verifying precipitation data were obtained for the area of Europe.

The objective verification procedures used in other sections of this report appraise the experiments in a rather different way, and are different in character from the synoptic verification procedures reported here. The synoptic features of principal interest in a medium-range forecast are not always the same as those in a short-range (up to 48 or 72 hours) forecast. It may be of greater practical use to examine such features as cyclone tracks and major changes of type, and in this synoptic evaluation, an attempt has been made to concentrate attention on these features of the forecasts.

When trying to assess the number of days from the start that a forecast is of use to a forecaster,it is difficult to settle an average time for the entire Hemisphere, or even for an area the size of Europe, due not only to the variations of different parameters at different places, but also to the requirements of the forecast, since different parameters often have different lengths of prognostic value. If, for instance, the forecasts were used for forecasting for oil rigs in the North-Sea, it would be very important to have accurate gale and storm warnings and also wind direction during the first 4 - 6 days, and after that a general trend. In this case the geopotential gradients are more important than the absolute height values; temperatures and precipitation would be of less importance. In other cases other combinations of parameters would be useful.

The synoptic assessment is made from the 1000 mb map unless otherwise stated. In cases where the value of the 500 mb forecast differs from that of the corresponding surface forecast, this will be mentioned.

Although an exact definition of what is meant by prognostic value is impossible to give, an effort has been made to set up some objective rules for defining this. These rules may appear somewhat vague but even so they cannot always be strictly followed. For instance, if a typical synoptic disturbance is more than about 12 hours out of phase with the verifying analysis during the first 3 days, or if the windspeed is more than 50 % wrong when the geostrophic wind exceeds 10 m/s, or if the wind direction differs by more than about 45 degrees, then these forecasts would not be acceptable if the deficiencies last for a longer period than 12 hours. The temperature forecast is not accepted if it differs by more than 5 degrees from the true value. For days 3 to 6, these figures are doubled, although with a smooth and continuous transition between the periods. During days 6 - 10 the forecasts are considered to be of at least some value if they reflect certain similarities with the analysis in the large scale patterns. If the cyclone tracks and baroclinic zones do not differ more than about 500 km, then this information is considered to be of some value even if the individual depressions and ridges are out of phase. If the last 3 - 5 days are summarised to give a mean value or a trend, then this sometimes can yield information which is superior to climatology at least. Sometimes even these criteria are not applicable e.g. when the pattern is not well defined or the contour distribution is very flat or stationary or the disturbances are mostly small scale. For these cases only a subjective assessment could be applied from case to case.

In a few cases the surface forecast would be rejected by the above rules at a certain location, but later in the forecast the prediction skill improves considerably again. In these cases, which are noted in the assessments, the reader is left to judge whether to accept the forecast or not during and after this temporary decline in skill.

Furthermore, it should be borne in mind that the model used in these experiments is not the final version of the Centre's model. The model is in the course of development and the subject of research with a view to making improvements in predictability.

We discuss three cases in chronological order. The first case (Section 5.2) is the least successful whereas the second and third (Sections 5.3 and 5.4) are more successful. At the end of each section there is a summary. In two cases this is preceded by a short note on forecasts of precipitation from the model. Section 5.5 summarizes briefly all of the experiments.

5.2 Case 1 - forecasts from 3.2.1976

a) Synoptic summary 3/2 - 25/2

During the first 5 days a blocking high is situated over Northern Europe and cyclone tracks are split over the Atlantic with one track leading to the north of Scandinavia and another track leading over Spain to the Mediterranean. On the 8th the fronts reach the British Isles and the high starts moving eastward. Two days later westerly winds extend over western and northern Europe, and the fronts move eastwards over Europe till the 12th. Meanwhile a cut off low is formed near Genoa and at the same time a low develops south of Iceland in association with an outbreak of arctic air from Greenland. This low moves to Central Europe on the 14th and fills quickly. The following ridge expands and links with the still persistent high centered near Moscow, and on the 16th the situation is similar to that at the beginning of the month, but this time the high becomes even more intense. It blocks the disturbances until the 21st when the first fronts pass Britain and Spain. They are then steered northward towards Scandinavia. On the 24th the situation is fairly similar to that on the 10th over Europe with fronts passing quickly over northern Europe in the strong south-westerly flow and high pressure dominating over the central and southern parts of the Continent.

b) Forecast evaluation

Day 0-2

The initial state of the atmosphere is shown in Fig. 5.2.1. During the following two days a filling low is incorrectly forecast to deepen east of Greenland in both models and starts moving south-eastward on the top of a longwave ridge towards the north of Scandinavia. On day 2 at 1000 mb there are some areas with unrealistically high gradients in both forecasts with winds in excess of 40 $msec^{-1}$ forecast over North America. Both forecasts intensify almost every low too much (Figs. 5.2.2, left). Over Europe, both forecasts are still acceptable according to the criteria discussed in the Introduction.

Day 3-5

The ridge over Northern Europe intensifies and progresses only slowly to the east in the analysis (Fig. 5.2.2, right). Both forecasts attack the blocking high from the southwest by a trough over Britain and by low pressure breaking through from near Greenland to Finland. On day 4 one finds this low east of Leningrad (Fig. 5.2.2, right), where in reality the centre of the blocking high is to be found. Both models incorrectly divide and redevelop the high with one part north of Scandinavia where the surface forecasts are misleading from around day 3, and another part over Turkey. Since, over Scandinavia, the error exceeds the limits according to the rules in para. 5.1 the forecasts are no longer useful for this area. This incorrect forecast of a major weather change is a significant error in both models. The southeasterly sea-level wind over southermost Europe is too strong in both forecasts, which are very similar in performance over this area. Both forecasts recover somewhat later in the forecast. Over the Atlantic and Pacific, however, there are some noticeable differences between forecasts, the ECM model being slightly more realistic. The best prognostic value in this period is for the central, western and south-western parts of Europe and this trend continues. Over the Middle East a new low is intensifying on the fifth day, associated with a strong outbreak of cold air from the north (Fig. 5.2.3-4). Neither of the models forecast this acceptably, instead they maintain the low near Sicily which the analyses show to have filled.

Day 6-7

Looking first at the eastern Pacific and North America on day 6 (Fig. 5.2.5) the models are not in phase with each other; there are also marked differences in amplitude, with the ECM model being slightly better. The GFD forecast shows a rapid easterly progression of some features, while the ECM forecast in general is a little slower. However, sometimes the 500 mb trough in the GFD model comes into better agreement with the analysed troughs, because the next trough upstream in the forecast has overtaken and has come into phase with the slower moving analysed trough. For example, the trough near the Great Lakes appears to be in correct phase on day 6 although the development was forecast wrongly. Over the Black Sea the observed cold upper trough intensifies and becomes persistent. The intensity of the trough is best indicated in the ECM forecast which also displays a slower and better phase speed. (Fig. 5.2.5). This can also be seen in the Hovmöller diagram near 45^{o} longitude (Fig. 5.2.6). The large Azores high has been displaced by day 5 too far southwards in both forecasts due to the low being forecast to approach Europe too quickly and on a track to the south of its correct position. The temperatures over the eastern Atlantic are accordingly lower in the forecasts both at 850 and 500 mb levels, and the resultant westerly winds over western Europe are not sufficiently well forecast from around day 5 or 6 (Fig. 5.2.3-5). The time-mean maps for days $4\frac{1}{2}$ - $7\frac{1}{2}$ (Fig. 5.2.7 , left) show the effect of this on the high over Europe. The mean flow forecast during this period has westerlies extending over France, the Benelux countries and north Germany, where the analysis shows the influence of the high persisting. The mean 1000 mb height forecast over eastern North America is over 16 dam low in places. Incorrect forecasts of the mean positions of major anticyclones and depressions over Asia and the Pacific are evident during this period of the forecast.

Charts showing the differences between the mean 1000 mb geopotential height and minimum 1000 mb height reached during the 3-day period including days 5,6 and 7 of the forecast (Fig. 5.2.7, right) show areas where large changes of geopotential height were recorded. These charts indicate regions of cyclonic, or (exceptionally) anti-cyclonic activity during this period. Thus the large centre near 160°E on the GFD forecast is associated with a depression forecast to first deepen, then fill and also move rather quickly northwards. The NMC analyses show a less intense low, which was slow-moving and filling. The forecast also had a low moving south eastwards and deepening rapidly over the North American Continent, while the analyses had less intense lows moving on an eastward track across this region, and the differences between the two can clearly be seen on the mean-min charts. The relative maximum just east of Newfoundland is associated with ridge development in the area. In the eastern Atlantic, the forecast shows a depression moving eastwards and filling, while the low in the analysis moved north to Greenland, then east near Iceland.

Investigation into possible uses of these mean-min and similar charts is continuing and will be the subject of a future report.

Day 8-10

The ECM forecast recovers somewhat again over Europe from around day 8 when the pressure is forecast correctly to fall, finally leading to an intense depression in the forecast on day 9-10 (Fig. 5.2.8). The GFD model is inferior due to an even faster and more incorrect easterly progression of the European low. However, both give a very wrong indication of type i.e. cold cyclonic rather than mild anticyclonic westerly, and the forecasts show little skill after days 5-6 for Europe. The east European high has a better position in the ECM model during the last 3 days and both models show some skill in the longwave pattern, the GFD model being slightly better over and around the north American Continent, whereas the ECM model forecast the Asian and Pacific features more accurately particularly the deep major 500 mb low north of Japan (Fig. 5.2.8).

c) Precipitation

Fig.5.2.9 and 5.2.12 show the observed precipitation over Europe during the first and third days respectively of the forecast period, with isohyet intervals of 2,5,10,15, 20,30,40 etc. mm. Figures 5.2.10-11,13-14 show the predicted accumulated daily amounts for the days corresponding to the ECM and GFD forecasts with isohyet intervals of 2 mm. There is a 6-hour difference between observations and forecast since the observations are made at 0600 GMT, but the effect of this is unlikely to be significant. The ECM model forecasts a noticeable increase in precipitation over the hemisphere between days 1 and 3 both with respect to intensity and areal distribution of rainfall. In the GFD forecast, this phenomenon is not so obvious, except for Europe where both models forecast a more widespread distribution of precipitation than observed over much of western Europe. However, the maximum values observed in the Mediterranean region are reasonably forecast by the models. Although the forecasts have not been compared with verification data for other areas, it is likely that the precipitation forecast for other parts of the hemisphere is too widespread. The ECM model produces less convective precipitation than the GFD forecast, and with different distribution. The GFD forecast shows a higher correlation between large-scale and convective precipitation.

d) Summary of forecasts from February 3 1976

Looking at the surface forecasts as a whole, one finds quite good resemblance in general between the ECM and GFD forecasts up to day 5, although there are some deviations mainly over the Pacific. From this time, the two forecasts start diverging from the analyses, and also from one another. For Europe, however, they are similar for the whole 10 days but useful only to 5 or 6 days, but over northern Europe only 3 to 4 days. Over Scandinavia , both models quickly and incorrectly divide and remove the strong blocking high after day 3, a high which the NMC analyses show persisted until day 5. Both forecasts have good and bad features. Perhaps the ECM model has a slightly more realistic evolution up to day 10 and certainly better and slower phase speeds of the synoptic waves, but the GFD model shows some skill in forecasting major systems, since it sometimes develops them in the right place at the right time, although the smaller scale troughs can be quite out of phase. Towards the end of the forecast period both models lower the pressure and temperature too much over the western parts of the two Continents, whereas the opposite distribution is found east of the Continents. The GFD forecast is, however, slightly superior to the ECM forecast over eastern Canada and the United States of America in this sense, whereas the ECM model is superior north of Japan. The ECM forecast has a better position of the Russian high during the last days of the integration. A noteable feature of both forecasts is the unrealistically high wind speed forecast at the 1000 mb level at times during the forecast period.

Continued on page 40

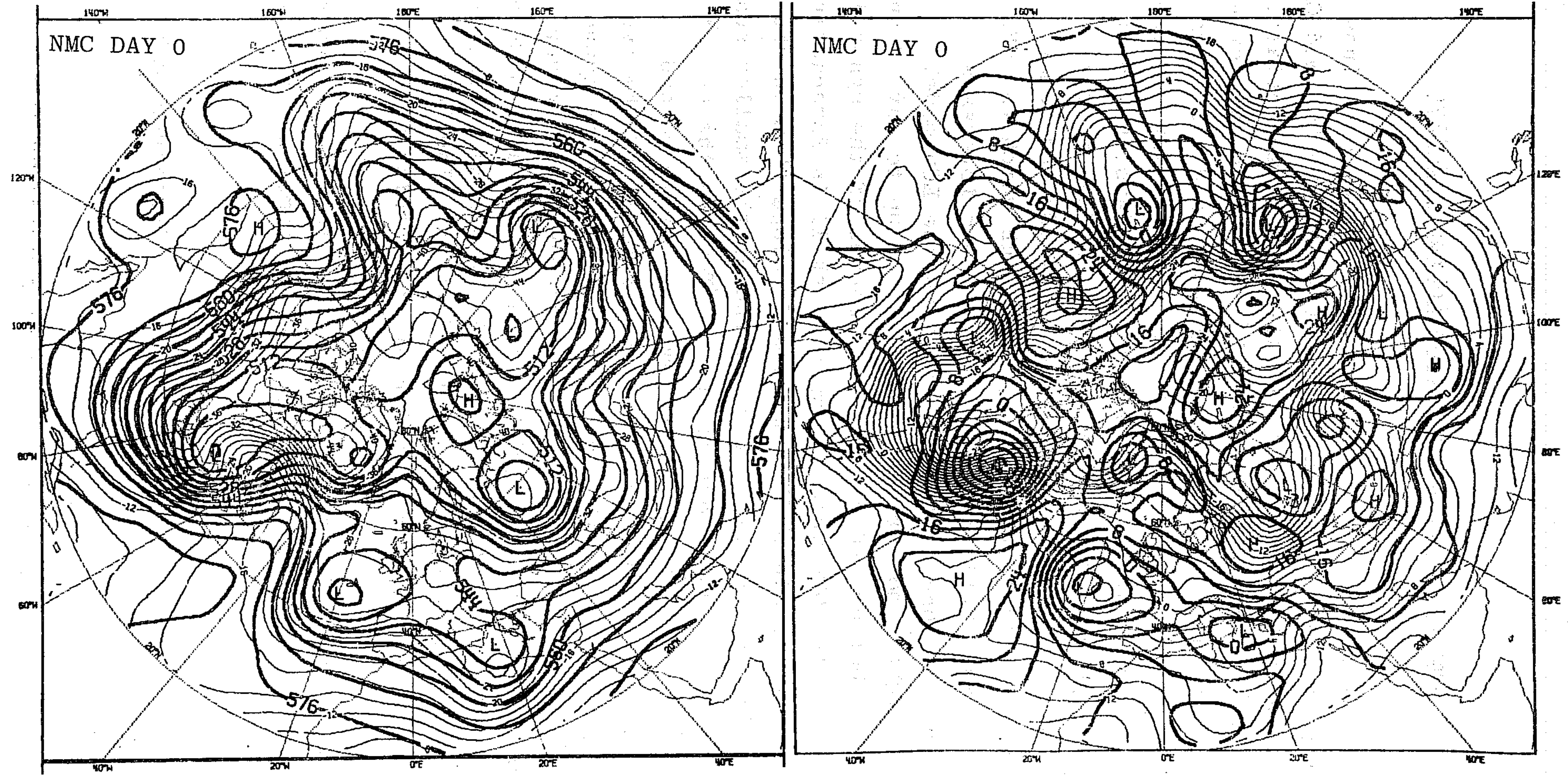


Fig.5.2.1 NMC analysis for initial day of forecast from 00Z on 3.2.76.
Left: 500 mb height field (thick lines, contour interval: 8 dekametres) and 500 mb temperature field (thin lines, contour interval: 2 ^{o}C).
Right: 1000 mb height field (thick lines, contour interval: 4 dekametres) and 850 mb temperature field (thin lines, contour interval: 2 ^{o}C.

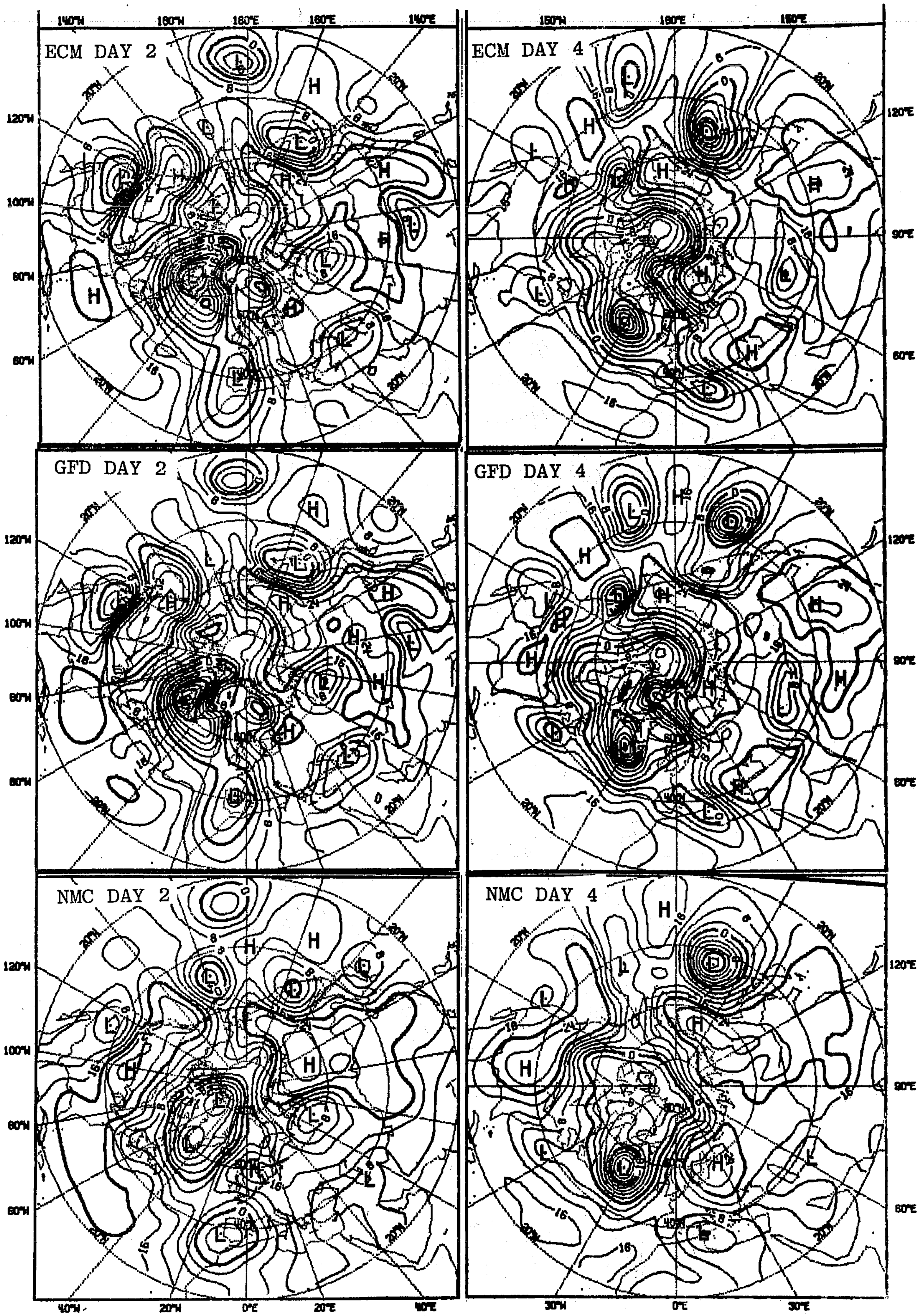


Fig.5.2.2 Left: 1000 mb height fields for day 2 of ECM and GFD forecasts from 00Z on 3.2.76 and corresponding NMC analysis. Right: 1000 mb height fields for day 4 of ECM and GFD forecasts from 00Z on 3.2.76 and corresponding NMC analysis. Contour interval as in Fig.5.2.1.

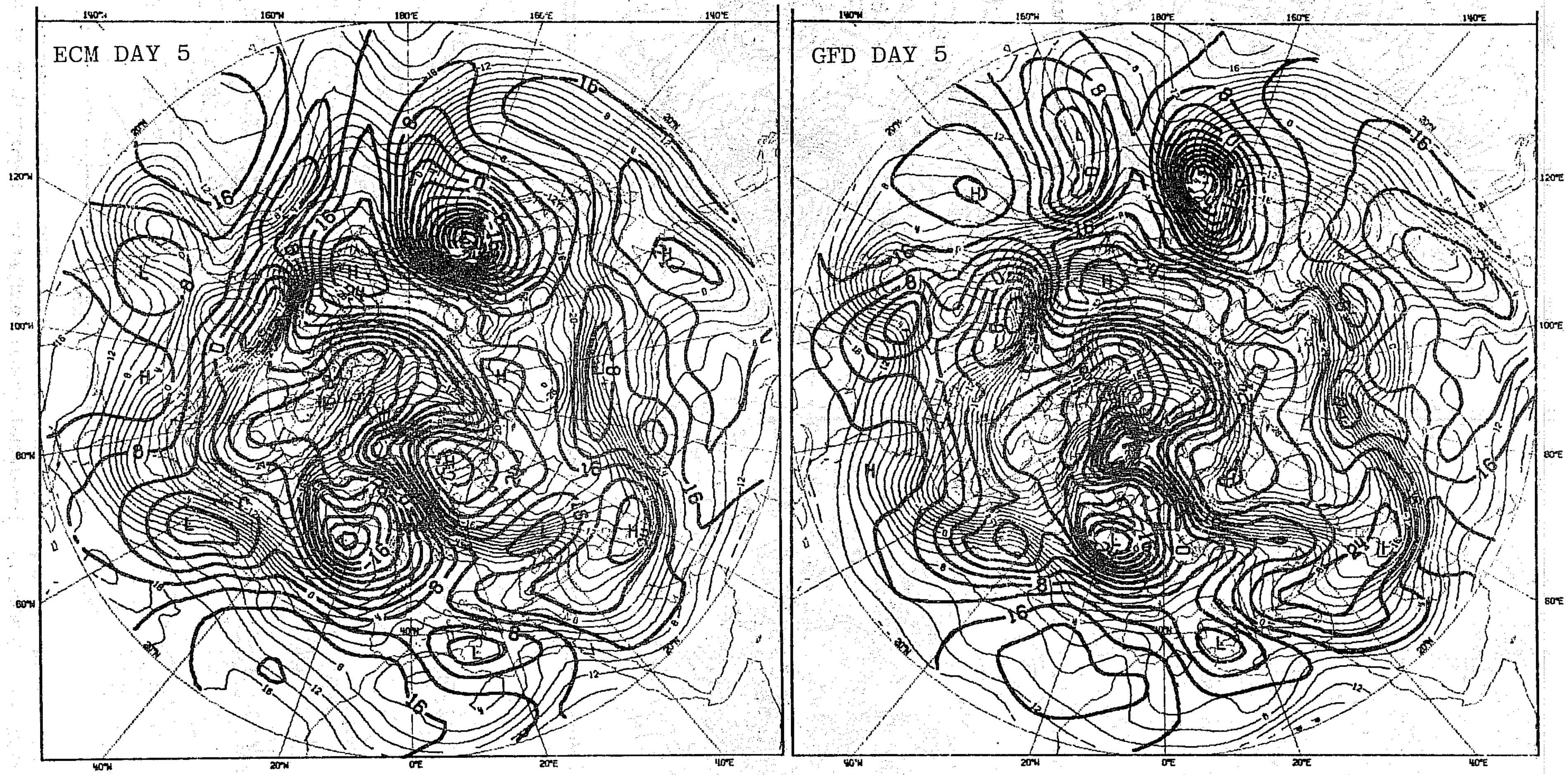


Fig.5.2.3. 1000 mb height and 850 mb temperature fields for day 5 of forecast from 00Z on 3.2.76. Left: ECM forecast. Right: GFD forecast. Contour intervals as in Fig.5.2.1.

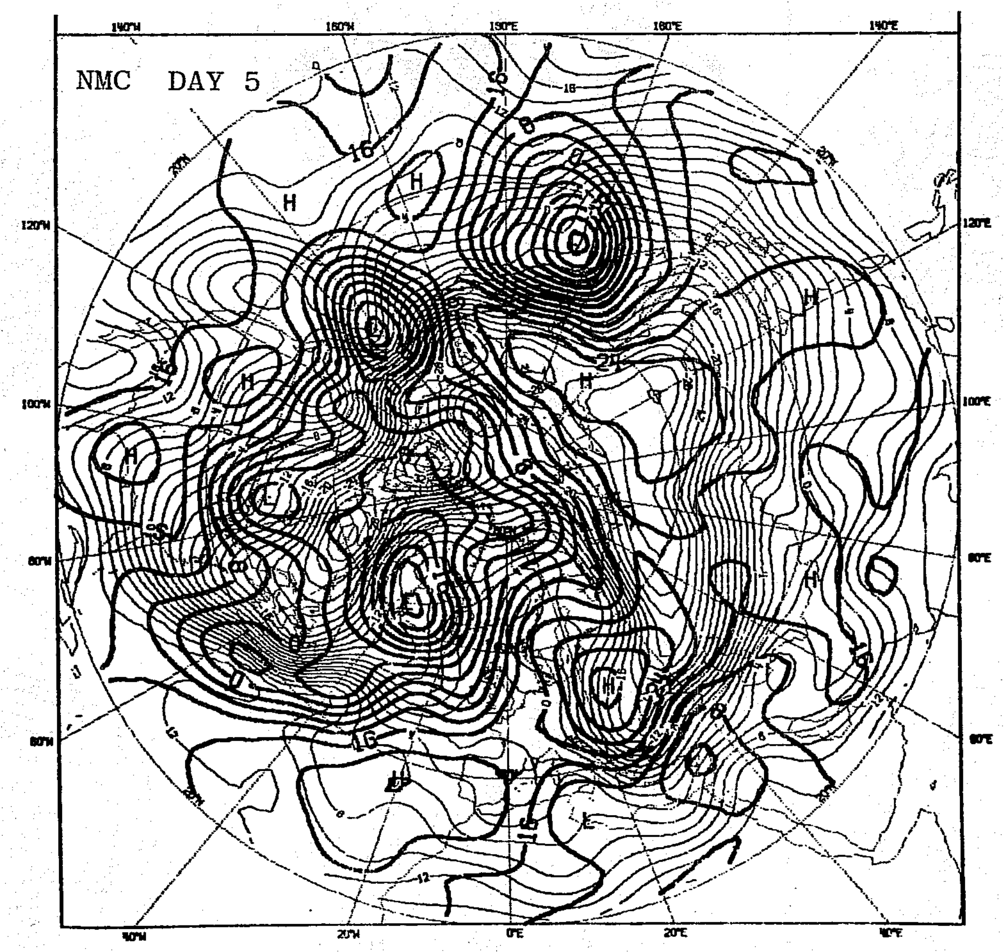
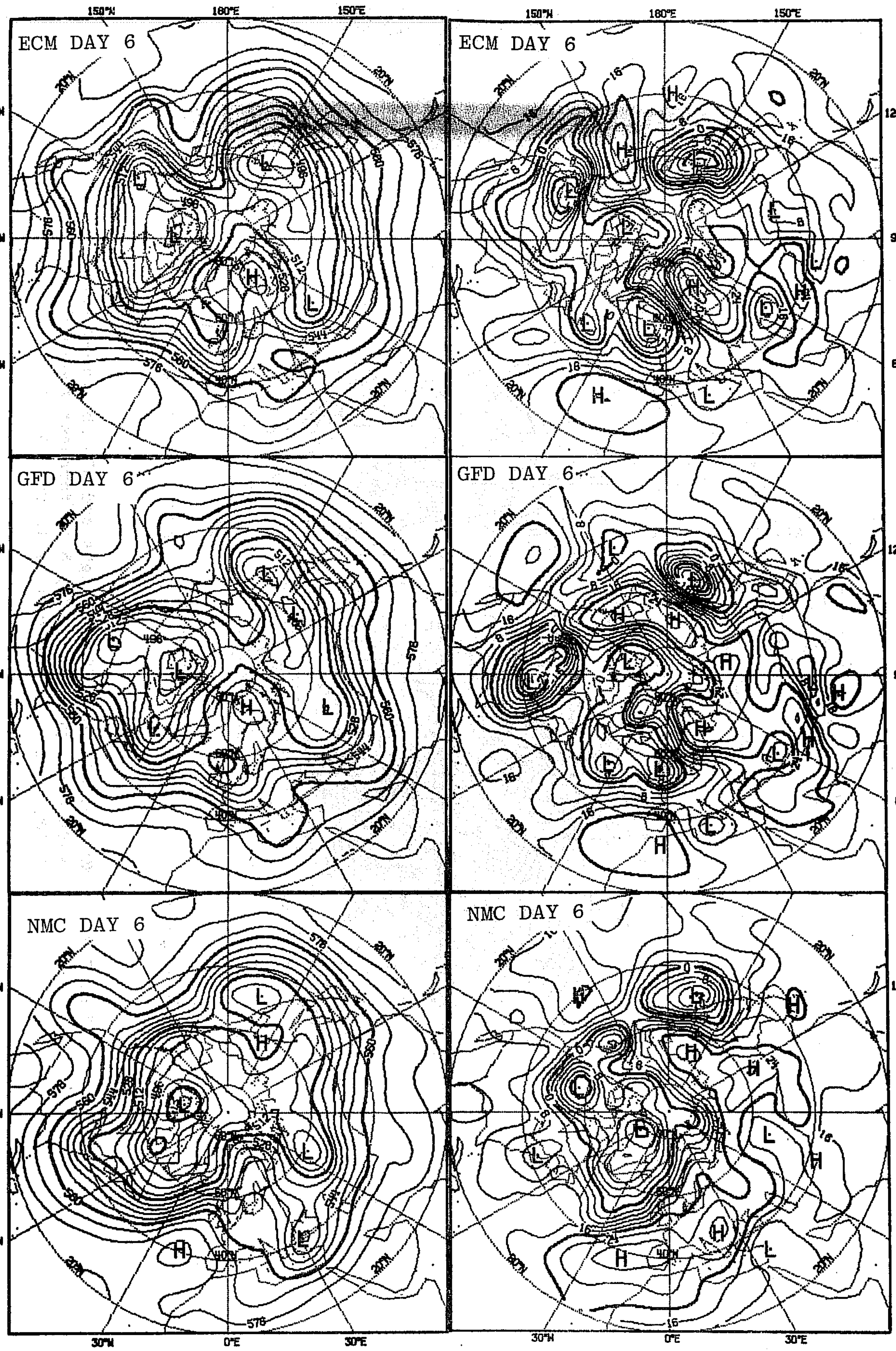


Fig. 5.2.4 NMC analysis of 1000 mb height and 850 mb temperature fields corresponding to day 5 of forecast from 00Z on 3.2.76. Contour intervals as in Fig. 5.2.1.



ECM DAY 6

ECM DAY 6

GFD DAY 6

GFD DAY 6

NMC DAY 6

NMC DAY 6

Fig.5.2.5 Left: 500 mb height fields for day 6 of ECM and GFD forecasts from 00Z on 3.2.76 and corresponding NMC analysis. Right: 1000 mb height fields for day 6 of ECM and GFD forecasts from 00Z on 3.2.76 and corresponding NMC analysis. Contour intervals as in Fig.5.2.1.

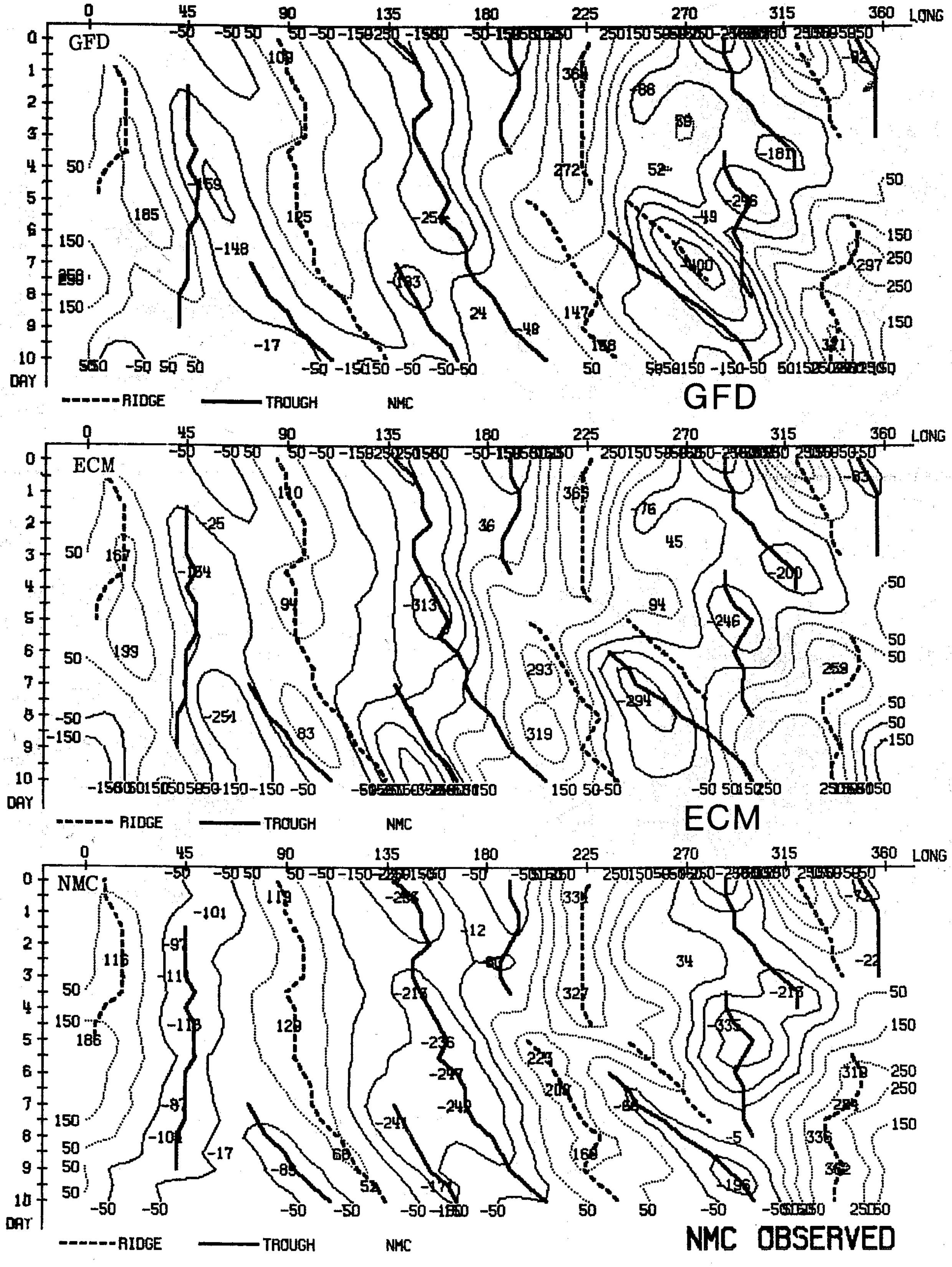


Fig.5.2.6 Hovmöller diagram at 500 mb for GFD and ECM forecasts from 00Z on 3.2.76 and corresponding NMC analysis. Wave number 1-9. Latitudinal mean between 35 °N and 55 °N.

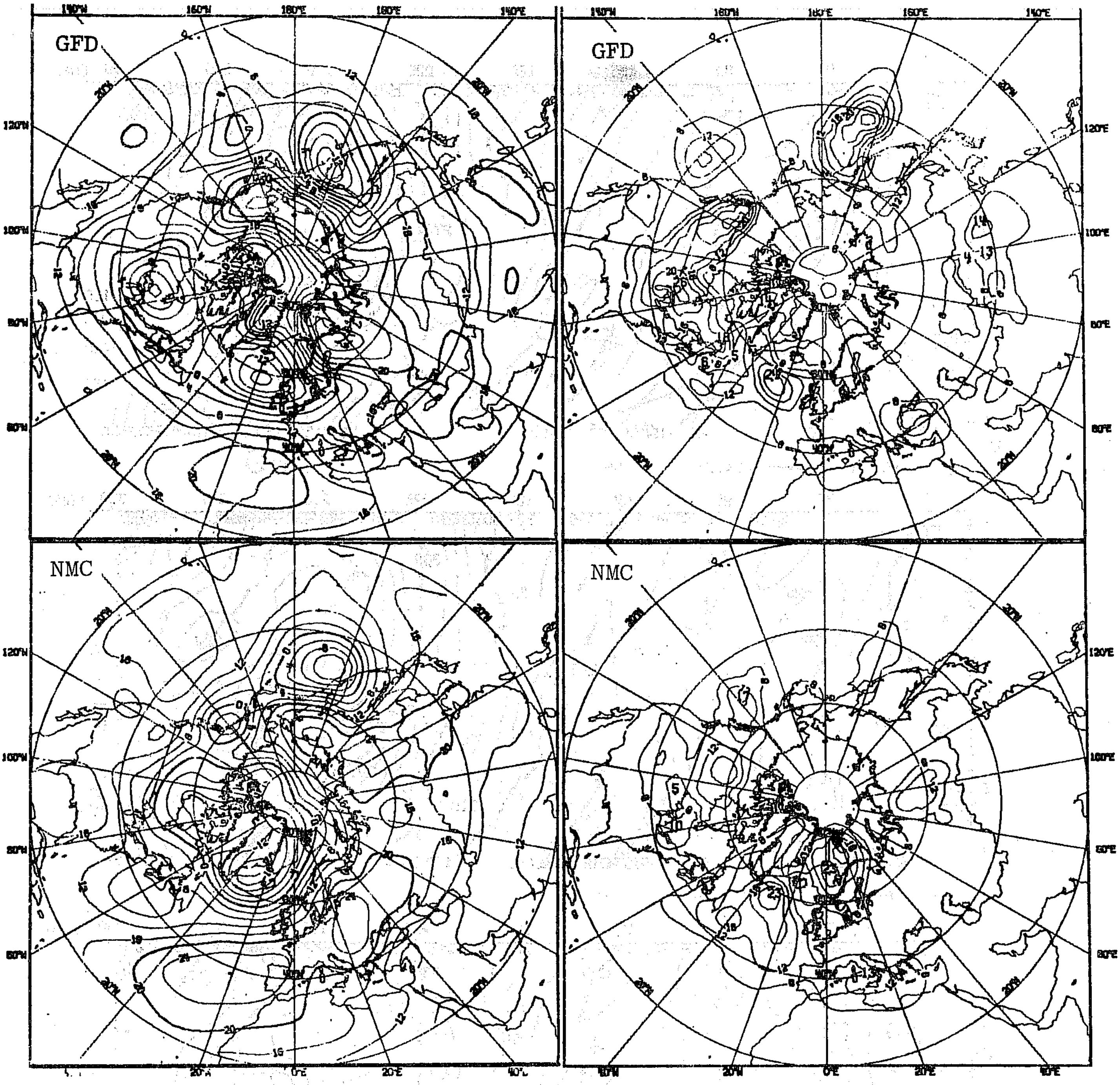


Fig.5.2.7 Left: time-mean maps for days $4\frac{1}{2}$-$7\frac{1}{2}$ (7.2.76 1200 GMT-10.2.76 1200 GMT) for GFD forecast from 00Z on 3.2.76 and corresponding NMC analysis. 1000 mb height fields. Contour interval: 4 dekametres. Right: 3 day mean-minimum 1000 mb maps for days $4\frac{1}{2}$-$7\frac{1}{2}$ (7.2.76 1200 GMT-10.2.76 1200 GMT) for GFD forecasts from 00Z on 3.2.76 and corresponding NMC analysis. Contour interval: 4 dekametres.

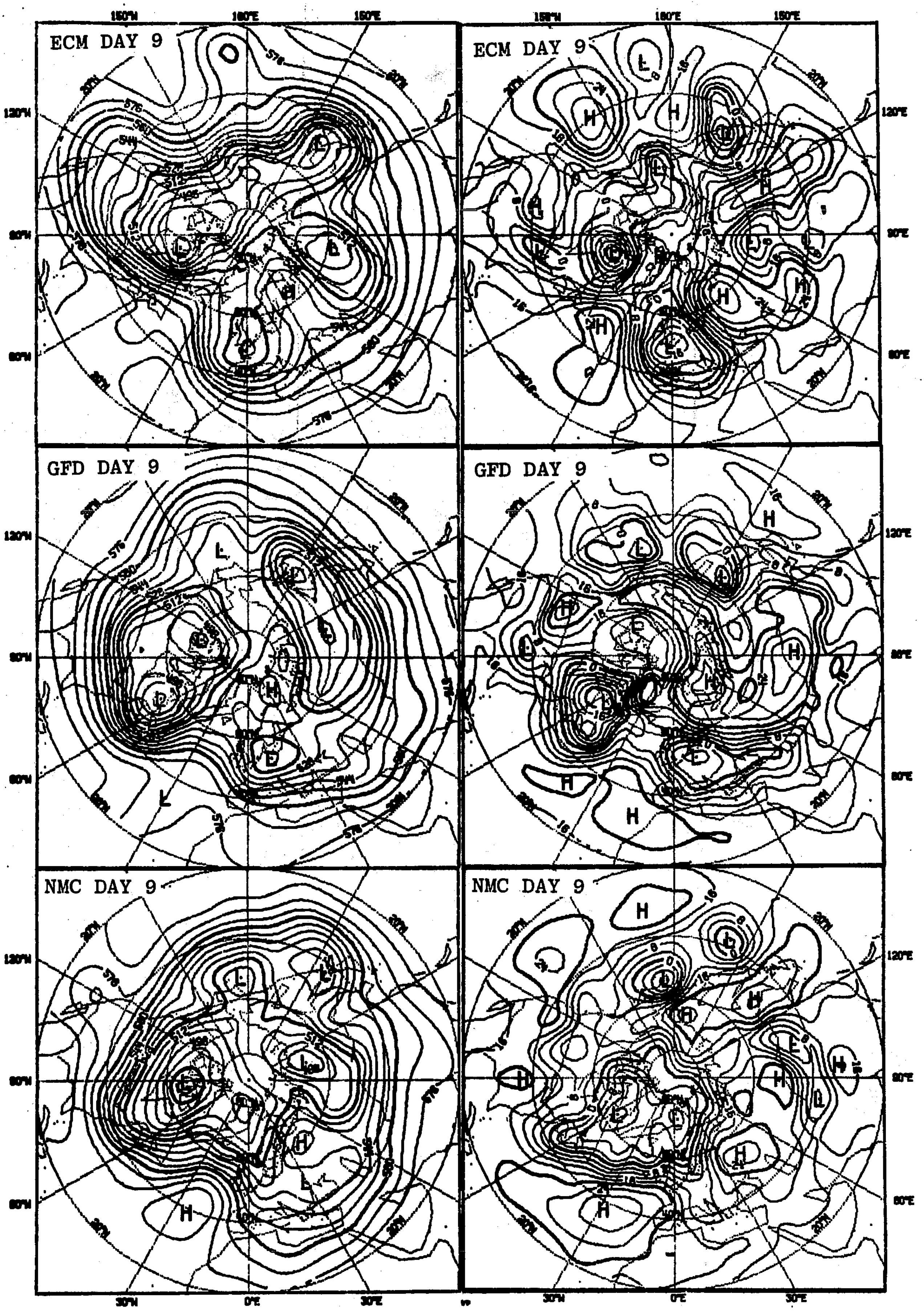


Fig.5.2.8 Left: 500 mb height fields for day 9 of ECM and GFD forecasts from 00Z on 3.2.76 and corresponding NMC analysis. Right: 1000 mb height fields for day 9 of ECM and GFD forecasts from 00Z on 3.2.76 and corresponding NMC analysis. Contour intervals as in Fig.5.2.1.

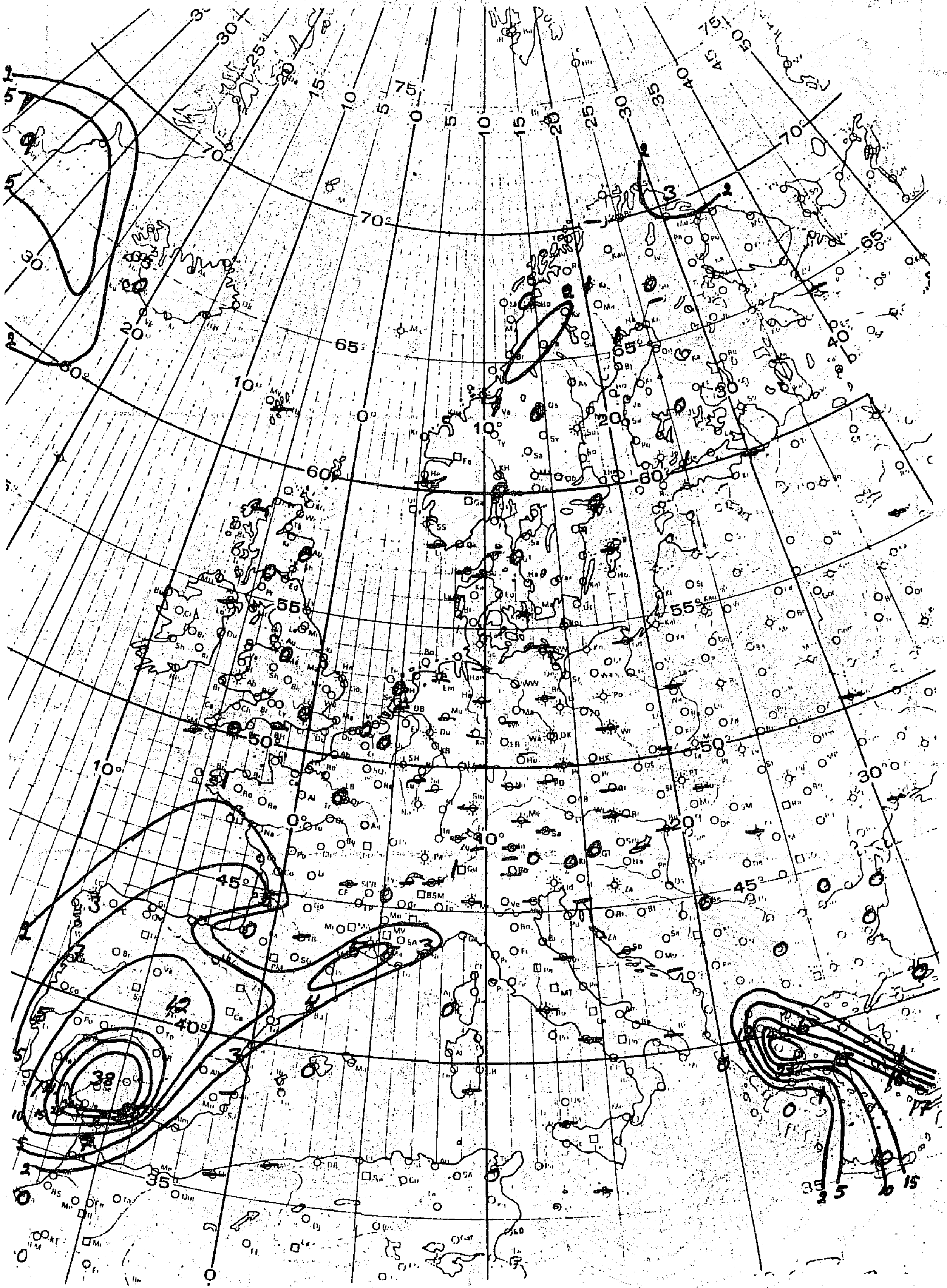


Fig.5.2.9 Observed 24 hours accumulated precipitation for period 3.2.76 0600 GMT-4.2.76 0600 GMT. Isohyet interval: 2,5,10,15,20,30,40,...etc.

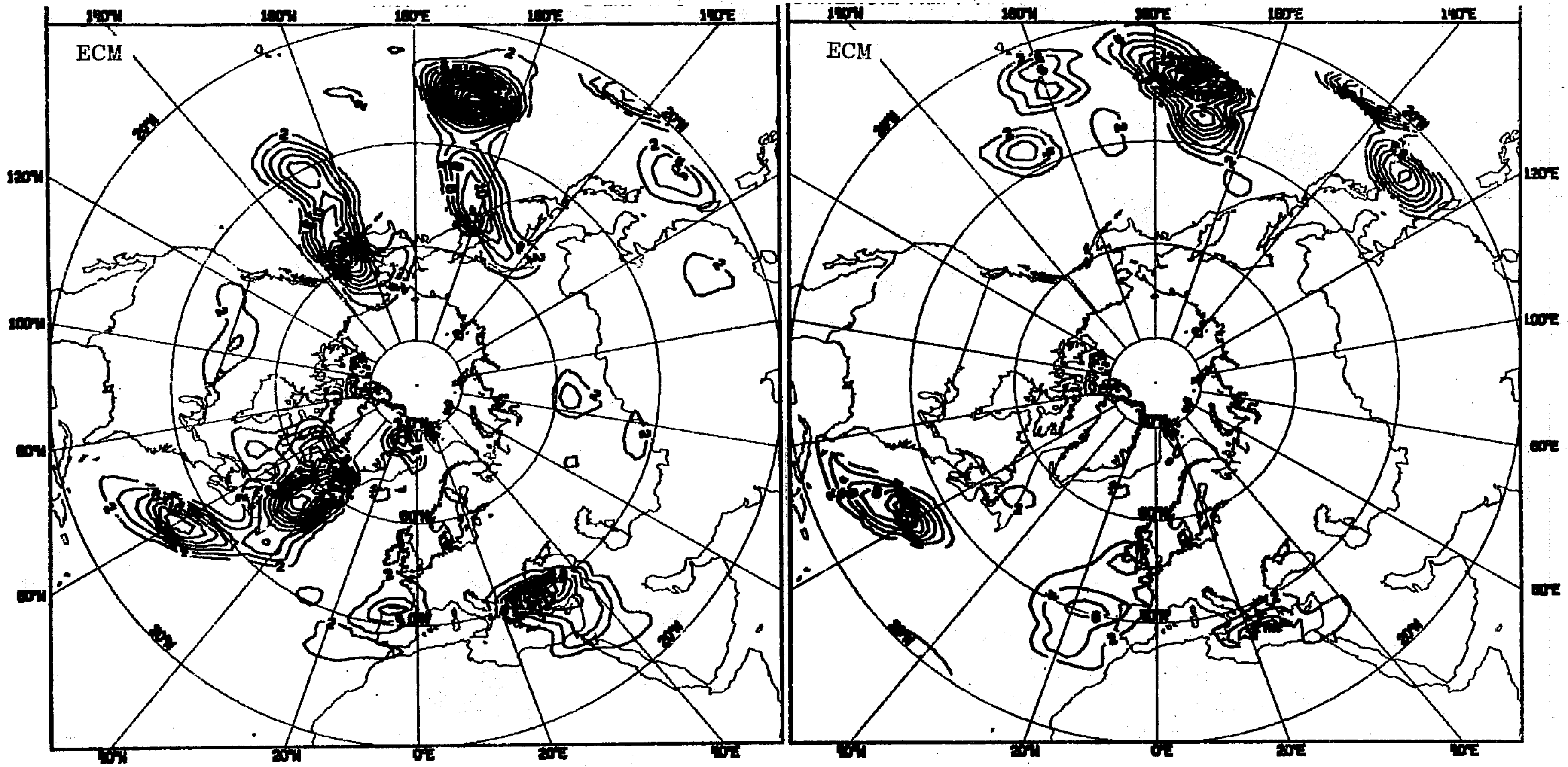


Fig.5.2.10 Left: ECM forecast accumulated large-scale precipitation for first 24 hours of forecast from 00Z on 3.2.76. Isohyet interval: 2 mm. Right: ECM forecast accumulated convective precipitation for first 24 hours of forecast from 00Z on 3.2.76. Isohyet interval: 2 mm.

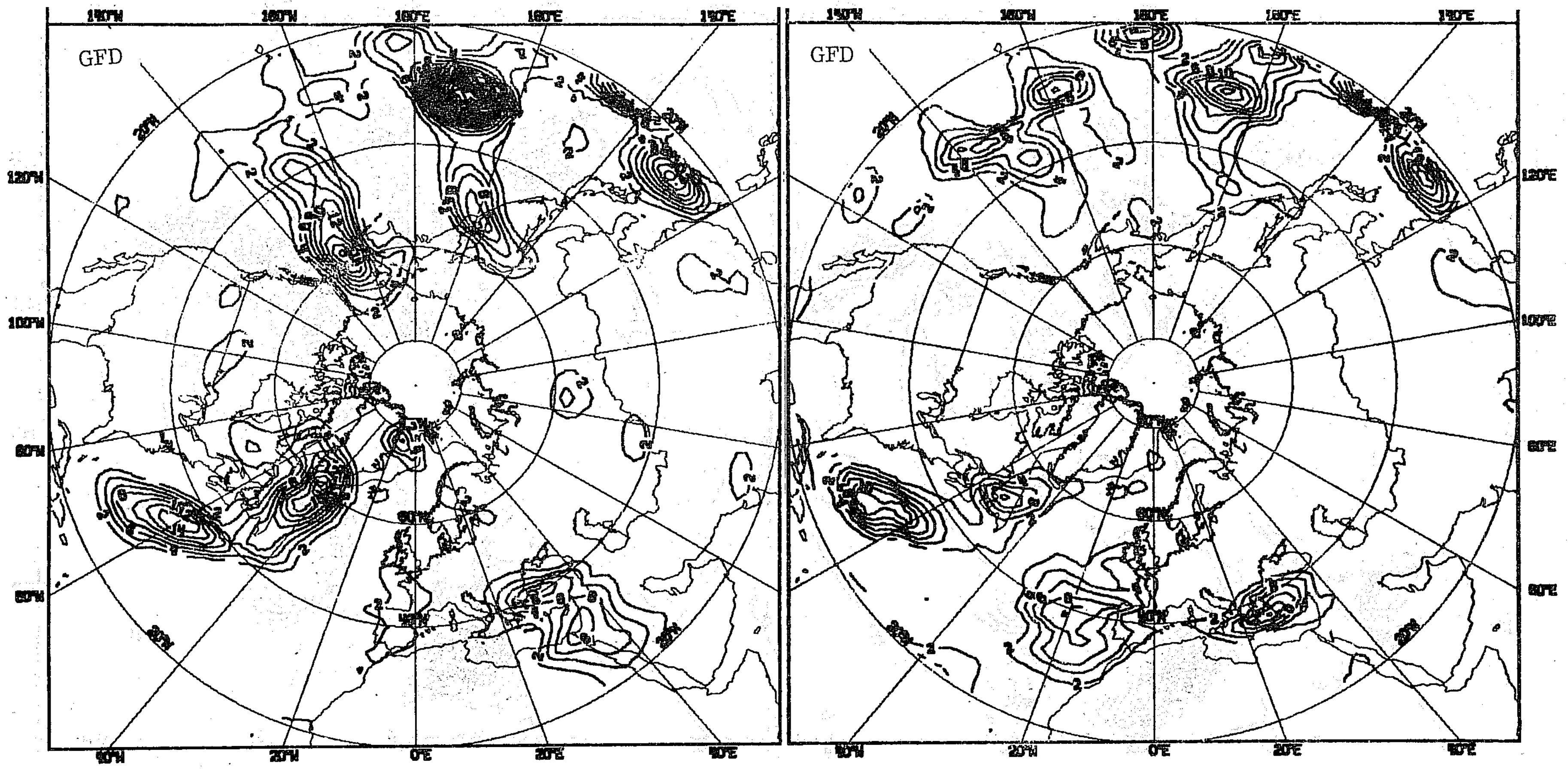


Fig.5.2.11 Left: GFD forecast accumulated large-scale precipitation for first 24 hours of forecast from 00Z on 3.2.76. Isohyet interval: 2 mm. Right: GFD forecast accumulated convective precipitation for first 24 hours of forecast from 00Z on 3.2.76. Isohyet interval: 2 mm.

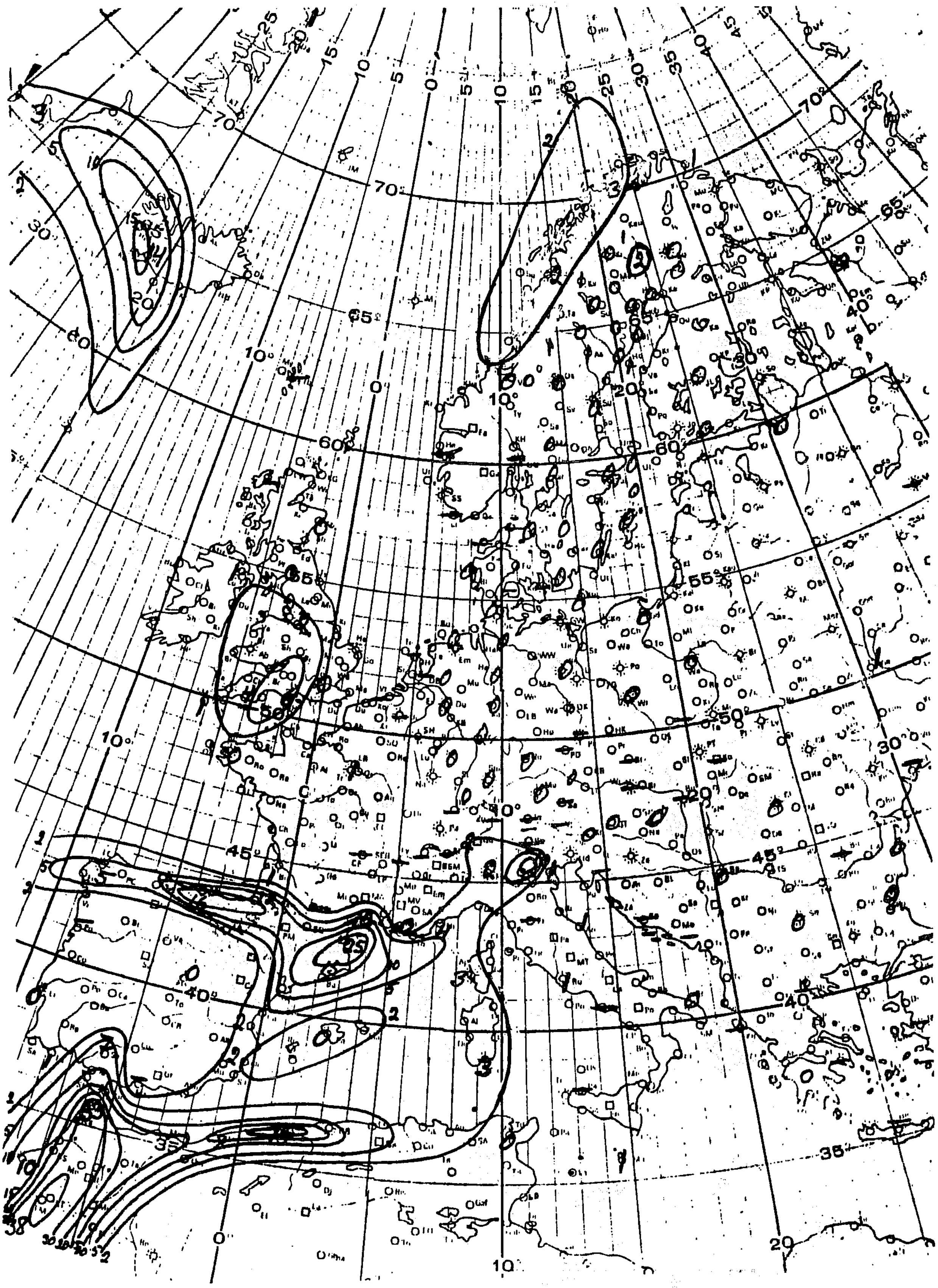


Fig.5.2.12 Observed 24 hour accumulated precipitation for period 5.2.76 0600 GMT-6.2.76 0600 GMT. Isohyet interval: 2,5,10,15,20,30,40,...etc.

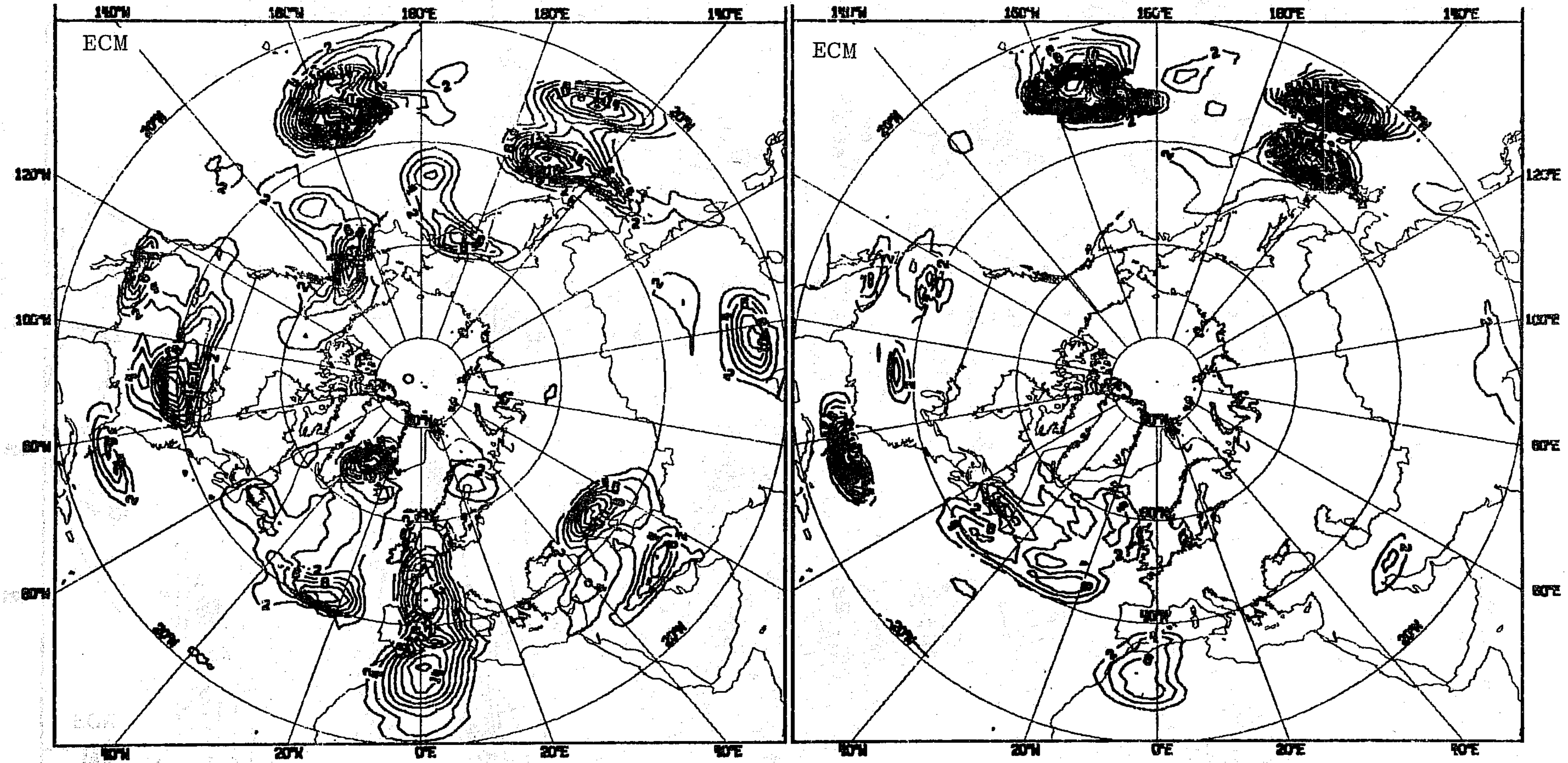


Fig.5.2.13 Left: ECM forecast 24 hour accumulated large-scale precipitation for day 2 of forecast from 00Z on 3.2.76. Isohyet interval: 2 mm. Right: ECM forecast 24 hour accumulated convective precipitation for day 2 of forecast from 00Z on 3.2.76. Isohyet interval: 2 mm.

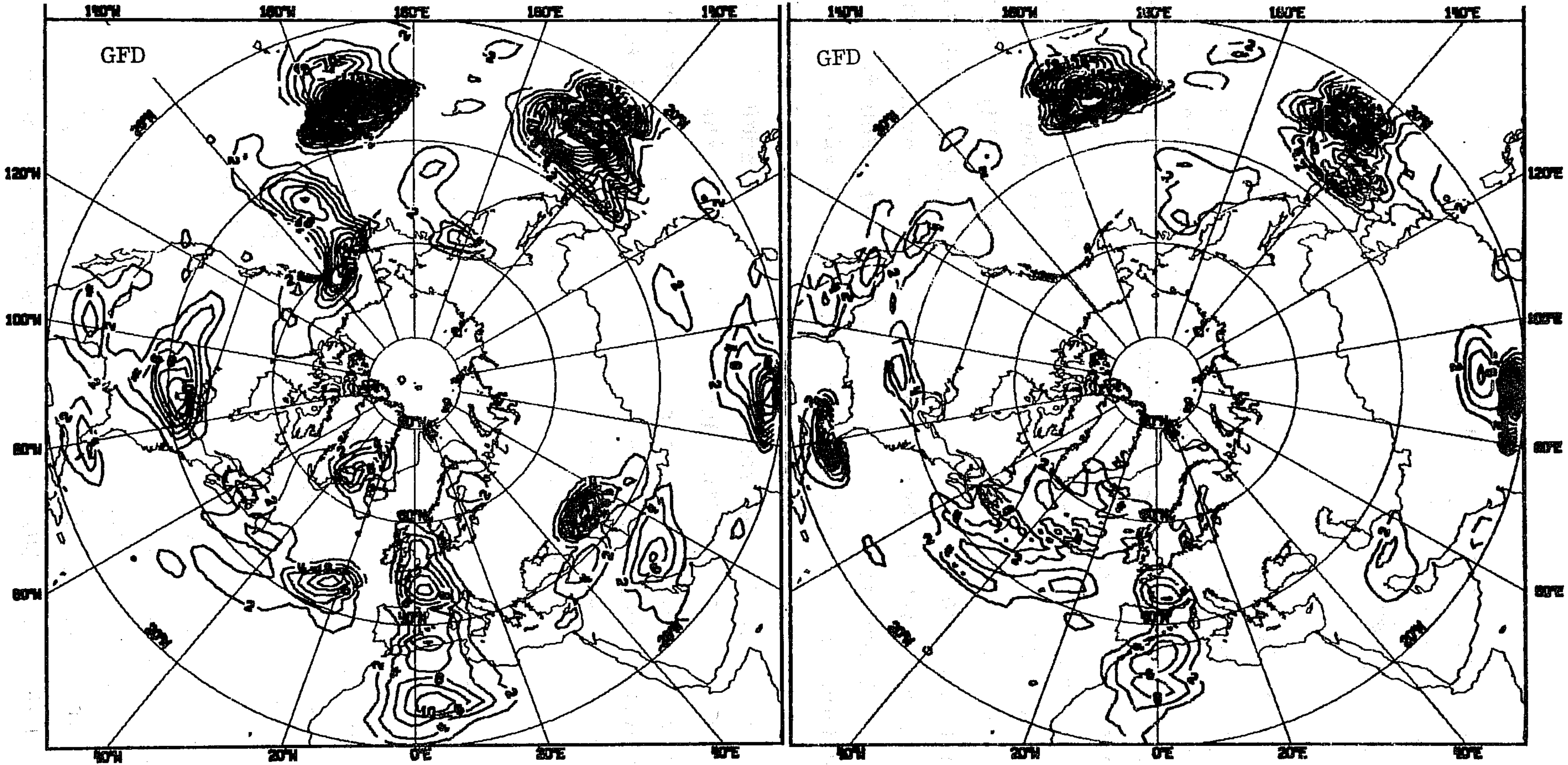


Fig.5.2.14 Left: GFD forecast 24 hour accumulated large-scale precipitation for day 2 of forecast from 00Z on 3.2.76. Isohyet interval: 2 mm. Right: GFD forecast 24 hour accumulated convective precipitation for day 2 of forecast from 00Z on 3.2.76. Isohyet interval: 2 mm.

5.3 Case 2 - forecasts from 6.2.1976

a) Synoptic summary

See 5.2a

b) Forecast evaluation

Day 0-3

The GFD and ECM forecasts are almost identical during the first three days. On day 1, the blocking surface anticyclone over Europe correctly intensifies by 7-8 dam in the models. During this period in both forecasts many lows are over-developed and move too quickly eastward. On day 3, although many features are moving too fast, the upper trough extending from Iceland to the North Sea and continuing over France is exactly in phase. (Fig.5.3.1-3). Although the 500 mb solutions are in good agreement with the analyses both in height and temperature over Greenland, the surface pressure is much higher and the 850 mb temperatures accordingly lower in the forecasts, due to interpolation errors over high ground (Fig. 5.3.1-3). At 500 mb there is, on some occasions, a tendency in both models, especially during days 3 and 4, for the temperature field to be in advance of the geopotential, for example, over California (Fig. 5.3.1-3),which may damp out the meridional flow too much. Both forecasts are still generally very good at this stage.

Day 4-7

On day 4 the influence of high pressure over Europe is, in general, correctly forecast, but in both models the high is predicted to be too intense to the southeast and over the Middle East. This has the effect of over-intensifying the westerly flow across central Europe to Russia. On day 5 the upper high has been cut off in a position too far to the north of Scandinavia in both models (Fig. 5.3.4-6), but the position of the ridge is good. Most depressions in the GFD forecast are now slightly more intense than the corresponding lows in the ECM forecast and both models in general forecast the pressure to be too low in centres of depressions. The 500 mb analysis on day 5 shows a slow moving closed low over the Black Sea with a cold centre, while the models develop troughs that move more quickly and are found further to the east on day 5.

On days 4-5 at upper levels over Europe a weak split in the jet stream is established with the weaker one over Spain and the other northward over Scandinavia. Fig. 5.3.6 shows the analysis corresponding to day 5 of the forecast. The forecasts show a more pronounced westerly current from Britain to Poland and the very intense surface low north of Scotland keeps moving on a slightly more southerly track (Figs. 5.3.4-5). However, the strong westerly gradient over the British Isles is very well predicted up to day 6 by both models, veering correctly to northwesterly on day 6 (Fig. 5.3.7, left). The high centred near Moscow in the analysis has been broken down too much and displaced to a more south-easterly position by the models.

On day 5 the pressure starts dropping rapidly over the Mediterranean area surrounding Italy. This is correctly forecast in the models and on day 6 a Genoa low is to be found in both forecasts (Fig.5.3.7, left). The low off the coast of Norway becomes slow moving over the Norwegian Sea in the forecasts in agreement with the analyses, but it fills too slowly and moves northward instead of southward. On day 6 the hemispheric forecasts are still very good, except over the Pacific, where both forecasts are out of phase with the analysis and with wrong amplitudes even in the major waves. Figs. 5.3.4-6 show that the forecasts had deteriorated already on day 5 for this area.

In the surface analysis a developing low south of Iceland on day 6 reaches the British Isles later the same day after having been more or less extinguished for a few days on top of the strong ridge over the Atlantic, (Fig.5.3.7, left). During days 6 and 8 it intensifies again, gaining vorticity in the northwesterly current towards the Alps. Although there are a number of similarities between the analysis and the ECM forecast on the 6th day, the development of this depression is not forecast to happen on day 7 in the ECM forecast and now this depression enters the European Continent as a weak trough,although exactly in correct phase (Fig.5.3.7,right). The GFD model forecasts a stronger trough from the North Sea over western France southwards to Spain where, over the northern part, there is a weak low. This low continues down to the Mediterranean till the next day and develops. Although a little too south it is a very good forecast at this late stage (Fig.5.3.7,right).

Days 8-10

The value of the ECM forecast, at least over southern Europe is becoming doubtful. There is an acceptable phaselag but this forecast shows a weaker trough over Europe, while the analysis shows an intense low near the Alps (Fig.5.3.8,left). On the other hand, the GFD model is nearer to the real atmosphere during day 8 and 9 over southern Europe (Fig. 5.3.8).

Until day 5 there is a rather good agreement for the large scale waves at the 500 mb level over the Pacific and North America. Between days 6 and 8 both models forecast incorrectly the intensities, phase-speeds and tracks for this area. However, on day 9 and 10 the ECM forecast in particular improves somewhat again over this area. From around day 5 both forecasts also seem to produce more kinetic energy in the meridional part of the synoptic waves. The phase error between geopotential and temperature at 500 mb is no longer observed. (Figure 5.3.8, right shows the actual and forecast situation at 500 mb on day 9).

c) Summary of forecasts from February 6 1976

The ECM surface forecast has good prognostic value up to around seven days over Europe, while the GFD forecast has value for about one day more. For other parts of the Northern Hemisphere both forecasts are useful to around 7 days, except for parts of the Pacific and eastern Asia where the forecasts are quite misleading from day 5. Although they improve slightly later, particularly the ECM forecast, the time of prognostic value cannot be significantly extended.

The value of the forecast depends upon the requirements of the user of the forecast. For instance, for ship routeing in the North Atlantic, the forecast would give useful guidance up to about day 6. For an oil-rig in the North Sea, both forecasts are good in both direction and speed to the morning of day 7. The accurate prediction of the westerly gale on days 3-5 is noteworthy. For agricultural forecasting over much of the south of Europe, the ECM forecast would be of value for around 7 days, the GFD model perhaps for one day longer. On the other hand, for wind forecasts on the Baltic or for ice breakers in the Gulf of Bothnia, the gradient is, for a short period, too strong around day 3, especially in the GFD model. After that the GFD forecast is useful for 7-8 days and the ECM model for 8-10 days.

As noted in other experiments, both models tend to lower the pressure too much over northern Europe after a few days, when strong depressions approach Europe on a track too far to the south, dividing the high into two parts, one moving too far north, the other one too far south. In the later stages of the ECM forecast, 500 mb troughs which approach and cross the west coast of North America are more intense than those observed in the analysis.

Continued on page 52

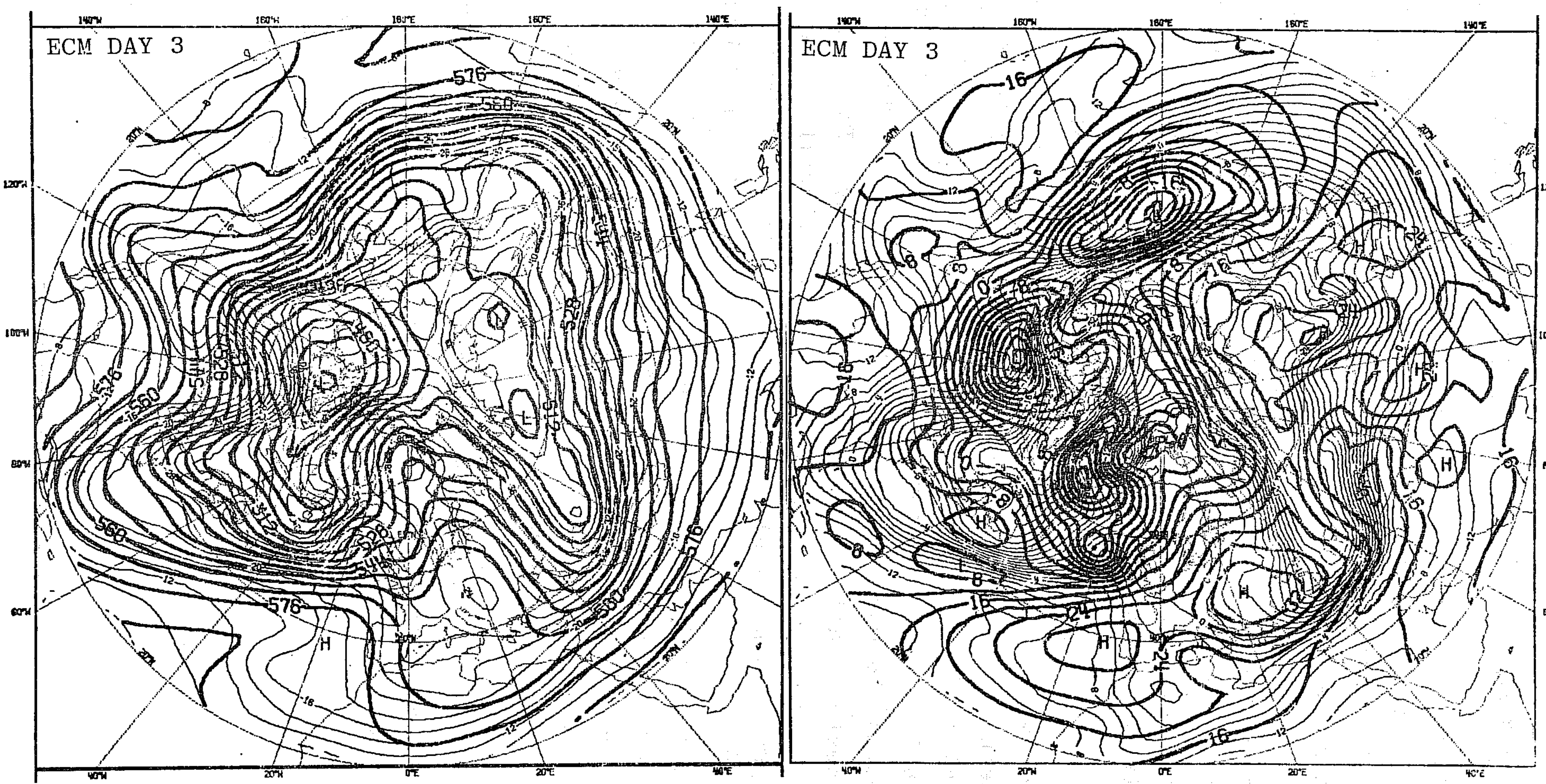


Fig.5.3.1 Left: 500 mb height field (thick lines, contour interval: 8 dekametres) and 500 mb temperature field (thin lines, contour interval: 2 °C) for day 3 of ECM forecast from 00Z on 6.2.76. Right: 1000 mb height field (thick lines, contour interval: 4 dekametres) and 850 mb temperature field (thick lines, contour interval: 2 °C) for day 3 of ECM forecasts from 00Z on 6.2.76.

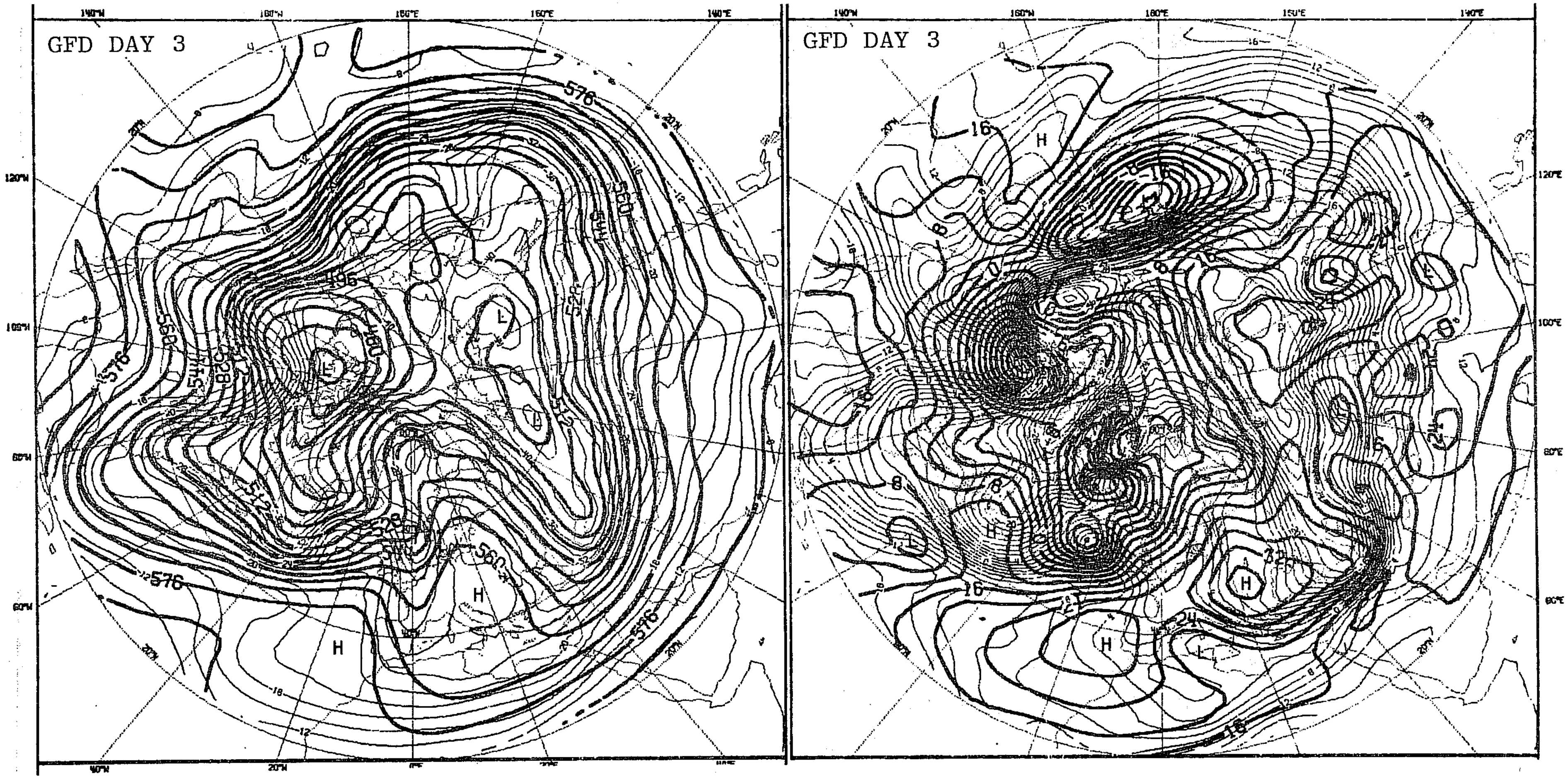


Fig.5.3.2 Left: 500 mb height and temperature fields for day 3 of GFD forecast from 00Z on 6.2.76. Right: 1000 mb height and 850 mb temperature fields for day 3 of GFD forecast from 00Z on 6.2.76. Contour intervals as in Fig.5.3.1.

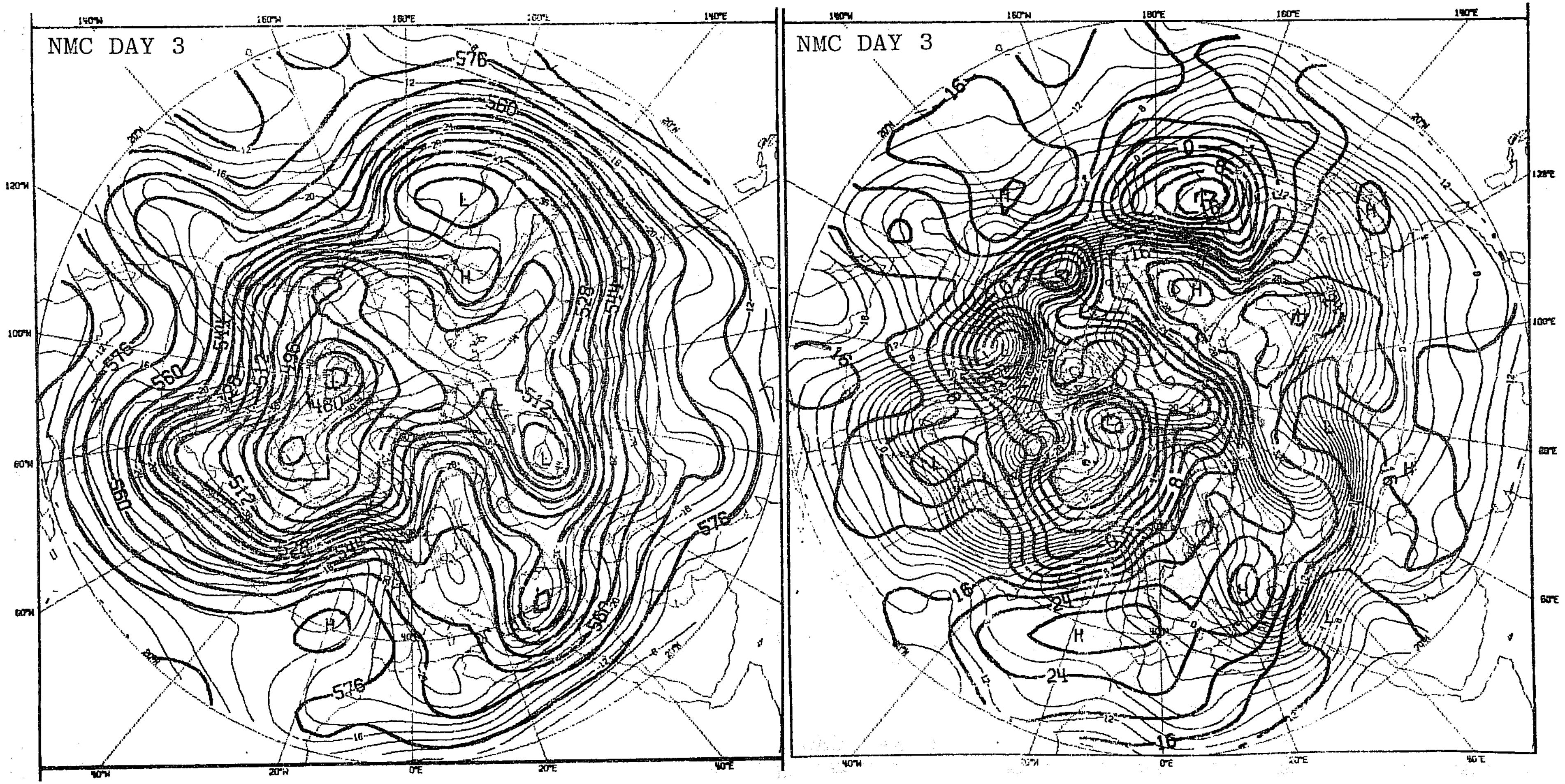


Fig.5.3.3 NMC analyses corresponding to day 3 of forecast from 00Z on 6.2.76. Left: 500 mb height and temperature fields. Right: 1000 mb height and 850 mb temperature fields. Contour intervals as in Fig.5.3.1.

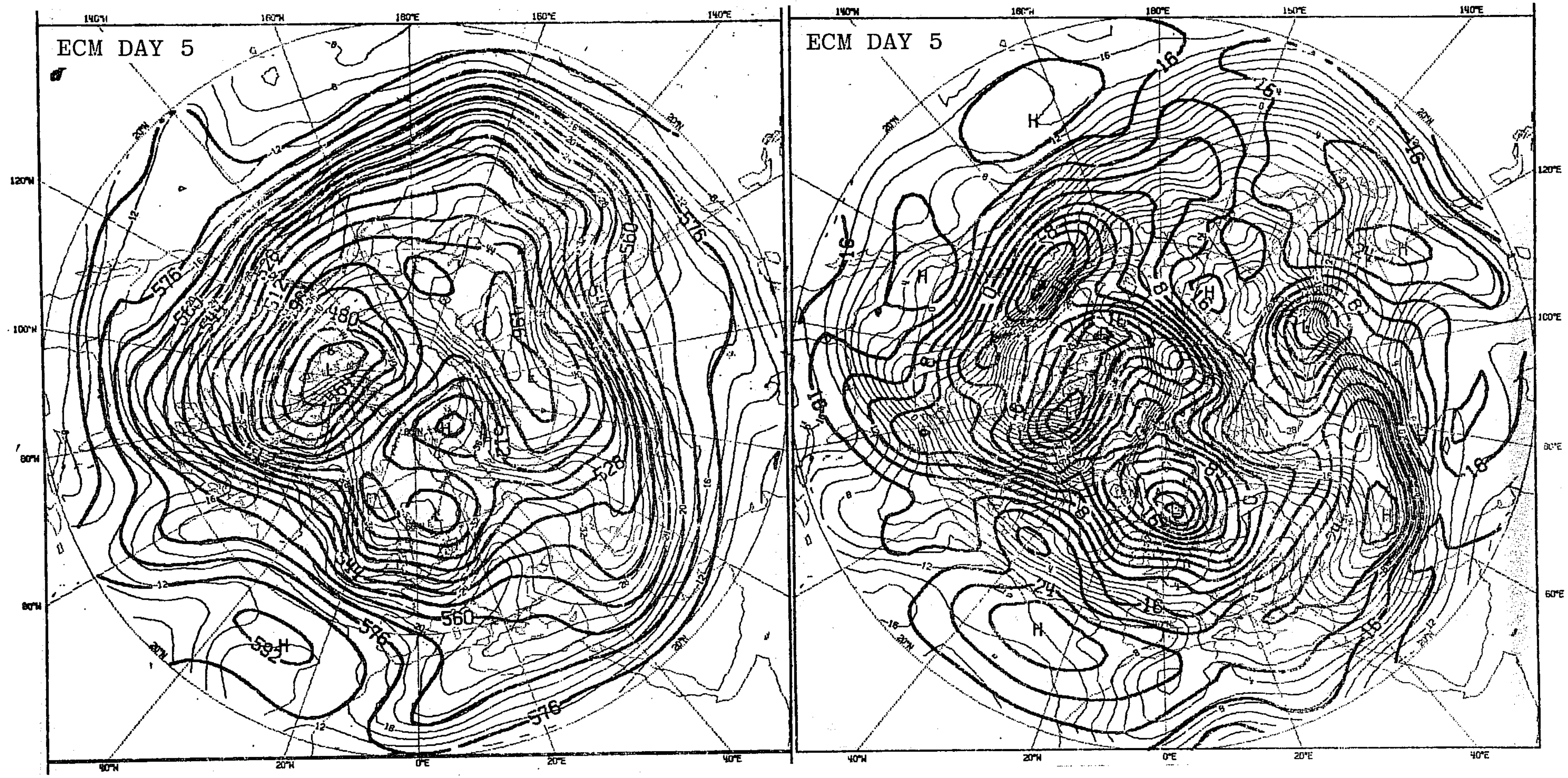


Fig.5.3.4 Left: 500 mb height and 500 mb temperature fields for day 5 of ECM forecast from 00Z on 6.2.76. Right: 1000 mb height and 850 mb temperature fields for day 5 of ECM forecast from 00Z on 6.2.76. Contour intervals as in Fig.5.3.1.

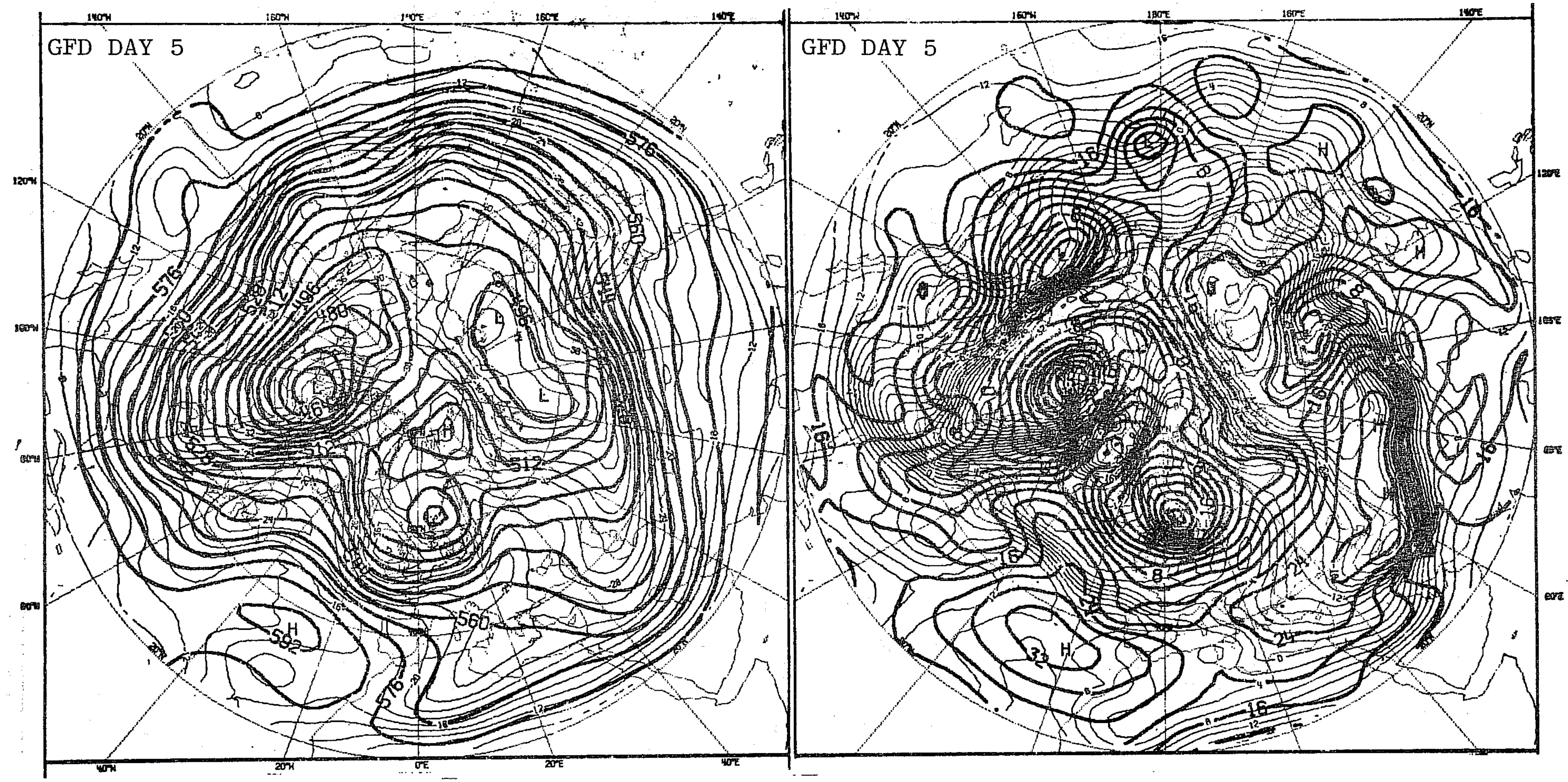


Fig.5.3.5 Left: 500 mb height and 500 mb temperature fields for day 5 of GFD forecast from 00Z on 6.2.76. Right: 1000 mb height and 850 mb temperature fields for day 5 of GFD forecast from 00Z on 6.2.76. Contour intervals as in Fig.5.3.1.

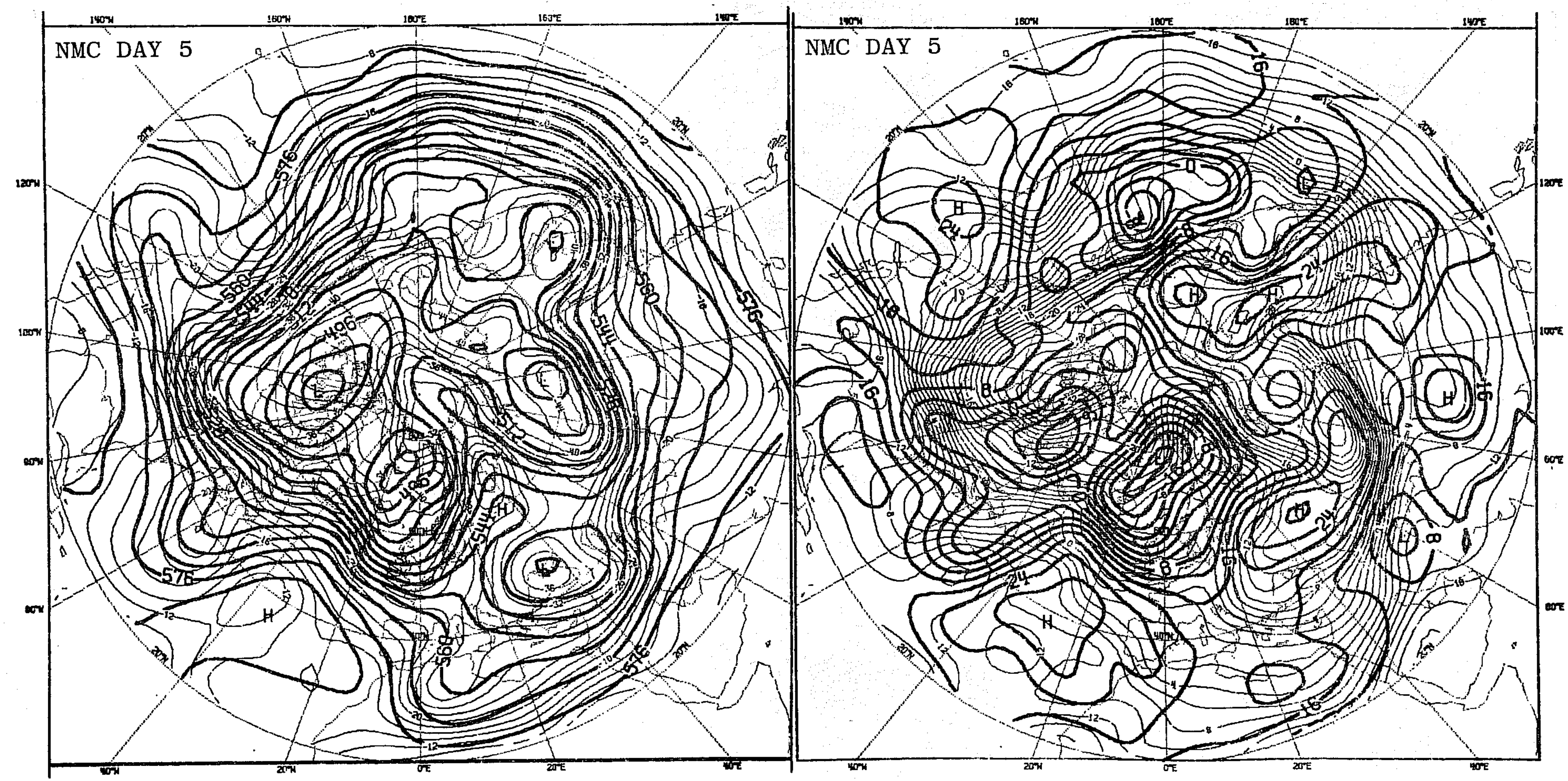
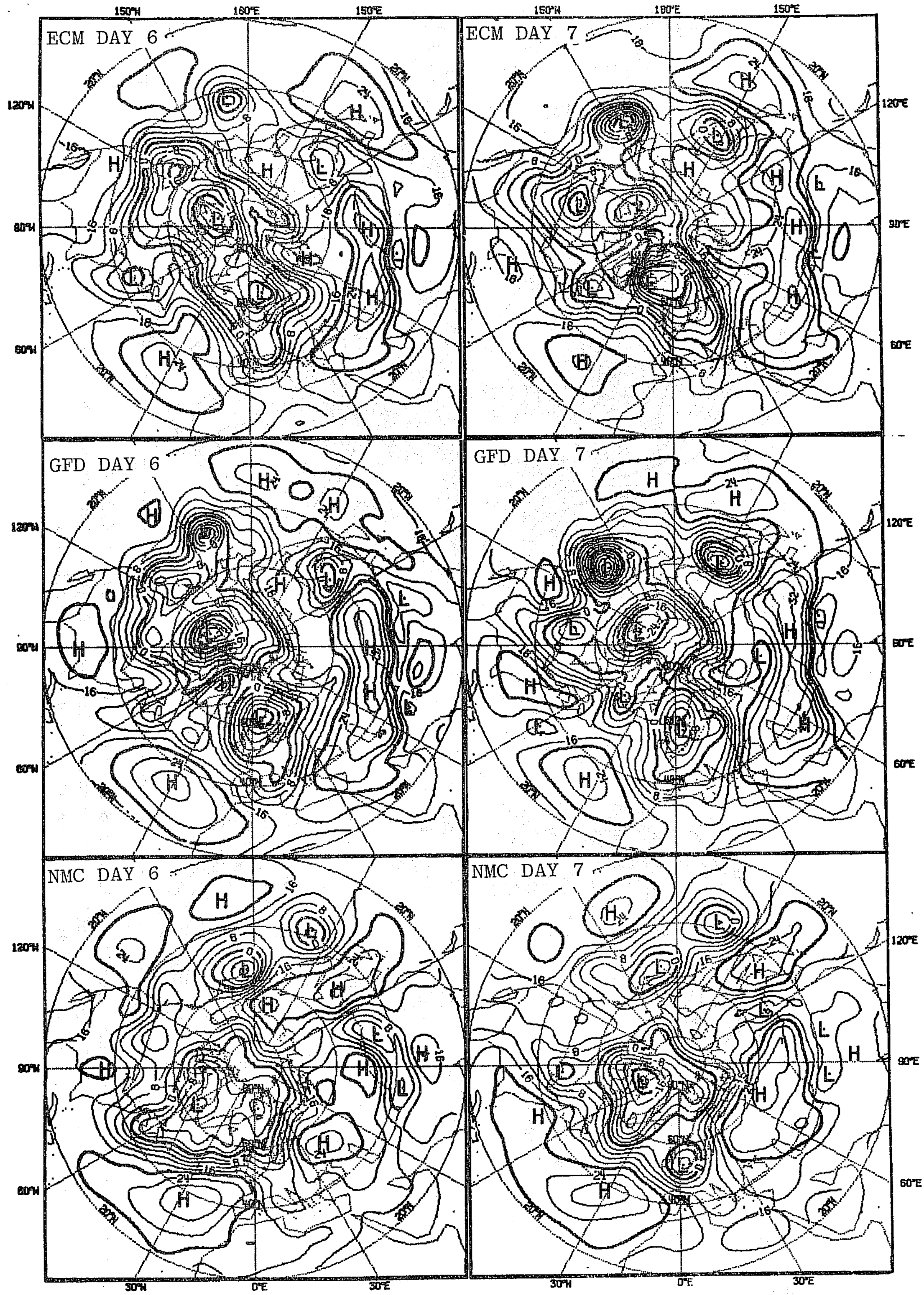


Fig.5.3.6 NMC analyses corresponding to day 5 of forecast from 00Z on 6.2.76. Left: 500 mb height and 500 mb temperature fields. Right: 1000 mb height and 850 mb temperature fields. Contour intervals as in Fig.5.3.1.


ECM DAY 6

ECM DAY 7

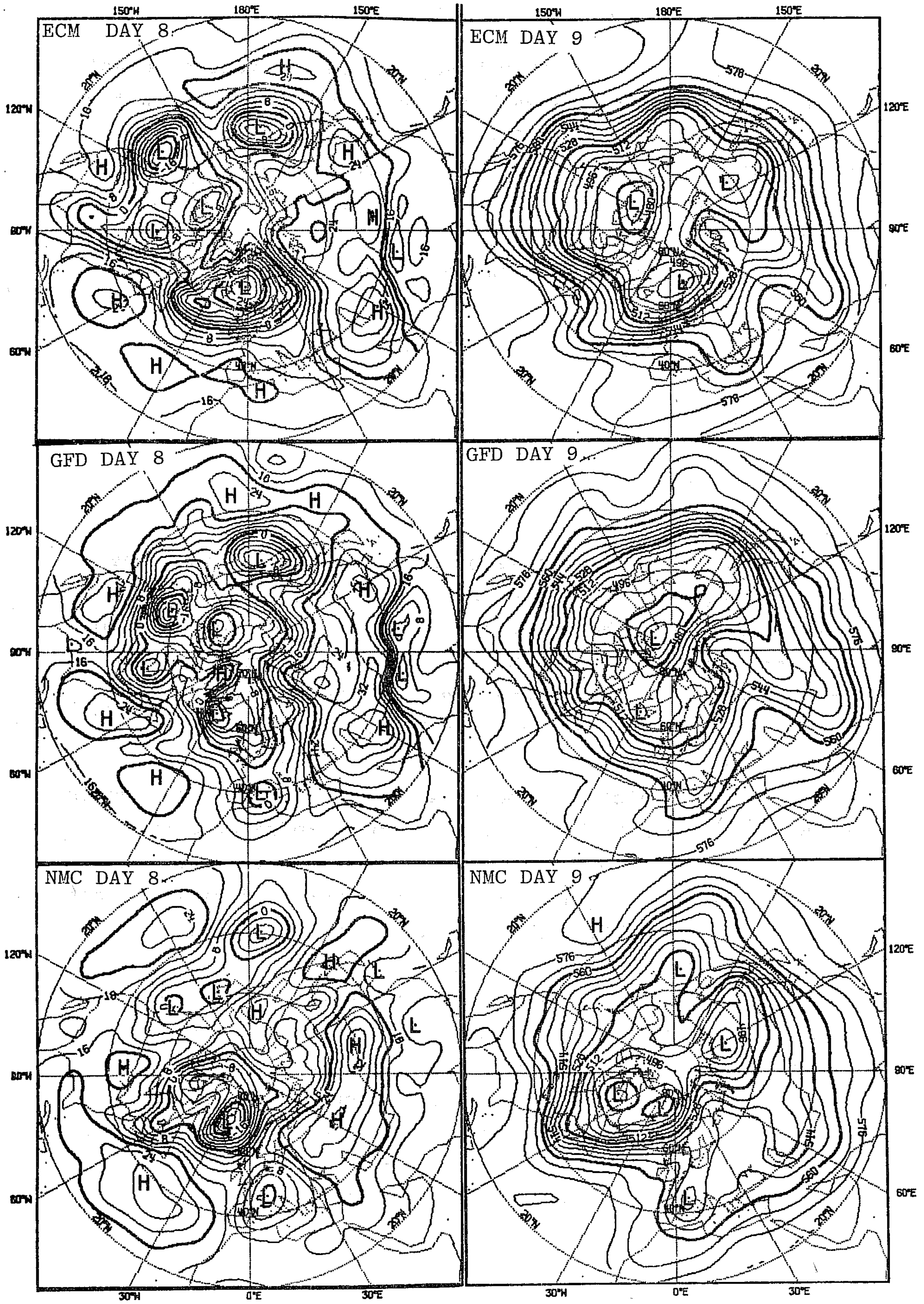
GFD DAY 6

GFD DAY 7

NMC DAY 6

NMC DAY 7

Fig.5.3.7 Left: 1000 mb height fields for day 6 of ECM and GFD forecasts from 00Z on 6.2.76 and corresponding NMC analysis. Right: 1000 mb height fields for day 7 of ECM and GFD forecasts from 00Z on 6.2.76 and corresponding NMC analysis. Contour interval as in Fig.5.3.1.



ECM DAY 8

ECM DAY 9

GFD DAY 8

GFD DAY 9

NMC DAY 8

NMC DAY 9

Fig.5.3.8 Left: 1000 mb height fields for day 8 of ECM and GFD forecasts from 00Z on 6.2.76 and corresponding NMC analysis. Right: 500 mb height fields for day 9 of ECM and GFD forecasts from 00Z on 6.2.76 and corresponding NMC analysis. Contour intervals as in Fig.5.3.1.

East of China, on the other hand, the major troughs do not seem to be forecast to reach the same amplitudes as in the analysis. During days 3-5 both models become too zonal, perhaps because at times the temperature field was not in correct phase with the contour field during the period.

5.4 Case 3 - forecasts from 15.2.1976.

a) Synoptic summary See 5.2a

b) Forecast evaluation

Day 0-3

After 24 hours, both models predict a sharp 500 mb trough on the Atlantic to be about 5° to the west of its true position. Many other features are also displaced by about the same amount. On day 2 the NMC analyses show a large anticyclone to be developing over Northern Europe with warm air advection on its western flank. The forecast temperatures are in good agreement with the observed, but the high is predicted to be slightly too weak by the models at 500 mb. Near 35°N 50°W the 500 mb temperatures are predicted to be about 2 degrees warmer in both forecasts, whereas the 850 mb temperatures are lower by 5 or 6 degrees (Fig. 5.4.1-2). This erroneous stabilising effect is also observed in another cold air outbreak south of 50°N over the Pacific. Over the eastern United States of America, however, this effect is not observed. Over China south of 35°N, both models make the atmosphere more unstable than observed, heating the lower atmosphere by an excess of 3-5 degrees. On day 3 both forecasts are very similar and they fail to fill an extensive depression not far from the North Pole, in turn associated with the suppression of the

European high and a strong westerly current between these two systems. In general, however, both forecasts are showing considerable skill at this stage.

Day 4

The European high is now correctly displaced to western Russia by both forecasts (Fig.5.4.3-4). It does not display sufficiently great an amplitude over Finland and northwestern Russia in either forecast at 500 mb. Temperatures are still too low on the Atlantic far south of Greenland, particularly at 850 mb where the forecast temperatures are around 6-8 degrees below the analyses (Fig. 5.4.4). However, northwest of Ireland there is a small but definite temperature excess forecast only at 850 mb mostly by the ECM model. Both forecasts are too slow with a 500 mb trough over the United States of America, but otherwise the contour pattern is advected too quickly eastward (Fig.5.4.3).
Both forecasts are acceptable for most parts of the hemisphere and certainly over Europe.

Day 5

On day 5 there is a very good 500 mb forecast by both models, of the trough from Greenland to the Azores, but at the surface the observed strong gradient and curvature west of Spain is a little weak in the ECM forecast and almost disappears in the GFD forecast.This is associated with lower temperature advection and gradients in connection with the low. (Fig.5.4.3-7).

Day 6

The predicted evolution over Europe and the Atlantic is still quite acceptable. The two forecasts continue to be fairly similar,but the ECM 500 mb forecast has correctly filled the Siberian low, while the GFD forecast did not fill it enough. On the other hand, the upper trough approaching Spain is over-intensified in the ECM forecast where the GFD model is superior. Figure 5.4.8 shows the ECM 500 mb forecast and the corresponding analysis.

Day 7

A very strong outflow of cold 850 mb air occurs in the ECM forecast south of the Caspian Sea and there is a surface high which is far too strong. The GFD model predicts the temperatures up to 14 degrees higher, although the differences at 500 mb are only 3-4 degrees. (Figs.5.4.9-11). The 500 mb predictions are preferable to the surface forecasts in this case.
Over Europe both forecasts are still generally acceptable.

Day 8-10

On day 8 the GFD 1000 mb forecast is in more correct phase than the ECM forecast which is slower (Fig.5.4.12). For most of Europe the ECM forecast is preferable since it predicts the high more accurately, although strong southeasterlies are incorrectly forecast on the southern side of the high. On day 9 the ECM forecast develops three waves over the Pacific instead of two. (Fig. 5.4.13). An almost contrary situation was found 3 days earlier when the analysis displayed three well developed troughs from the United States of America to Japan and the forecasts one less (Fig.5.4.8). These deficiencies earlier in the forecasts now affect the Atlantic and Europe. However, although somewhat out of phase during the last two days, the forecast is still generally reasonable, with the exception of the failure to correctly forecast the redevelopment of the anticyclone west of Ireland. The same applies also for the GFD forecast. This error was found at different stages in the two previous February cases as well in both models. Now, once again both models develop a strong zonal jet with only small perturbations that progress eastward towards towards Spain instead of towards Scotland and Scandinavia. The 500 mb flow for day 10 shows that both models seem to have recovered and the GFD model in particular is significantly improved, taking into account only Asia and the Pacific (Fig.5.4.14,left).

c) Time-mean maps

Figure 5.4.15 shows time-mean maps averaged over 3 days at the end of the forecast period, for the ECM forecast and for the analyses. The incorrect forecast of a slow-moving trough over the Iberian peninsula where the analyses show a ridge, resulted in the strong 1000 mb southeasterly flow over central and southeastern Europe, and a southerly, rather than south-westerly flow over northwestern Europe. The 500 mb mean flow in the analysis (Fig.5.4.15,left) shows a trough near 80°W. This is associated with the eastward movement of a well-developed trough which crossed 80°W between day 7 (Fig. 5.4.9-11) and day 9 (Fig.5.4.13). The ECM mean flow (Fig.5.4.15,left) is much more zonal in this area, reflecting the weaker troughs in the forecast (Figs.5.4.9-11 and 5.4.13). However, most of the large-scale features of the flow (the blocking ridge over eastern Europe, the low centred near 100°E for example) have been well forecast. Transient features of the flow (e.g. the 1000 mb low on day 9 centred near 60°W, Fig. 5.4.13) are not of course represented well on the 3-day time-mean chart. Several of the features mentioned here will be referred to again in the discussion of systematic errors (Section 8).

d) Precipitation

In Section 5.2d we noted that both models appear generally to over-estimate the extent of precipitation with respect to its areal distribution and the same remarks apply here. Figures 5.4.16 and 19 show the observed precipitation from days 1 and 2 of the forecast, Figs. 5.4.17-18, 20-21 show the precipitation as forecast by the models. The ECM forecast underestimates the amounts of convective precipitation in the Mediterranean on day 1 of the forecast, where the GFD forecast is more accurate. The maximum intensity of precipitation is low in both forecasts. Over northern Scandinavia the ECM forecast is slightly better, where it predicts more large-scale amounts than the GFD model. The synoptic situation here suggests mostly large-scale precipitation with reinforecement due to lifting over the mountains. Neither of the models forecasts the convective precipitation penetrating over the Iberian peninsula on day 2, while no precipitation was observed over Iceland on day 1, suggesting perhaps that the models are forecasting too much precipitation over the oceans.

e) Summary of forecasts from February 15 1976

These forecasts from February 15 1976 are perhaps the best among the selected experiments, particularly when considering the 500 mb flow rather than the 850 mb and 1000 mb and considering the evolution over Europe. Except for small scale details and temperatures, both models provide some useful guidance for 8-9 days for Europe. However, for other parts of the hemisphere, such as the Pacific and North America, neither forecast can be used more than 5-6 days, although they improve during the last two days.

A similar feature to the two previous February cases is the trend of the depressions to approach southern Europe at the end of the forecast instead of being forced to move on much more northerly tracks.

Continued on page 78

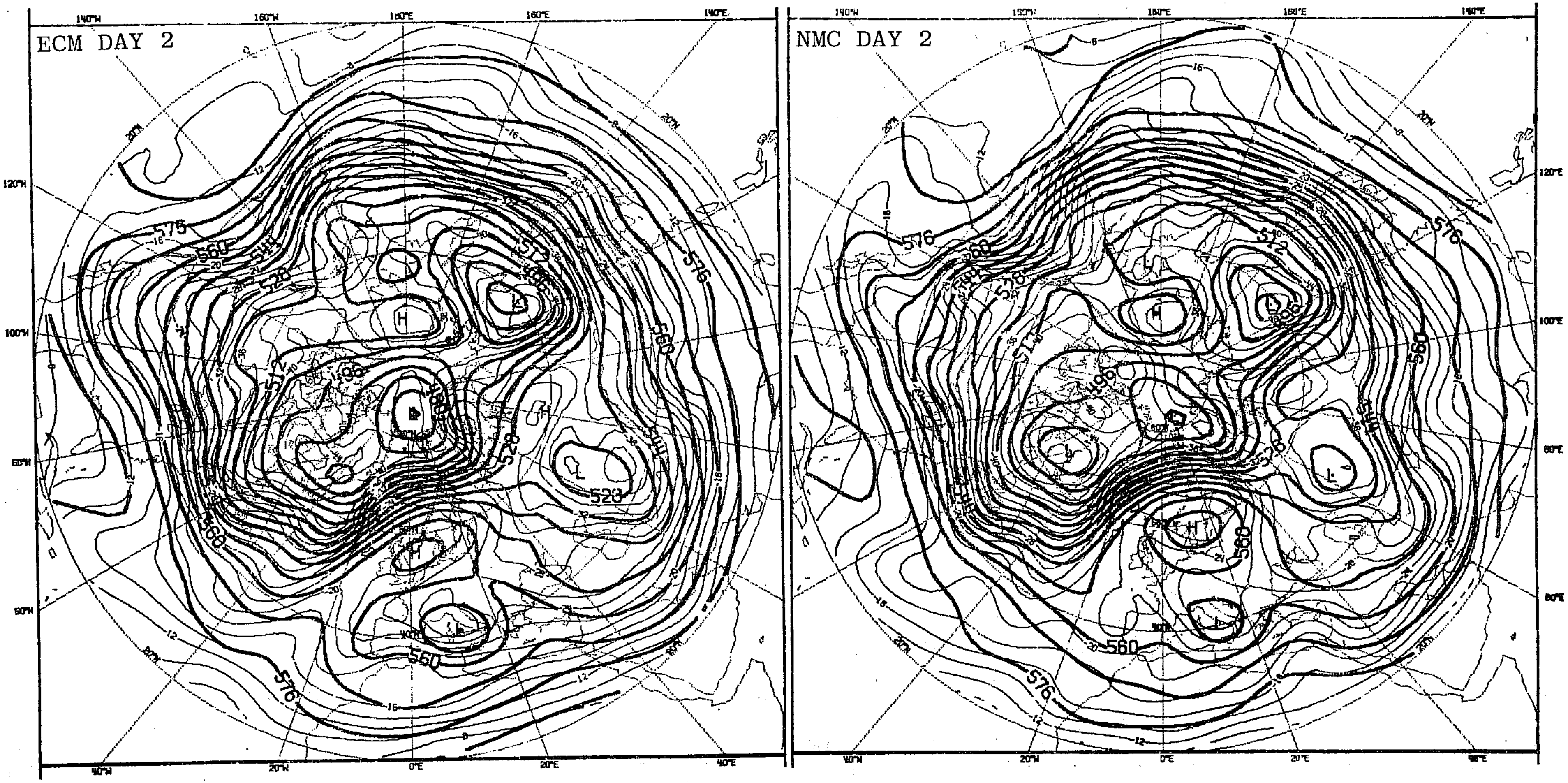


Fig.5.4.1 500 mb height fields (thick lines, contour interval: 8 dekametres) and 500 mb temperature fields (thin lines, contour interval: 2 °C) for day 2. Left: ECM forecast from 00Z on 15.2.76. Right: corresponding NMC analysis.

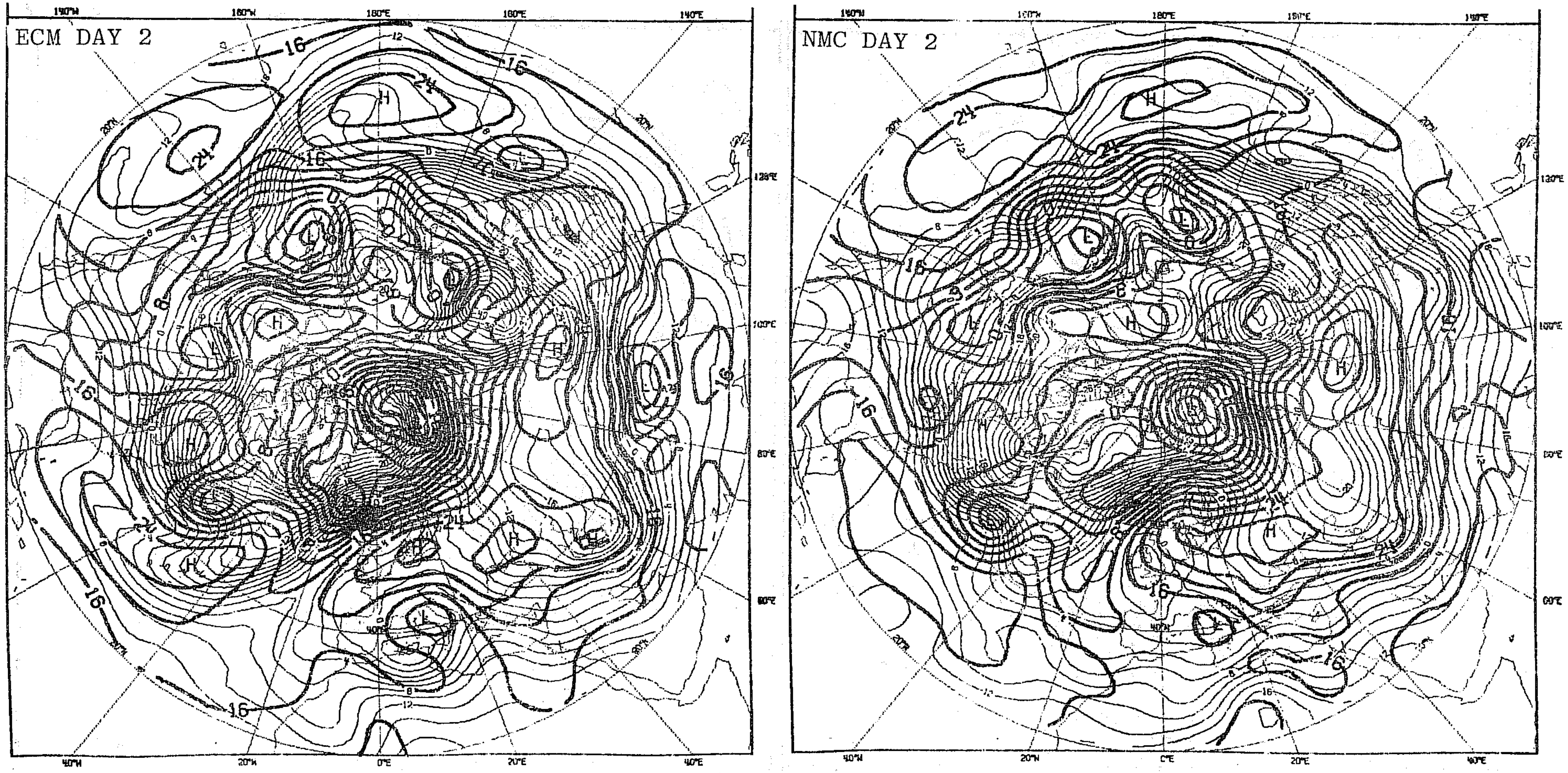


Fig.5.4.2 1000 mb height fields (thick lines, contour interval: 4 dekametres) and 850 mb temperature fields (thin lines, contour interval: 2 °C) for day 2. Left: ECM forecast from 00Z on 15.2.76. Right: corresponding NMC analysis.

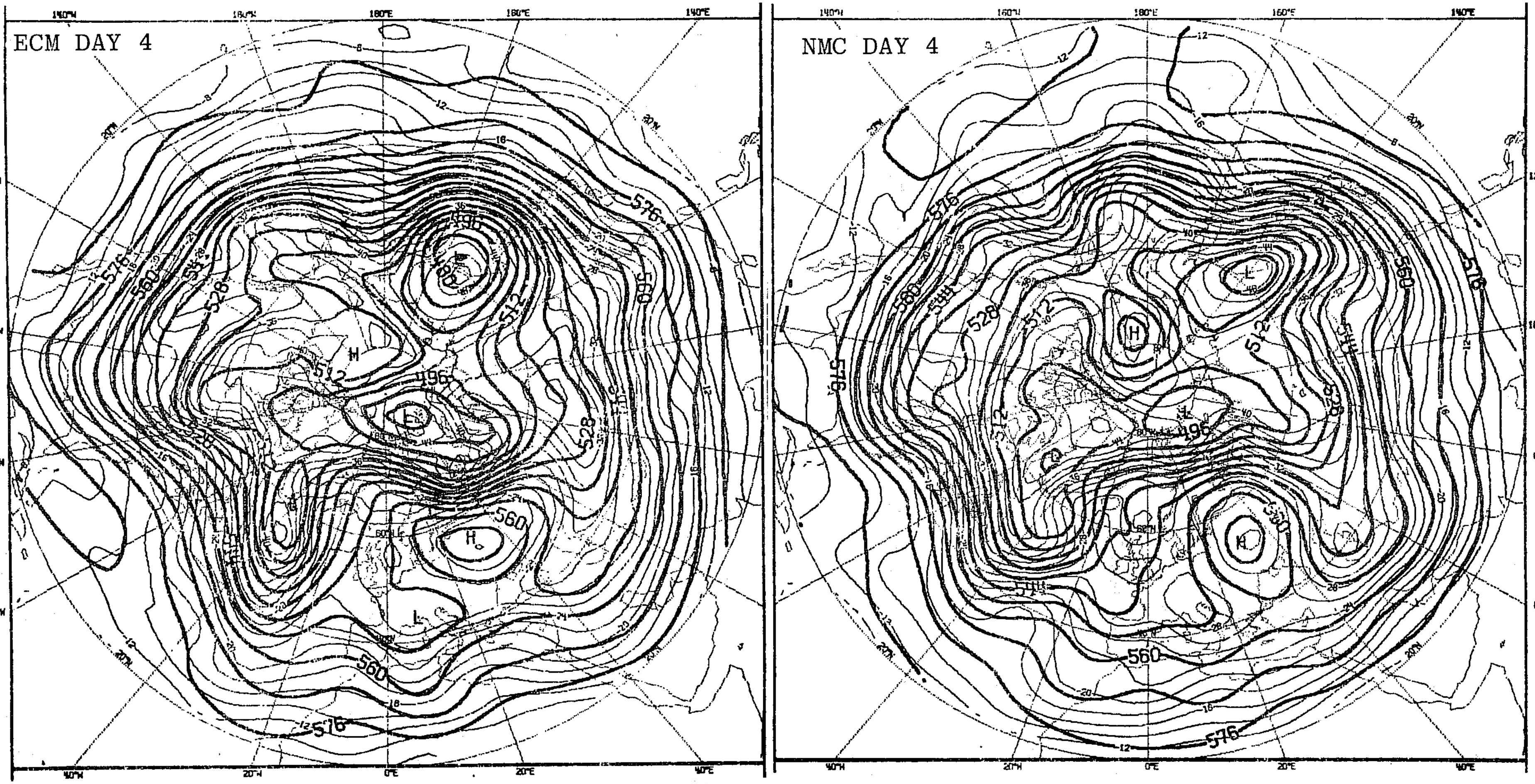


Fig.5.4.3 500 mb height and temperature fields for day 4. Left: ECM forecast from 00Z on 15.2.76. Right: corresponding NMC analysis. Contour interval as in Fig.5.4.1.

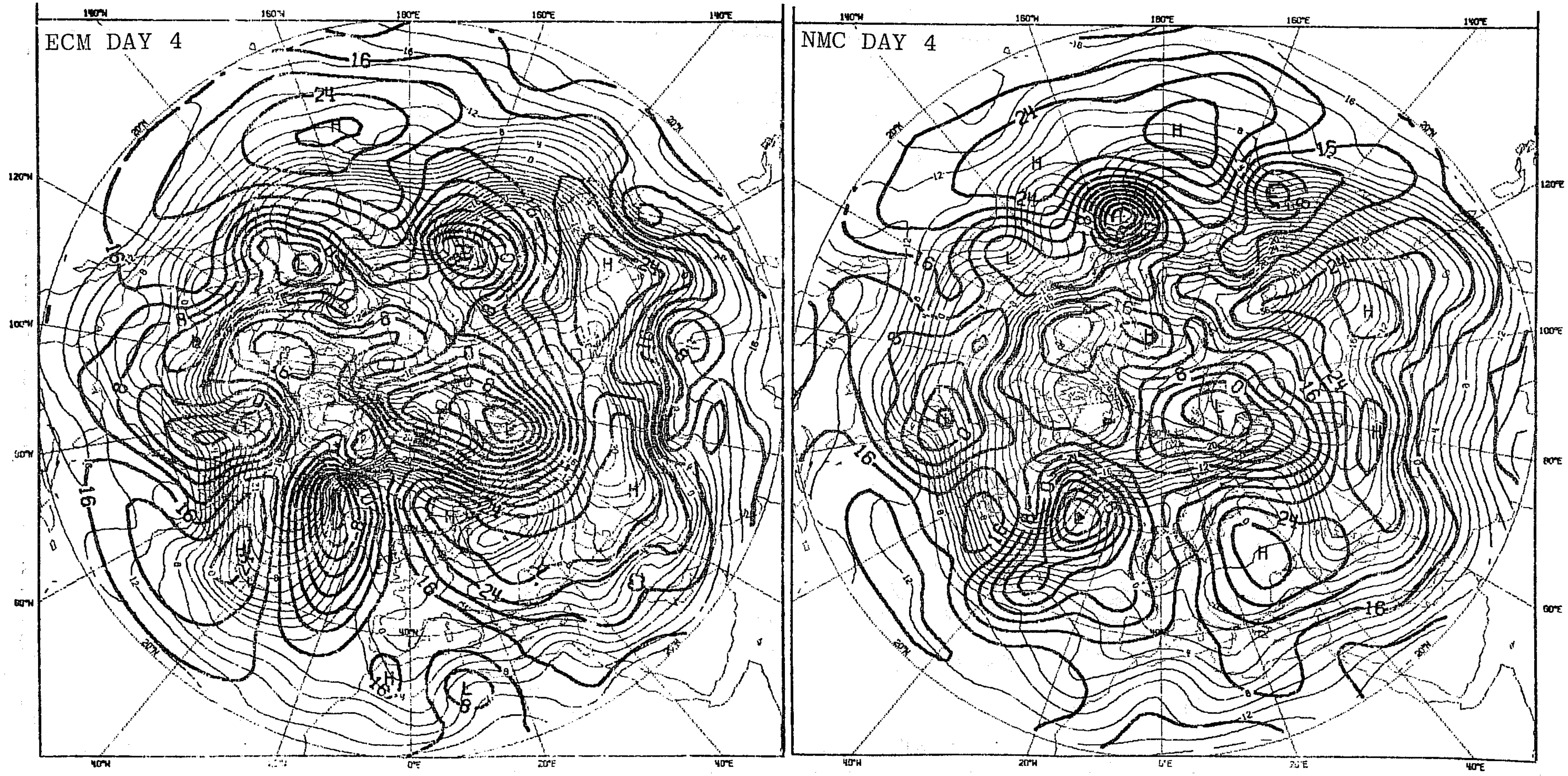


Fig.5.4.4. 1000 mb height and 850 mb temperature fields for day 4. Left: ECM forecast from 00Z on 15.2.76. Right: corresponding NMC analysis. Contour intervals as in Fig.5.4.2.

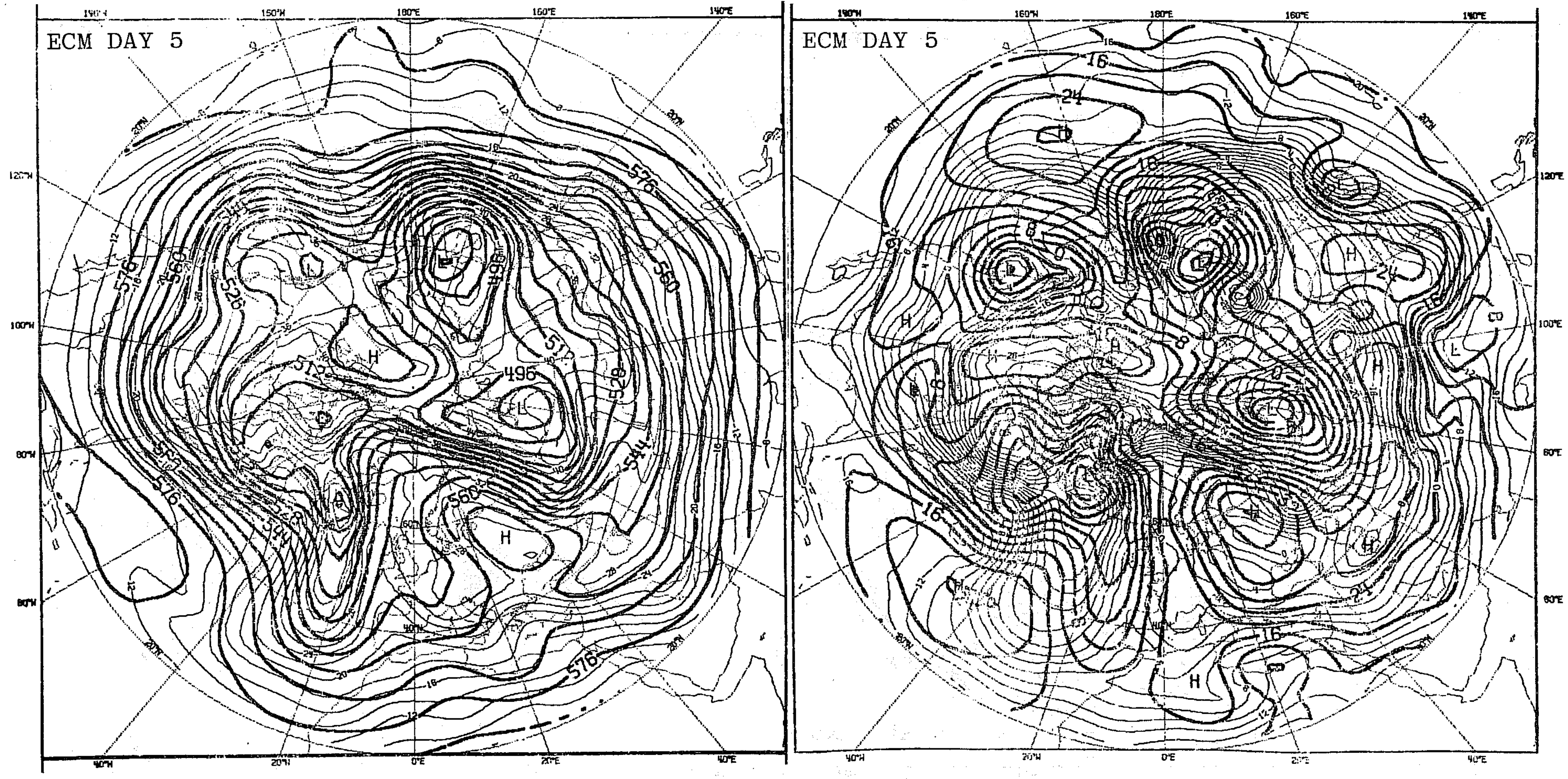


Fig.5.4.5. Left: 500 mb height and temperature fields for day 5 of ECM forecast from 00Z on 15.2.76. Right: 1000 mb height and 850 mb temperature fields for day 5 of ECM forecast from 00Z on 15.2.76. Contour intervals as in Figs.5.4.1 and 2.

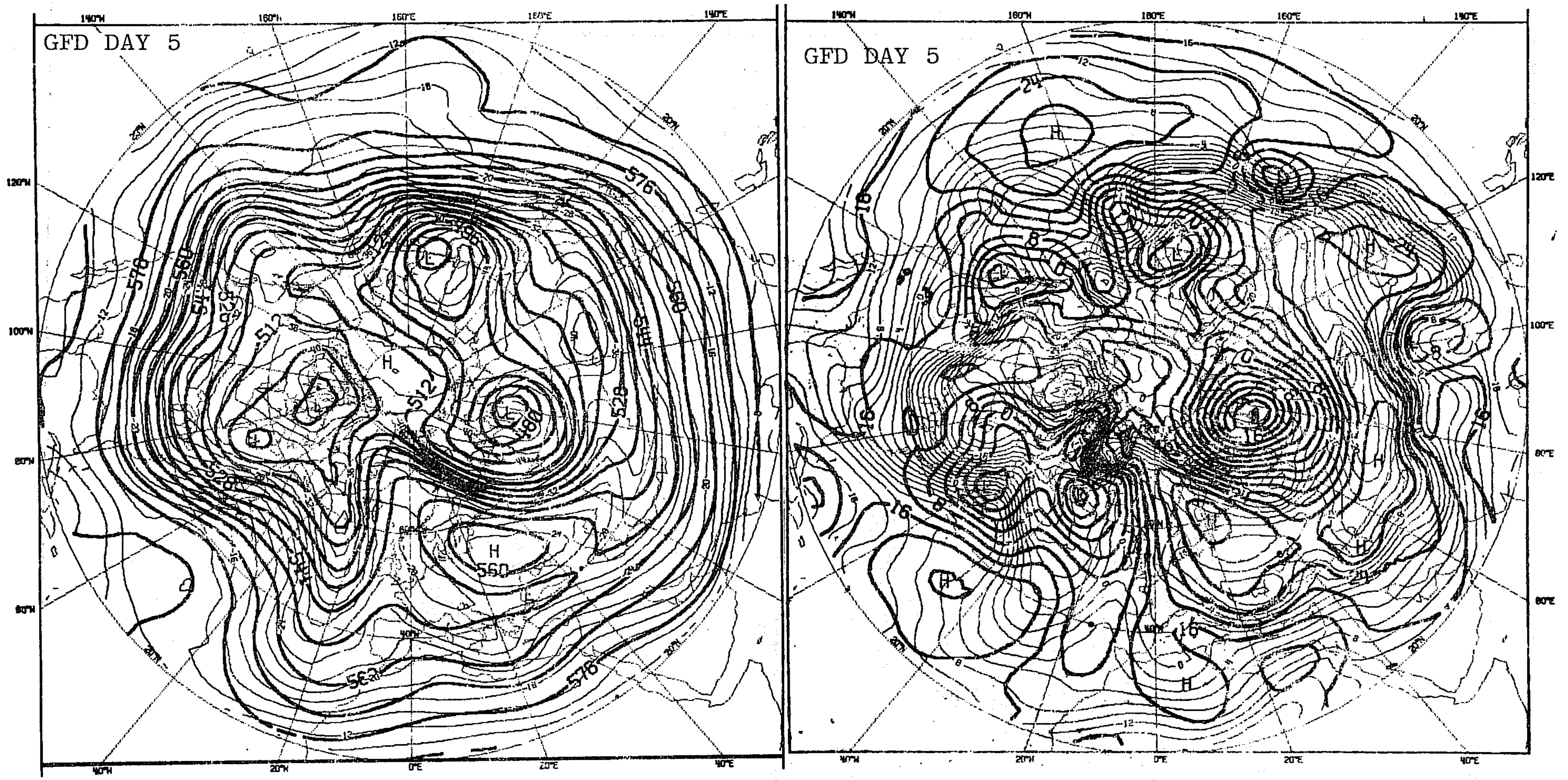


Fig.5.4.6 Left: 500 mb height and temperature fields for day 5 of GFD forecast from 00Z on 15.2.76. Right: 1000 mb height and 850 mb temperature fields for day 5 of GFD forecast from 00Z on 15.2.76. Contour intervals as in Figs.5.4.1 and 2.

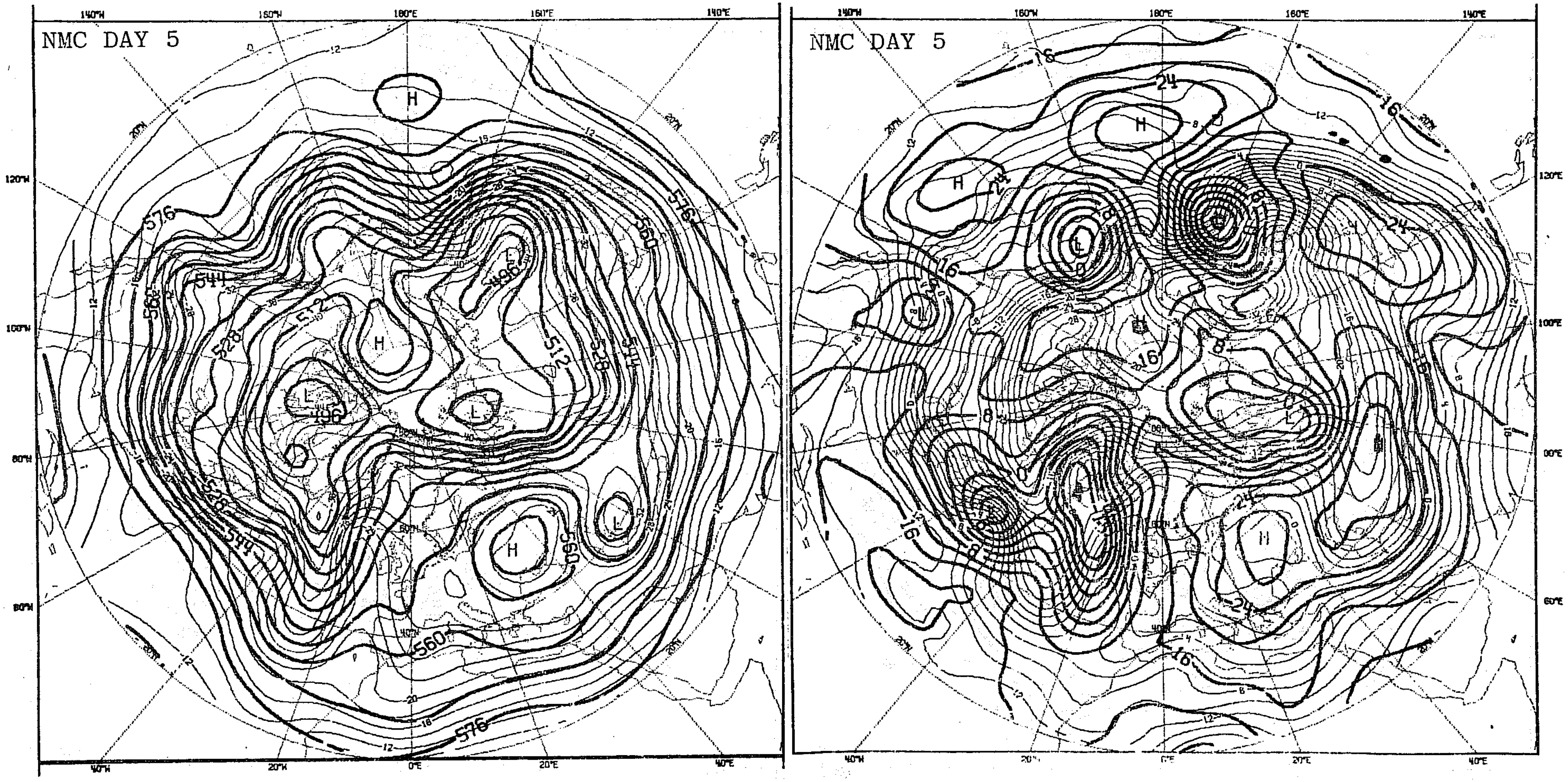


Fig.5.4.7 NMC analyses corresponding to day 5 of forecast from 00Z on 15.2.76. Left: 500 mb height and temperature fields. Right: 1000 mb height and 850 mb temperature fields. Contour intervals as in Figs.5.4.1 and 2.

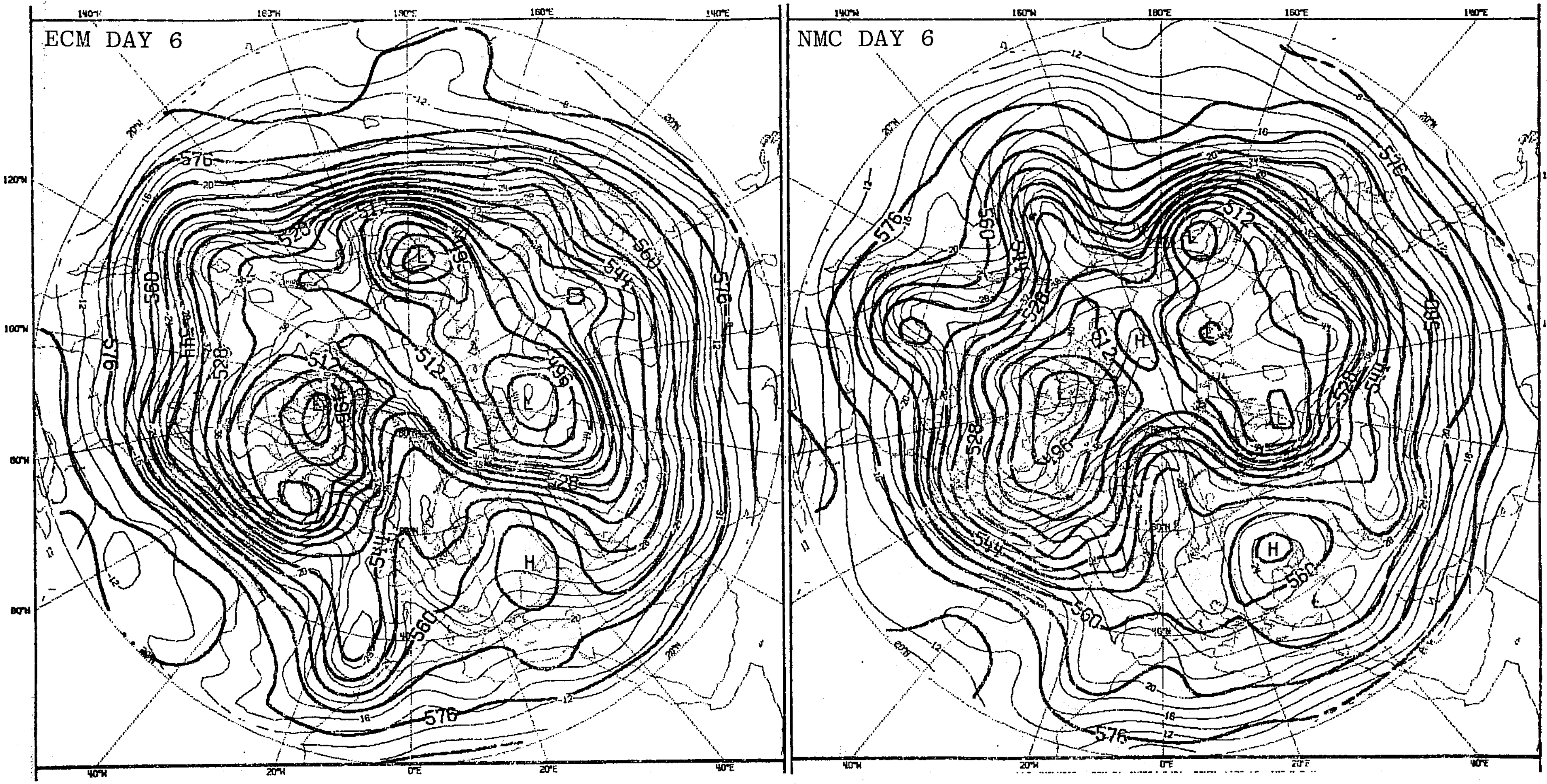


Fig.5.4.8 500 mb height and temperature fields for day 6. Left: ECM forecast from 00Z on 15.2.76. Right: corresponding NMC analysis. Contour intervals as in Fig.5.4.1.

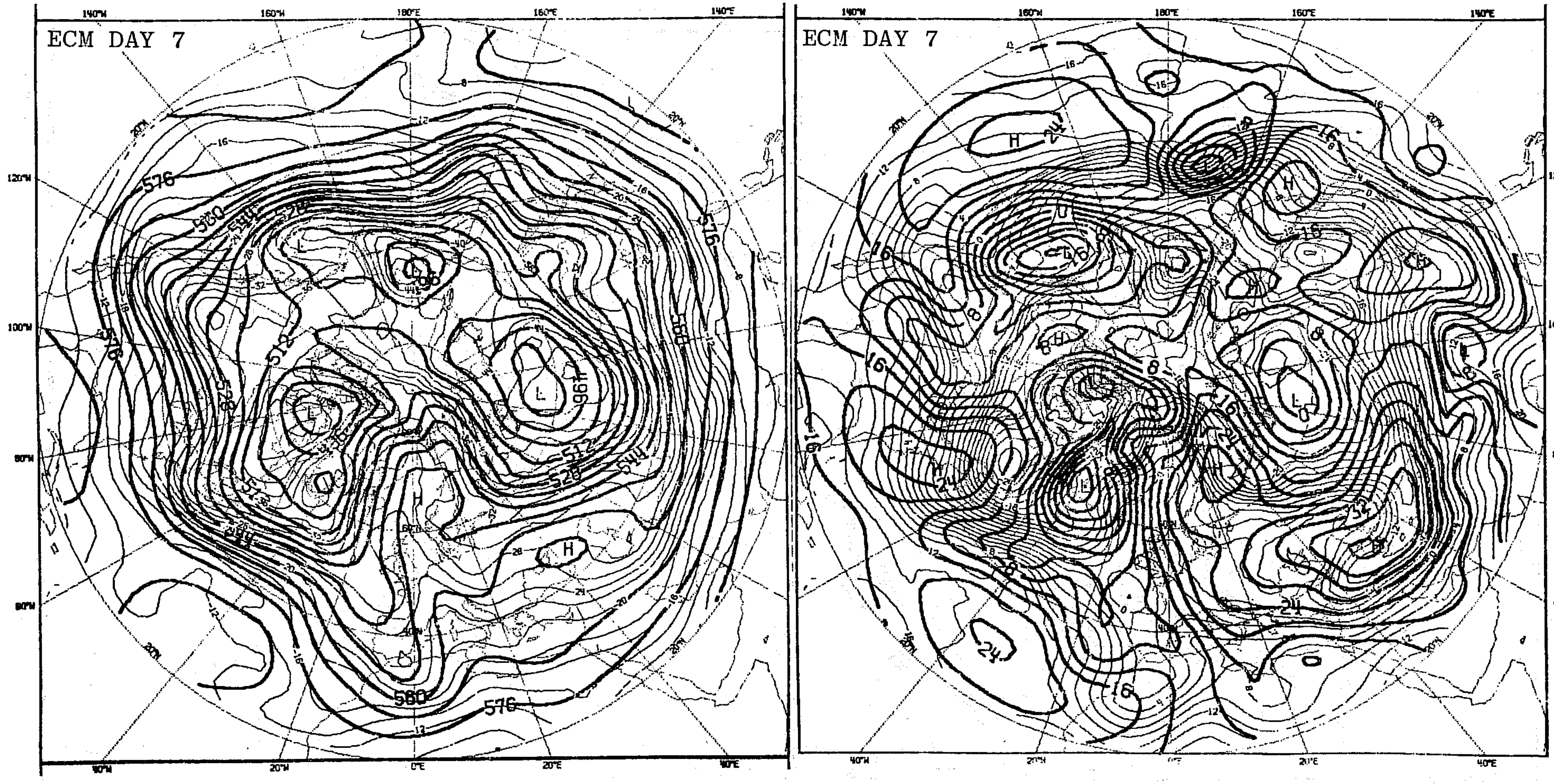


Fig.5.4.9 Left: 500 mb height and temperature fields for day 7 of ECM forecast from 00Z on 15.2.76. Right: 1000 mb height and 850 mb temperature fields for day 7 of ECM forecast from 00Z on 15.2.76. Contour intervals as in Figs.5.4.1 and 2.

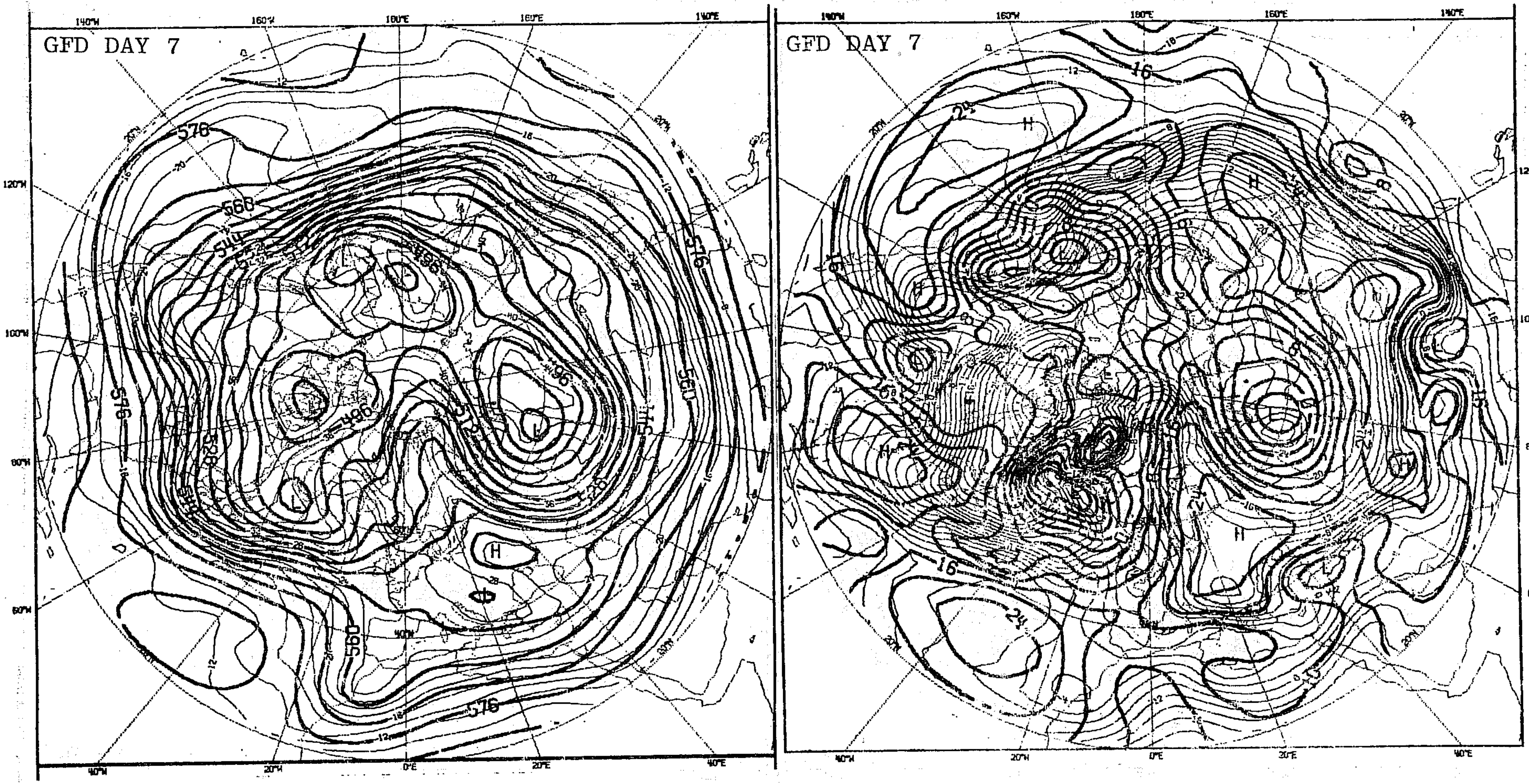


Fig.5.4.10 Left: 500 mb height and temperature fields for day 7 of GFD forecast from 00Z on 15.2.76. Right: 1000 mb height and 850 mb temperature fields for day 7 of GFD forecast from 00Z on 15.2.76. Contour intervals as in Figs. 5.4.1 and 2.

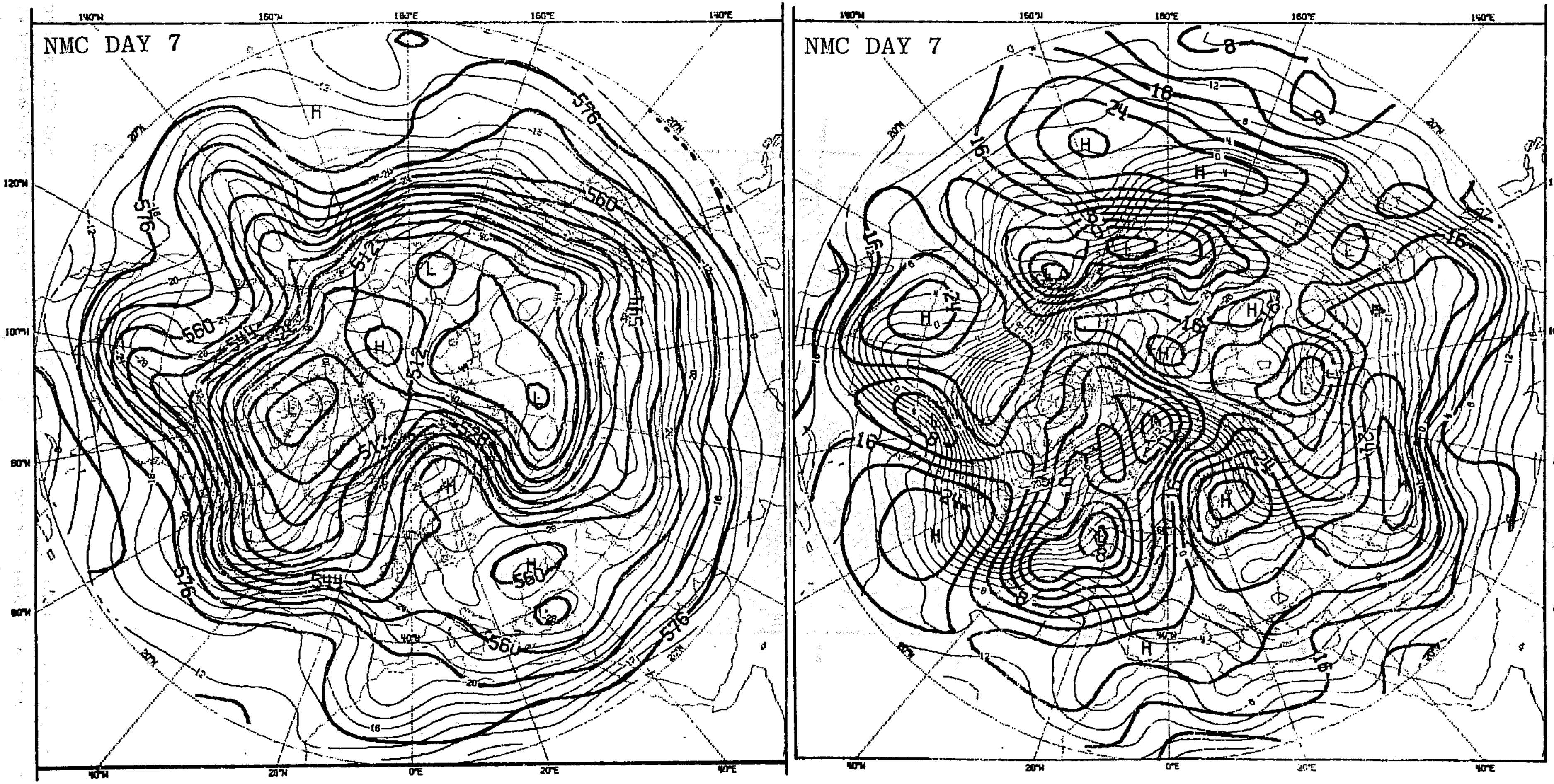


Fig.5.4.11 NMC analyses corresponding to day 7 of forecast from 15.2.76. Right: 1000 mb height and 850 mb temperature fields. Contour intervals as in Figs. 5.4.1 and 2.

Day 8
5 2

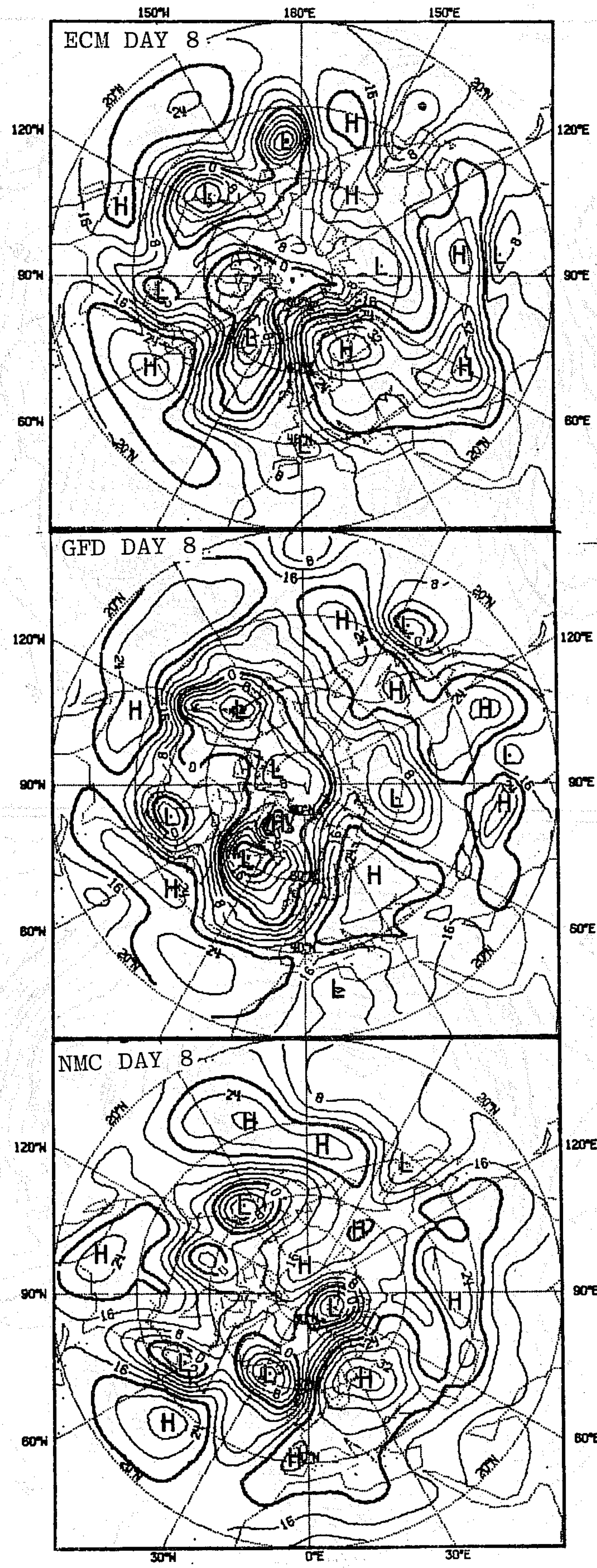


Fig.5.4.12 1000 mb height fields for day 8 of ECM and GFD forecasts from 00Z on 15.2.76 and corresponding NMC analysis. Contour interval as in Fig.5.4.2.

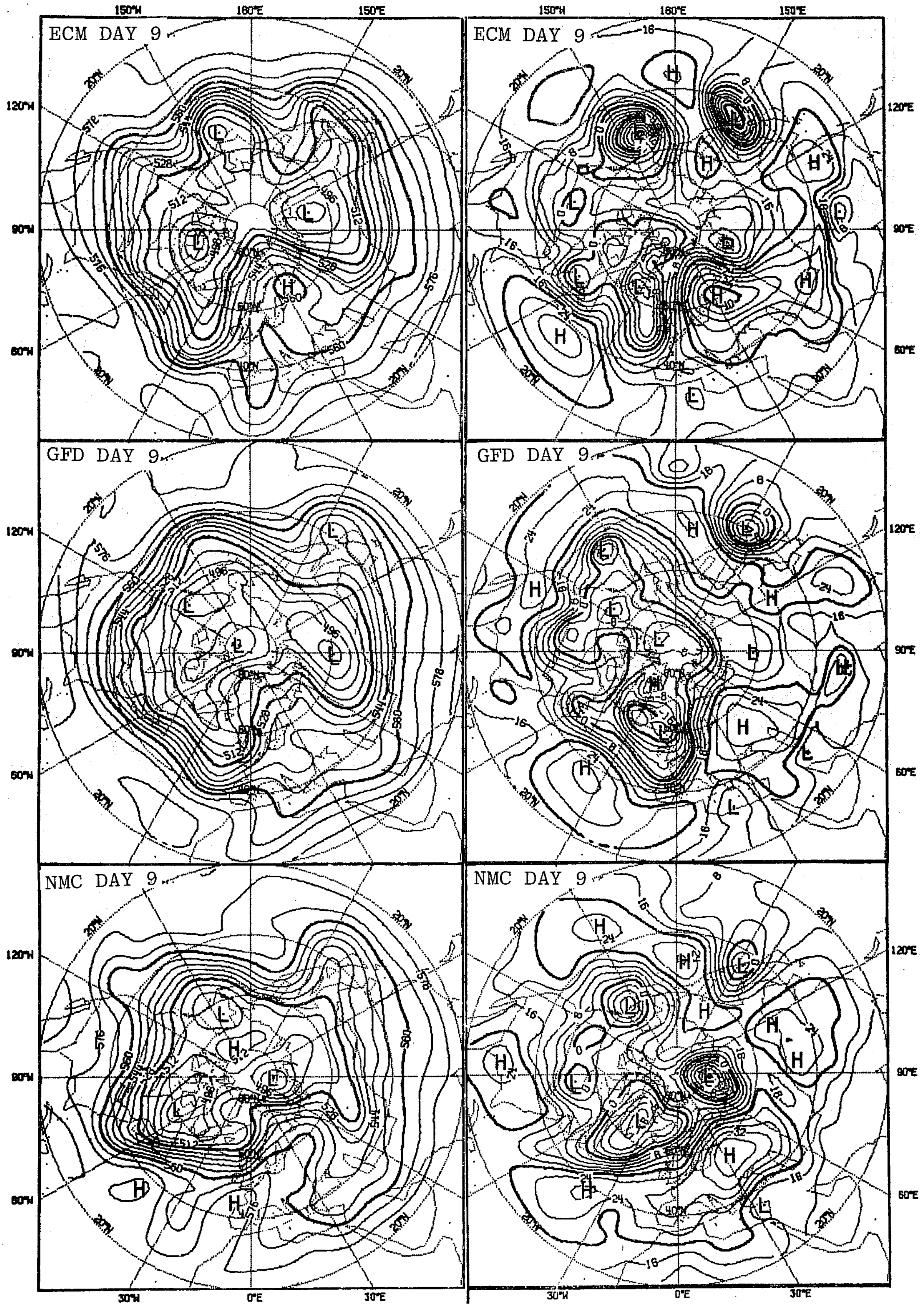


Fig.5.4.13 Left: 500 mb height fields for day 9 of ECM and GFD forecasts from 00Z on 15.2.76 and corresponding NMC analysis.
Right: 1000 mb height fields for day 9 of ECM and GFD forecasts from 00Z on 15.2.76 and corresponding NMC analysis.
Contour intervals as in Figs.5.4.1 and 2.

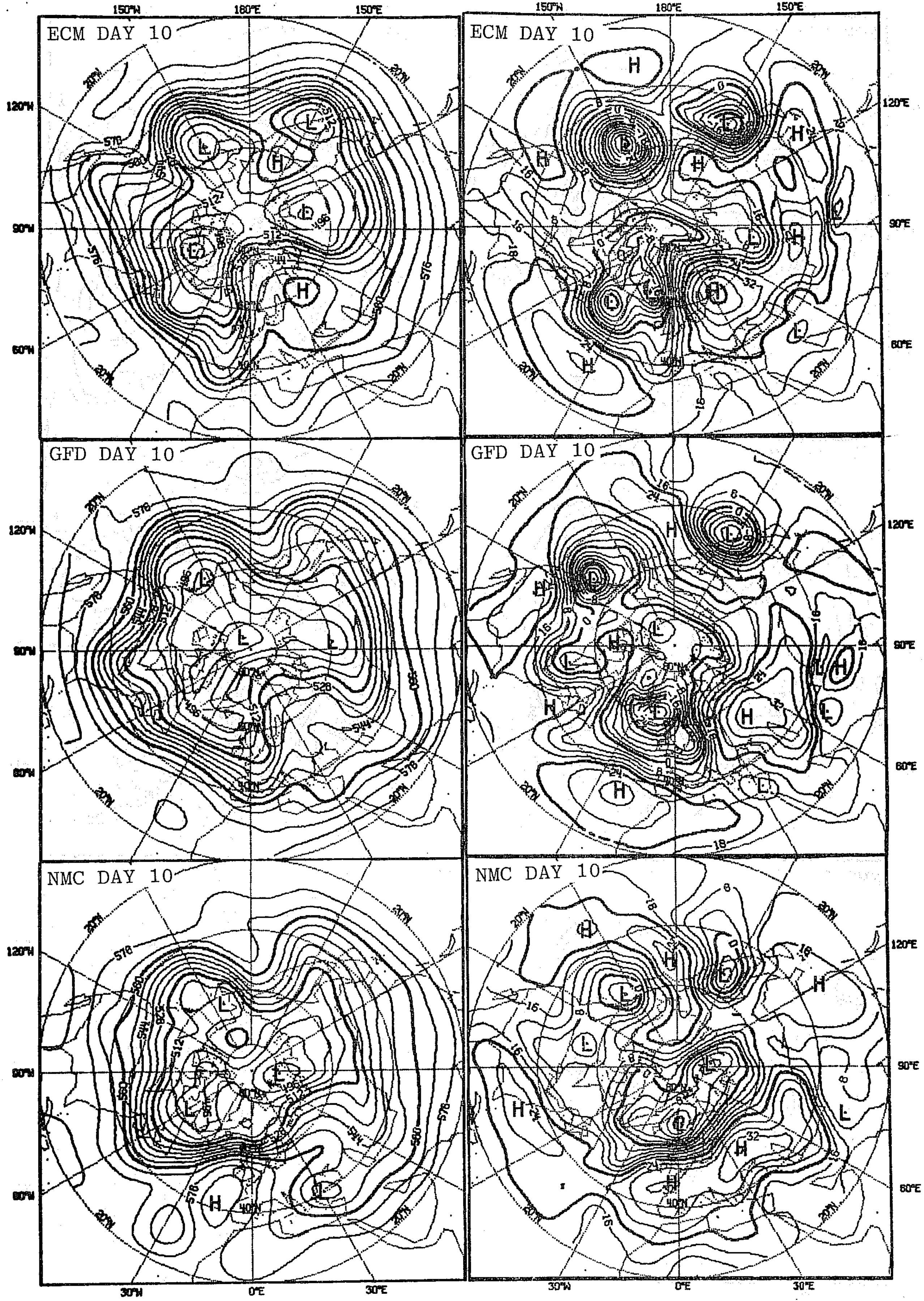


Fig.5.4.14 Left: 500 mb height fields for day 10 of ECM and GFD forecasts from 00Z on 15.2.76 and corresponding NMC analysis. Right: 1000 mb height fields for day 10 of ECM and GFD forecasts from 00Z on 15.2.76 and corresponding NMC analysis. Contour intervals as in

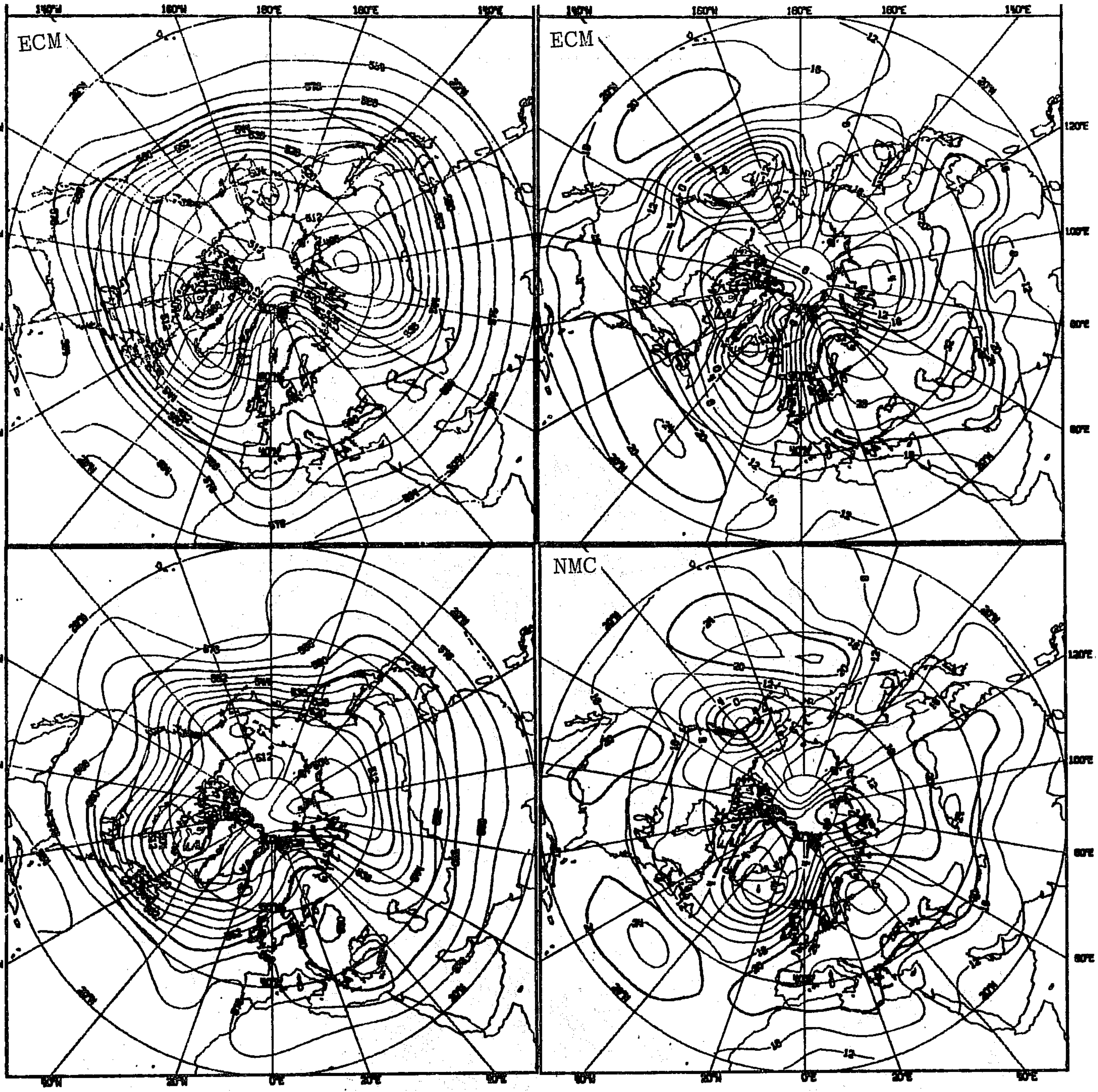


Fig.5.4.15 Time mean maps for days $6\frac{1}{2}$-$9\frac{1}{2}$ (21.2.76 1200 GMT-24.2.76 1200 GMT). ECM forecast from 00Z on 15.2.76 and corresponding NMC analysis. Left: 500 mb, contour interval: 8 dekametres. Right: 1000 mb, contour interval: 4 dekametres.

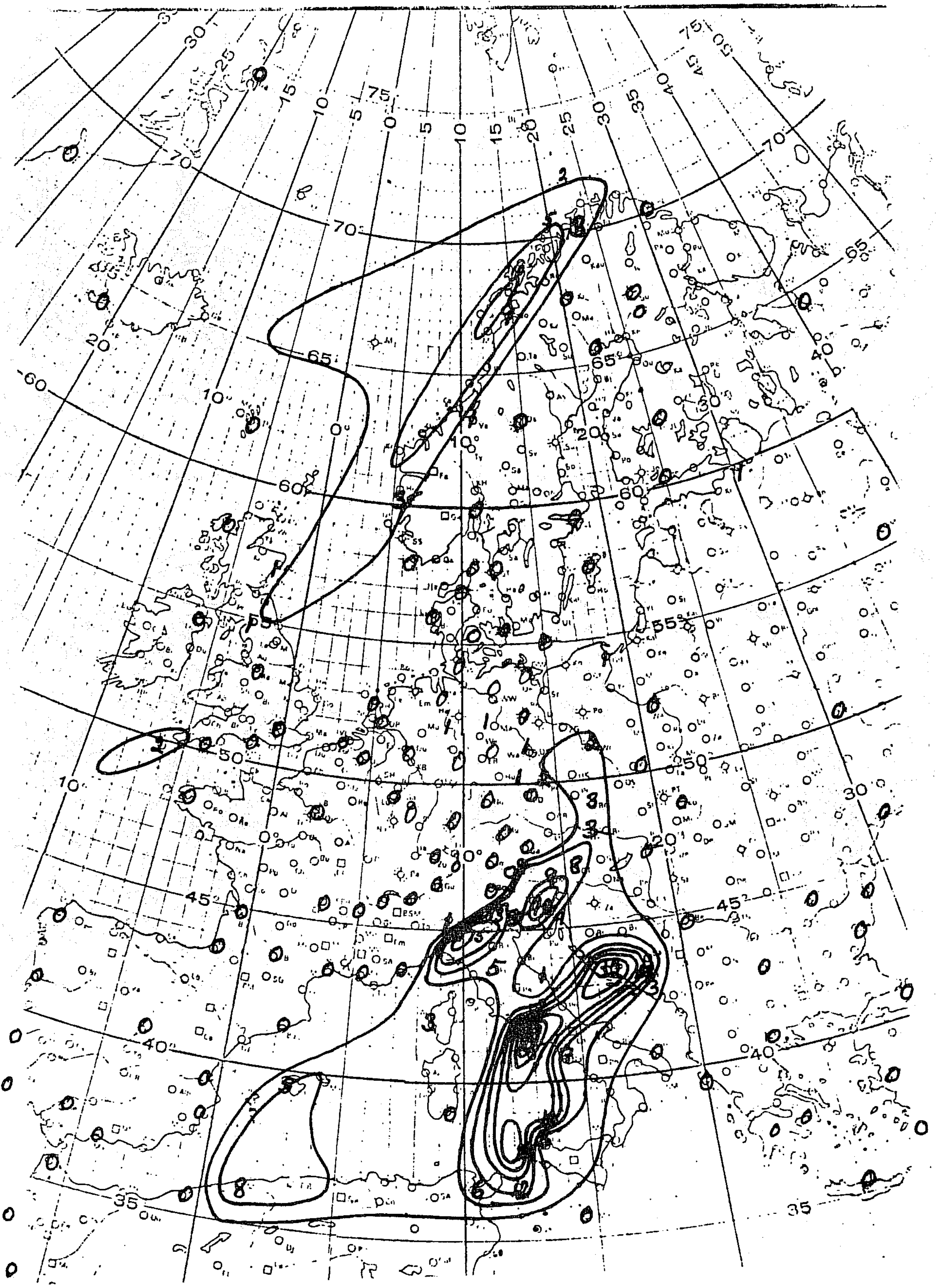


Fig.5.4.16 Observed 24 hour accumulated precipitation for period 15.2.76 0600 GMT-16.2.76 0600 GMT. Isohyet interval: 2,5,10,15,20,30,40,...etc.

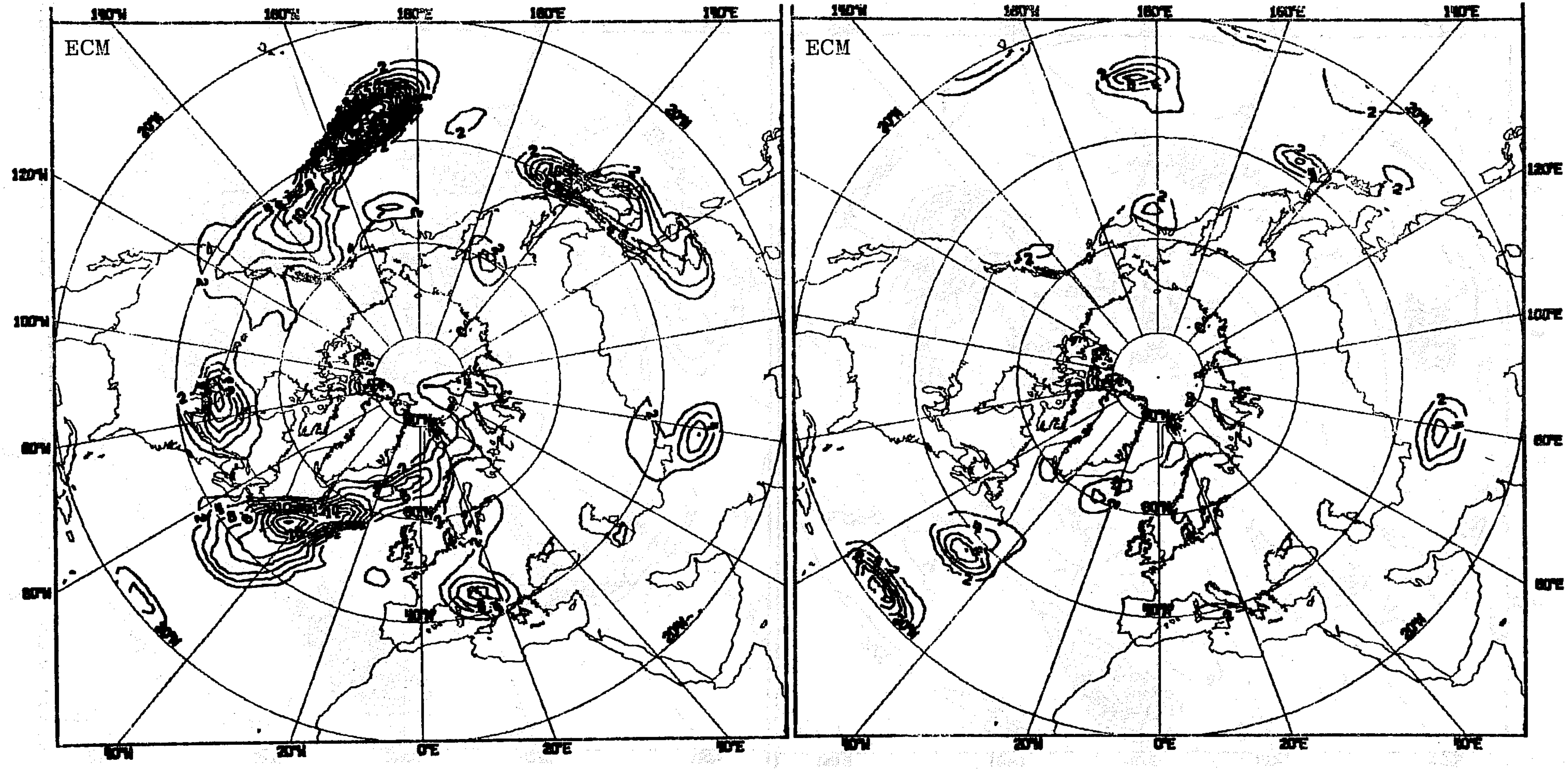


Fig.5.4.17 Left: ECM forecast accumulated large-scale precipitation for first 24 hours of forecast from 00Z on 15.2.76. Isohyet interval: 2 mm. Right: ECM forecast accumulated convective precipitation for first 24 hours of forecast from 00Z on 15.2.76. Isohyet interval: 2 mm.

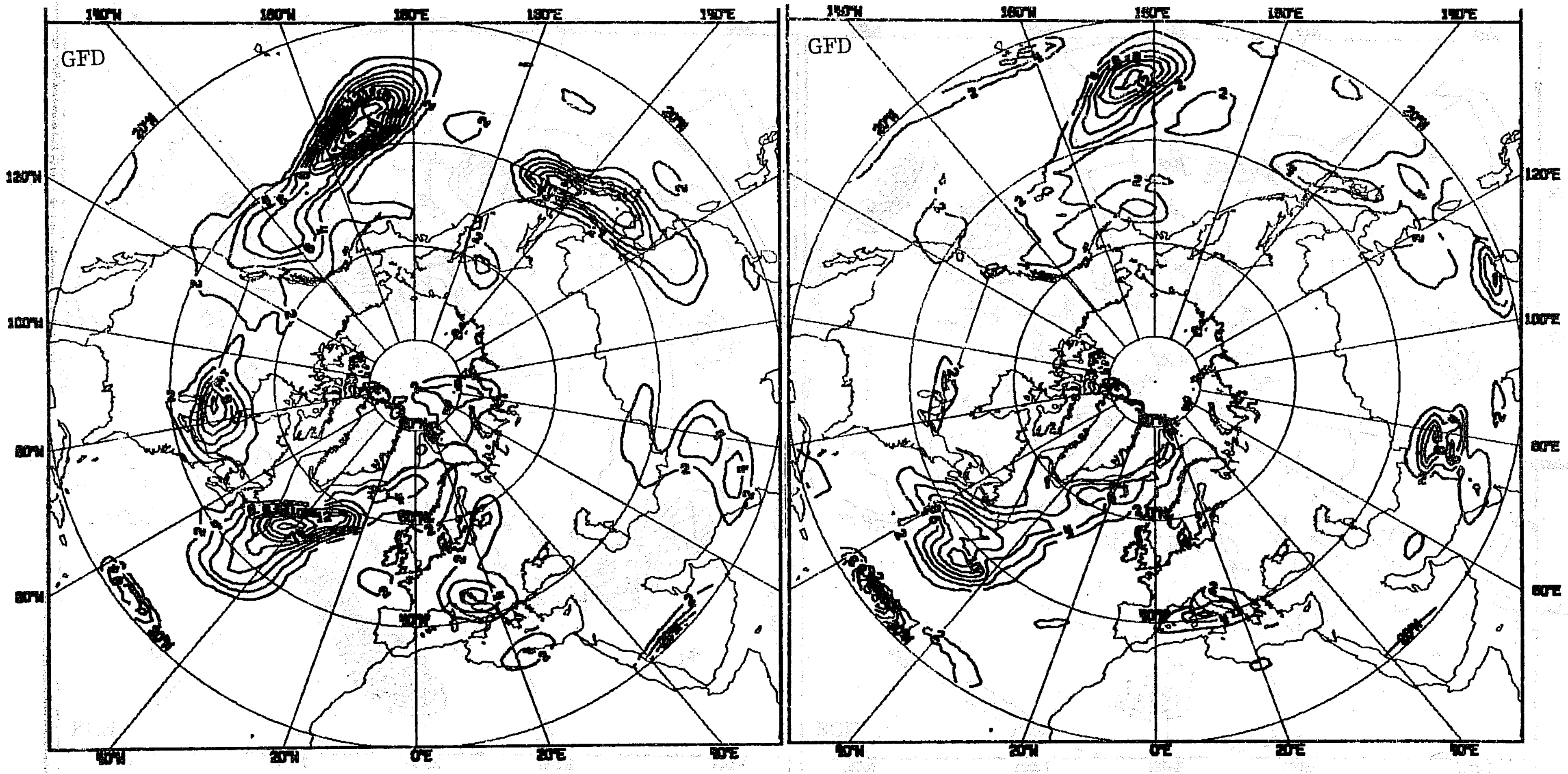


Fig.5.4.18 Left: GFD forecast accumulated large-scale precipitation for first 24 hours of forecast from 00Z on 15.2.76. Isohyet interval: 2 mm. Right: GFD forecast accumulated convective precipitation for first 24 hours of forecast from 00Z on 15.2.76. Isohyet interval: 2 mm.

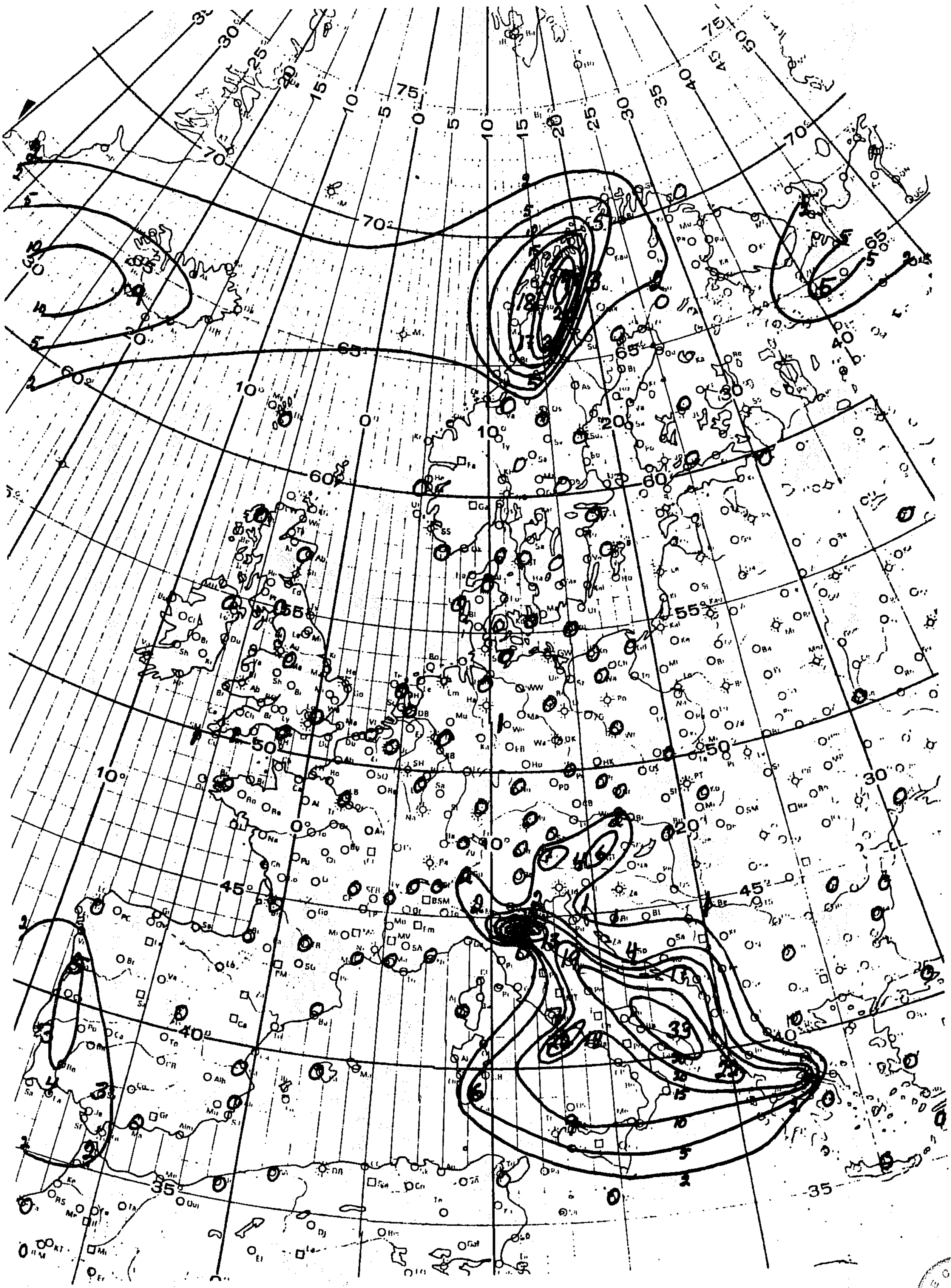


Fig.5.4.19 Observed 24 hour accumulated precipitation for period 16.2.76 0600 GMT-17.2.76 0600 GMT. Isohyet interval: 2,5,10,15,20,30,40,...etc.

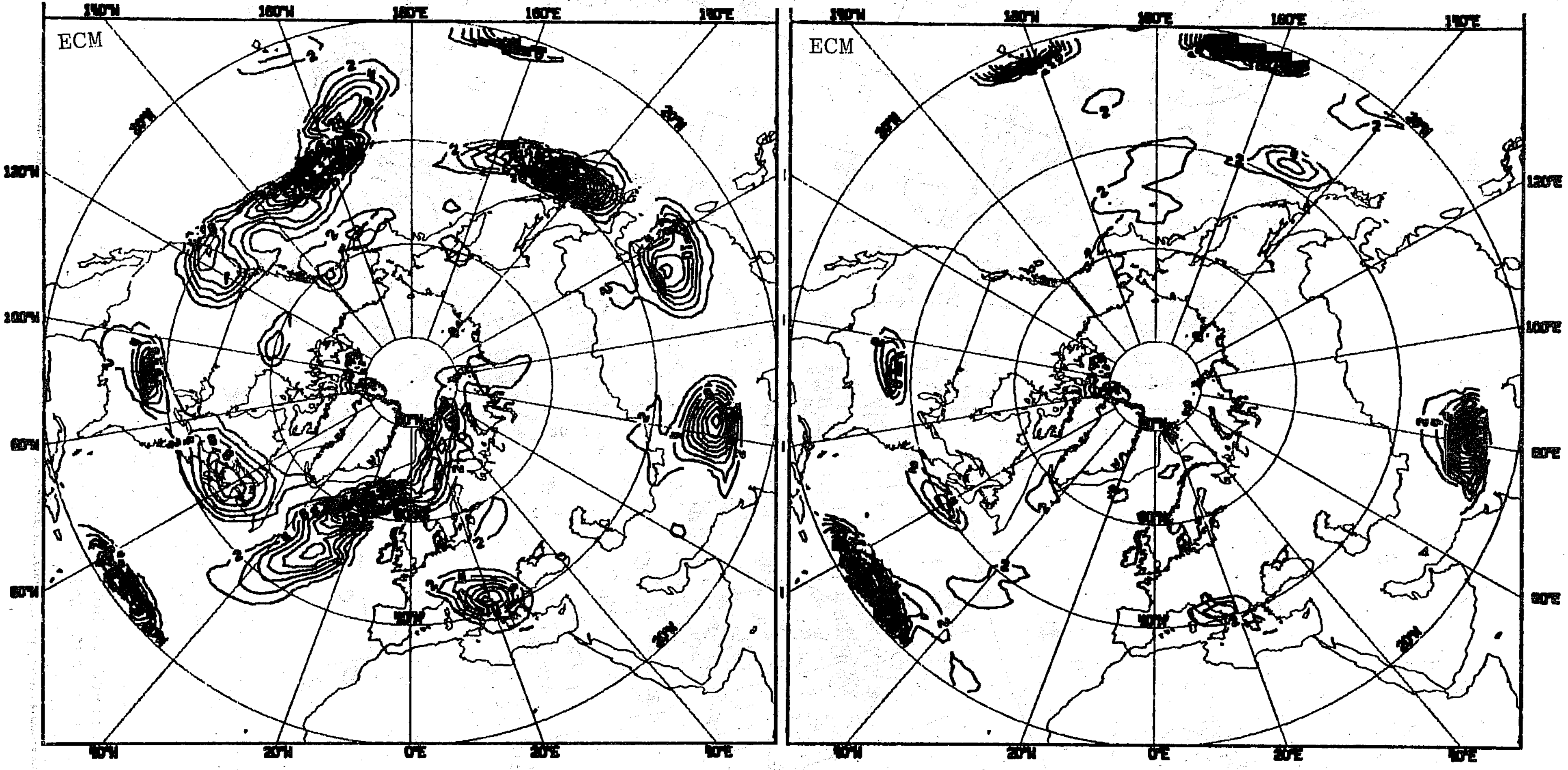


Fig.5.4.20 Left: ECM forecast 24 hour accumulated large-scale precipitation for day 1 of forecast from 00Z on 15.2.76. Isohyet interval: 2 mm. Right: ECM forecast 24 hour accumulated convective precipitation for day 1 of forecast from 00Z on 15.2.76. Isohyet interval: 2 mm.

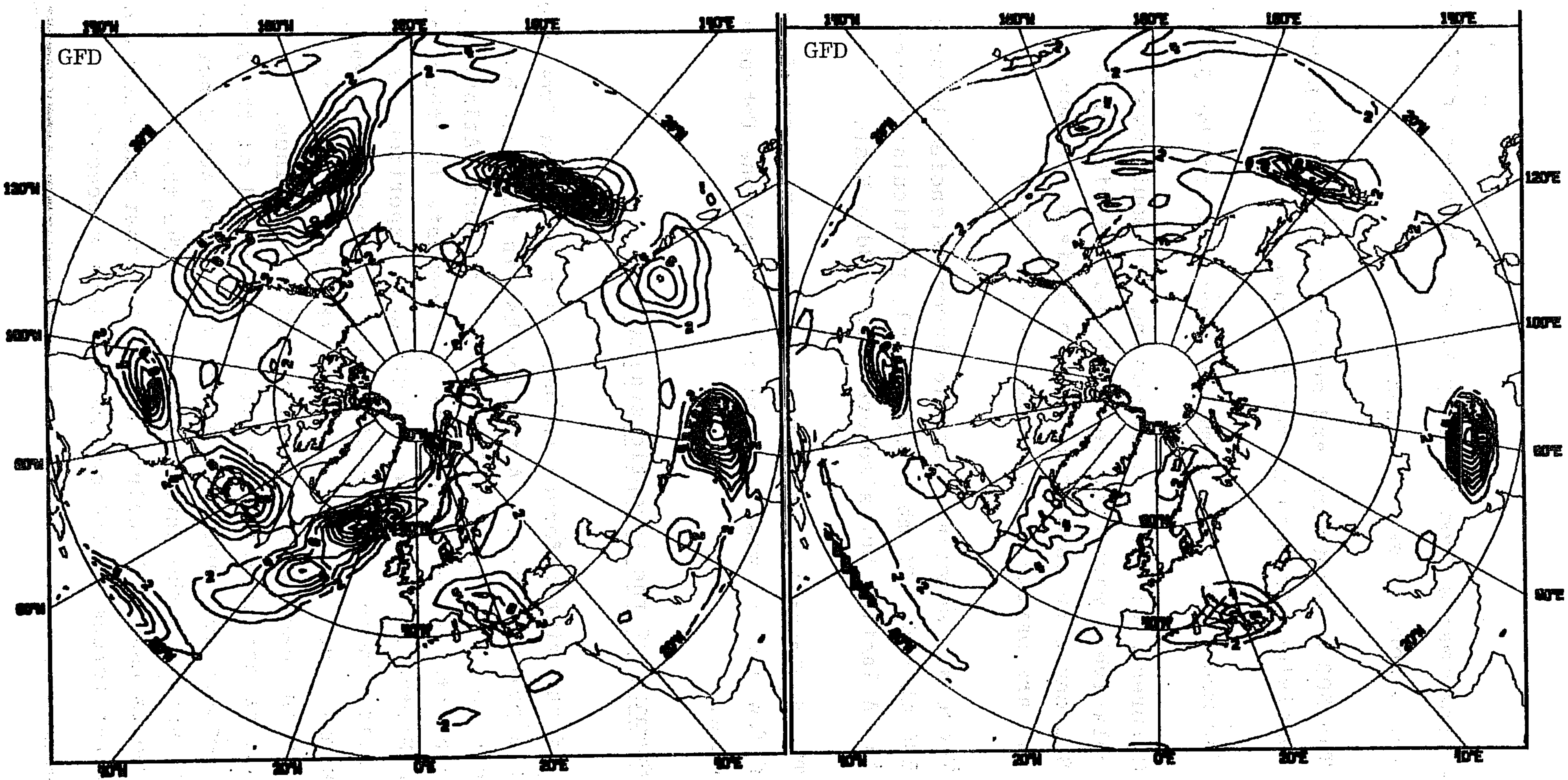


Fig.5.4.21 Left: GFD forecast 24 hour accumulated large-scale precipitation for day 1 of forecast from 00Z on 15.2.76. Isohyet interval: 2 mm. Right: GFD forecast 24 hour accumulated convective precipitation for day 1 of forecast from 00Z on 15.2.76. Isohyet interval: 2 mm.

These experiments from February 15 are, however, different from the others in several ways :

i. The 500 mb isotherms and geopotential remain generally consistent throughout the forecast.

ii. The models do not become too zonal around day 3-5.

iii. Outbreaks of cold air occurring far south over relatively warm open sea with deficiencies up to 8°C cause an incorrect rise in surface pressure and are partly responsible for iv. below.

iv. There are sometimes notable inconsistencies between 500 mb and 1000 mb geopotential, particularly over the Atlantic, but also over the Middle East.

v. The dominant high does not move away or decay after a few days as was observed in the previous February case.

5.5 Summary of the synoptic evaluation

Three of the winter forecasts have been assessed from the synoptic point of view. Attention has in the main been confined to the region of Europe and the North Atlantic, although note has been taken of developments in other parts of the Northern Hemisphere. Individual experiments show some variationsin skill in different geographical areas, for example over the Pacific where the level of skill was lower in the case of all three experiments. Thus, conclusions concerning the accuracy and skill of the forecasts reached for limited geographical areas would not in general apply to the entire hemisphere. The three experiments assessed include two of the most successful of the winter forecasts and one rather less successful.

Considering temperature and geopotential forecasts, on the average over the three ECM and three GFD forecasts assessed, northern Europe would have had useful synoptic guidance from the forecasts for around 6 or 7 days. The worst of the experiments for Europe are ECM and GFD from 3.2.1976, both useful for 4 days. The best were the experiments from 15.12.1978, both of which could be useful for Europe (with some reservations) for 8-9 days. Both 500 mb forecasts from the 6th February became too zonal after a few days, but improved again in the later stages of the integrations, becoming more meridional. The phase speeds of the forecasts varied somewhat, being in some cases too slow, in others too fast, and also changing during the period of the forecasts. However, the phase speeds are generally forecast quite well. The 500 mb solutions were rather good from the 15 February, the 1000 mb geopotential forecasts were not so successful. This was due to an incorrect temperature forecast particularly at the 850 mb level with excessive outbreaks of cold air from the north but with correct warm air advection from the south in both forecasts. This error was found predominantly over the Atlantic and Pacific. It favoured the amplification of the 500 mb troughs in the forecasts, at the expense of the surface forecast. The predicted excess of cold air at 850 mb did not occur to the same extent in the other experiments. In many cases the 850 mb temperature gradients in the baroclinic zones were predicted to be sharper than were observed, with some excess of cold air north of the fronts, but also with warmer air south of the fronts, but also consistent with the often much stronger surface developments observed in the forecasts compared to the analysis.

Occasionally, also, there is a lack of consistency between the 500 mb isotherms and geopotential, such that the cold air is sometimes found to the east of the countour trough, a situation favourable for filling the trough.

This happened particularly in both forecasts from the 6th February. Thus it seems likely that the excessive surface developments often seen, especially in the GFD forecasts, are associated with an excess of warm air in the lower troposphere close to the centre of the lows, and sharper horizontal temperature gradients. In two of the experiments the blocking high was rather persistent but was incorrectly divided by a baroclinic current, one part being pushed too far north of Scandinavia, the other too far south towards the Black Sea. In all experiments the jets attacking the high were forecast too far south and the redevelopment of the high over the eastern Atlantic was not correctly forecast. The systems which were incorrect from around day 5 over the Pacific were affecting the Atlantic areas and Europe during the later stages and this contributed to the decay of forecast skill in the last five days over Europe.

An attempt has been made to verify precipitation distributions and amounts for day 1 and 3 of the experiments from 3 February (Section 5.1c) and for days 1 and 2 of the experiments from 15th February (Section 5.3d). The month of February 1976 was a rather dry month for most of Europe due to the predominance of high pressure, making verification difficult. However, both experiments starting from 3 February erroneously developed precipitation over western Europe, in association with the deterioration of these forecasts.

The experiments from the 15th February appear to have overpredicted the quantities of precipitation over the north east Atlantic, in accordance with the prediction of developments stronger than observed. In the Mediterranean region, amounts of precipitation of convective origin were underestimated by the ECM model, and in general the ECM model produced less convective precipitation than did the GFD model. Otherwise, the areal distribution of large-scale precipitation appears to be overestimated, but the predicted amounts agreed reasonably well with the available verifying data over Europe.

6. Downstream energy propagation

The synoptic discussion of the previous section took the forecasters' view of the experiments. In this and later sections we fix attention on other aspects of the forecasts which merit further discussion. The discussion here will centre on the phenomenon of downstream energy propagation. This phenomena seems to be predictable for surprisingly long periods.

6.1 Downstream energy propagation

The phenomenon of downstream energy propagation has been discussed by several authors, e.g. Rossby (1945), Hovmöller (1949), Yeh (1949), van Loon (1962) and Simmons and Hoskins (1978). The latter paper contains a good bibliography. Baumhefner and Downey (1978) discuss forecasts of this phenomenon by some research and operational forecast models. Simmons and Hoskins consider an isolated low in an otherwise undisturbed unstable zonal flow. They show that a succession of lows and highs forms downstream. The rate at which new members of the family of highs and lows appear implies a propagation velocity intermediate between the 500 mb and 300 mb mean zonal wind. Fig. 6.1.1 shows a Hovmöller diagram for the 500 mb geopotential, at latitude 45 N, for the February 15 - March 4 1976. Troughs and ridges are marked by solid and dashed lines. Two, and perhaps three, episodes of the phenomenon of downstream propagation of energy may be distinguished. They have been indicated by the labels I, II, III on the heavy lines. The applicability of the description to episode III may be questionable.

The paper by Simmons and Hoskins did not discuss effects which might give a preferential location for the occurrence of the phenomenon.

However, in Fig. 6.1.1 the phenomenon may be interpreted, to some extent, as pulsations of quasi-stationary features.

We have four pairs of ten-day forecasts covering the period, viz. forecasts from 15, 18, 22, 25 February. Fig. 6.1.2 shows the Hovmöller diagrams for the observations and the two forecasts beginning on 15 February. For convenience the observed trough and ridge lines have been copied onto the forecast Hovmöller diagrams; we have also copied the heavy line I (ray I) onto the forecast diagrams.

A number of features, common to both forecasts, are evident. Both models catch the inception of the phenomenon quite well at 60 E at day 1½. However, it is clear that both models produce the phenomenon with an east west wavelength which is too long (by 15^{o}) and a propagation speed which is too large (50^{o}/day as against 40^{o}/day). By day 6 the third low of both forecast families occurs just where the third ridge of the observed family is located. After this point the phenomenon seems to dissipate in the models but the observations show it continuing to form a large low at 300 E at day 9.

Fig. 6.1.3 shows the Hovmöller diagrams for the observations and forecasts beginning on 18 February. There are now two episodes (I and II) of the phenomenon to be found in the ten day period. Both models show a serious forecast failure with episode I over North America. At day 3 in the observations there are two troughs close together at 260 E and 300 E. Between day three and day six the eastern trough fills and the western trough deepens and moves eastward under the influence of the energy propagation from upstream. Neither model forecasts that sequence of events correctly. They both deepen the easternmost of the original pair of troughs and by day 7 both models show a trough at 325 E where a ridge is observed. This development of the wrong trough also occurred in the forecast from the 15th. However, both forecasts for episode II, which begins on day 7 at 135 E are quite remarkable in the success of the forecast phases and amplitudes between day 7 and day 9. Fig. 6.1.4 shows the 500 mb height field at day 9. The success of the forecast for the western hemisphere is evident.

Fig. 6.1.5 shows the Hovmöller diagrams for forecasts and observations for the ten-day period beginning on 22 February. Episode I is successfully forecast in the first three days. The phases of episode II are also very successfully forecast by both models from day 2½ to day 6½. However the amplitudes of the episode are rather over-estimated by both models.

Fig. 6.1.6 shows the Hovmöller diagrams for the forecast from 25 February. Neither forecast shows the development of the trough at 280 E at day 2½. Both models do show the development of a trough in this location one day later.

The mutual interconnection of events along the line III in Fig. 6.1.1 is less clear cut than the earlier episodes. The forecasts from February 25 have not been very successful in predicting it. The forecast failure starts between days 4 and 6 with a general and erroneous eastward movement of the quasi stationary troughs and ridges between OE and 180 E. Again the models are very similar in this regard.

To summarize this discussion we make the following points.

The similarity of the two models in their treatment, both successful and unsuccessful, of these phenomena is remarkable and suggests that the phenomena are mainly of dynamical origin. Secondly, there is evidence from the forecasts from 18 February that this particular phenomenon can be predictable nine days ahead. Thirdly, forecast success seems to depend strongly on the initial data. If we consider the longitude band 270-300 E we had a good three-day forecast from the 22 February and a poor one from the 25 February. In the same longitude band we had poor six-day forecast from the 18th and a good six-day forecast from 22 February. Variability in the quality of the initial data must contribute to this variation in forecast quality.

Finally we have seen in the forecasts, particularly in that from 25 February, the necessity to maintain correctly the location of the quasi stationary troughs for a successful forecast of the phenomenon of energy propagation. This aspect will be reviewed from a rather different point of view in the discussion of the systematic errors in Section 9.

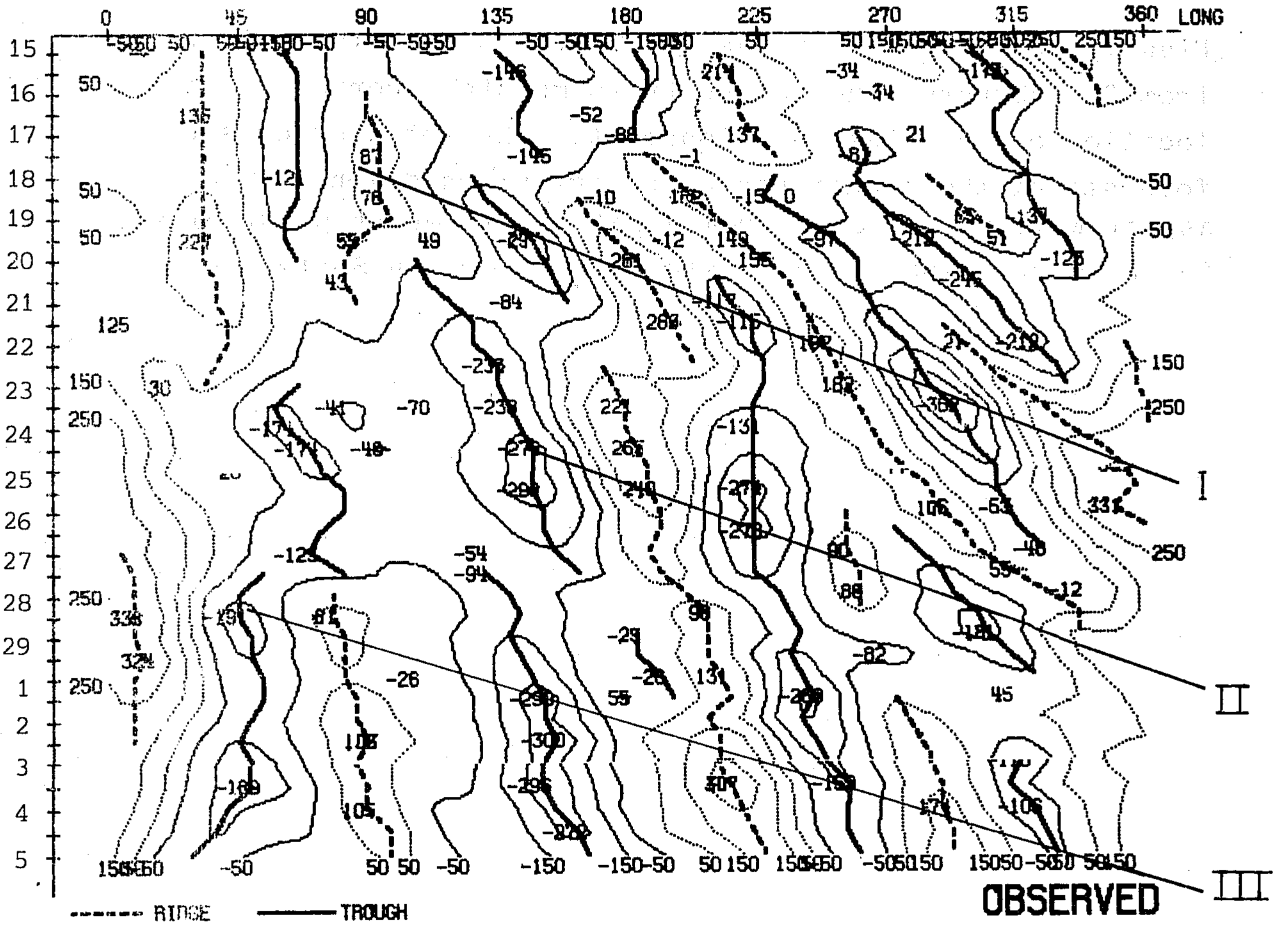


Fig.6.1.1 Trough-ridge diagram for wavenumbers 1-9, of observed 500 mb height at 45 N for the period 15.2.76 to 5.3.76. The contour interval is 100 m with contours at ± 50 m. Heavy solid lines denote trough lines, heavy dashed lined show ridge lines. The straight lines marked I, II, III show episodes of downstream intensification.

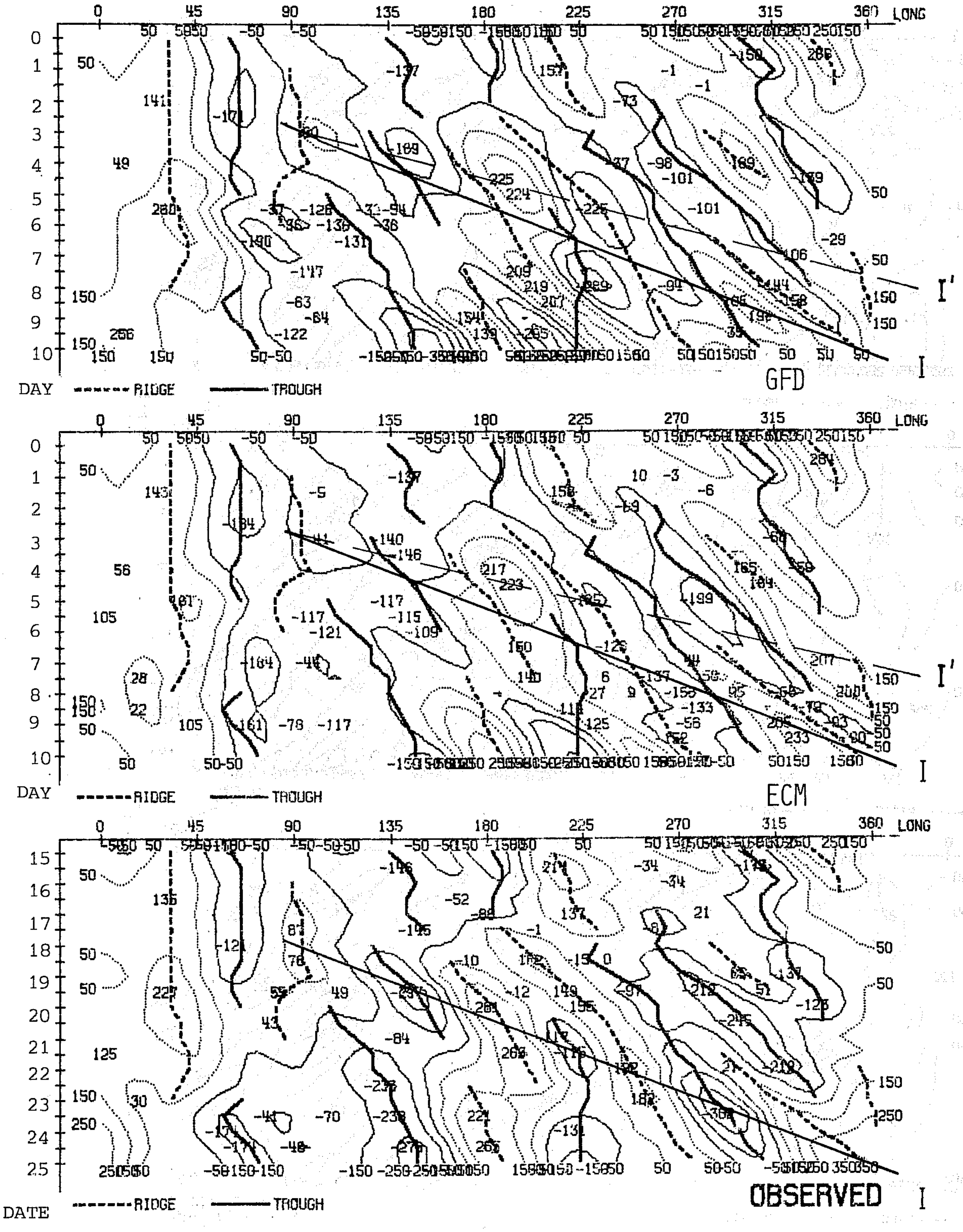


Fig.6.1.2 Trough-ridge diagrams for forecast and observations of wavenumbers 1-9 of 500 mb height at 45 N for the 10 day period from 15.2.76. The contour interval is 100 m. The trough (solid heavy lines) and ridge lines (dashed heavy lines) of the observed field have been copied onto the forecast fields. The solid line I from Fig.6.1.1 has been copied onto all three panels. The dashed lines I' indicate the phenomenon as it occurs in the models.

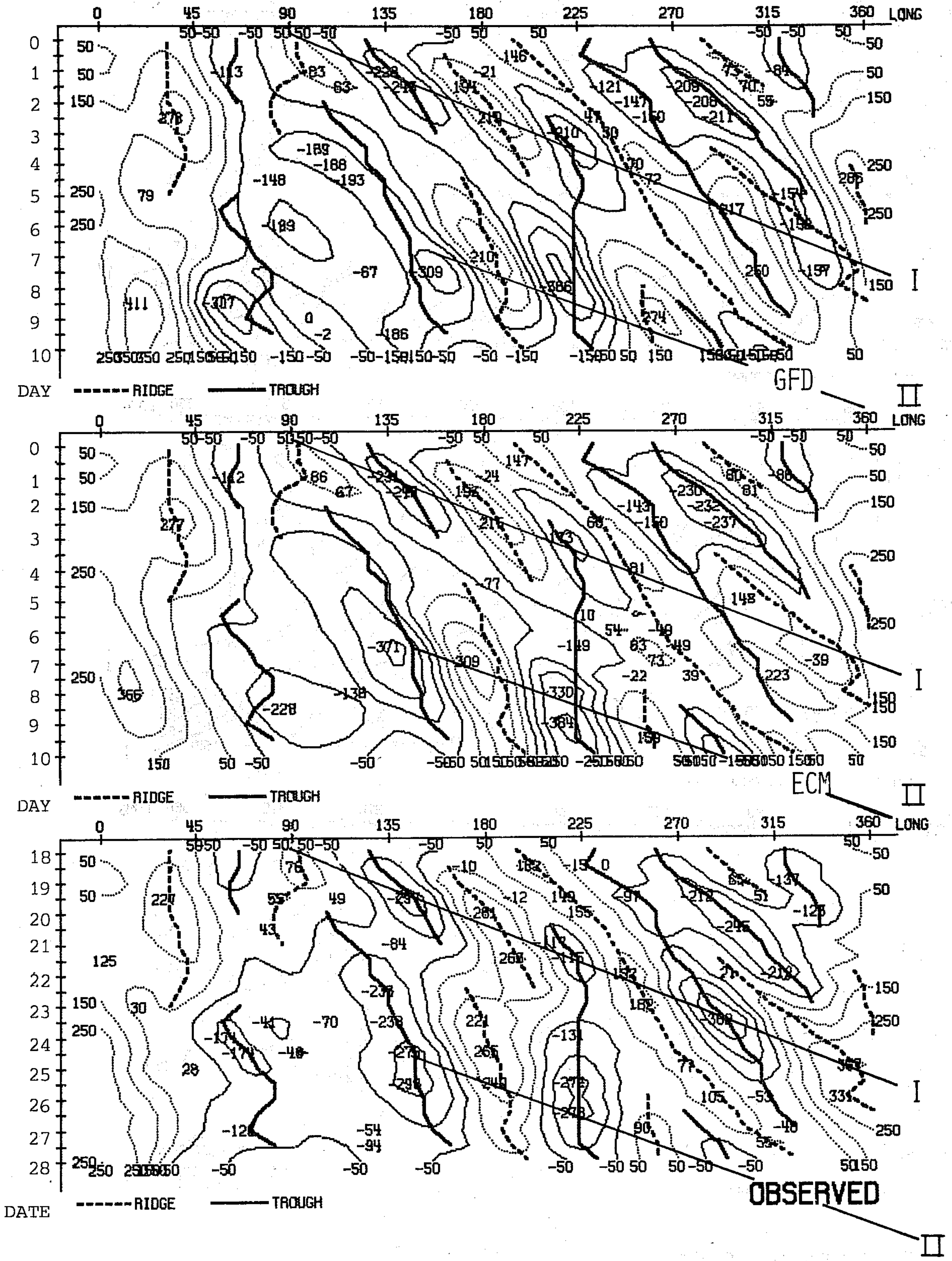


Fig.6.1.3 As Fig.6.1.2 for the forecasts and observations from 18.2.76.

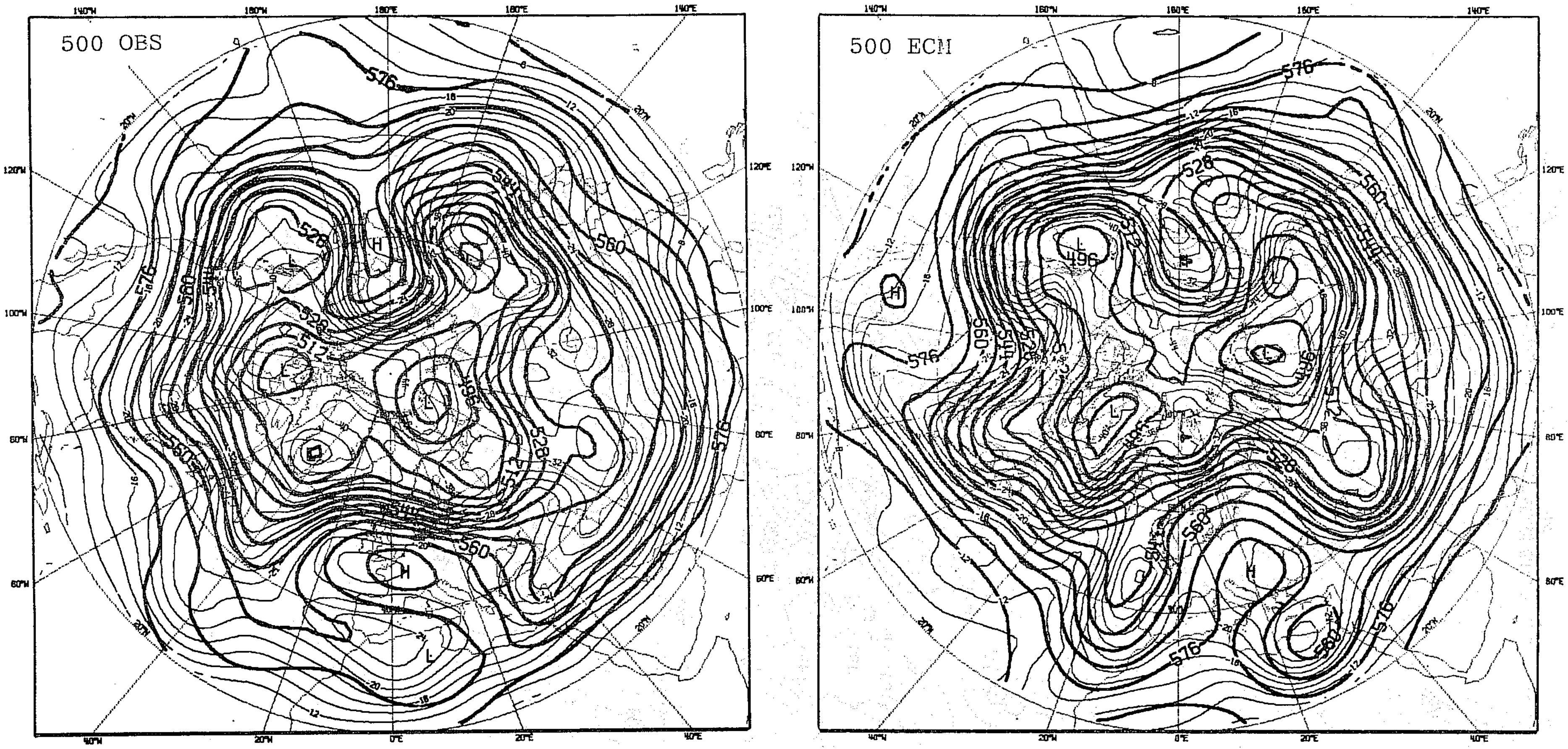


Fig.6.1.4 Nine day 500 mb forecast (valid 0Z, 27.2.76) from 18.2.76 (right) and verifying analysis. Heavy lines are geopotential in intervals of 8 dam and thin lines are temperature in intervals of 2K.

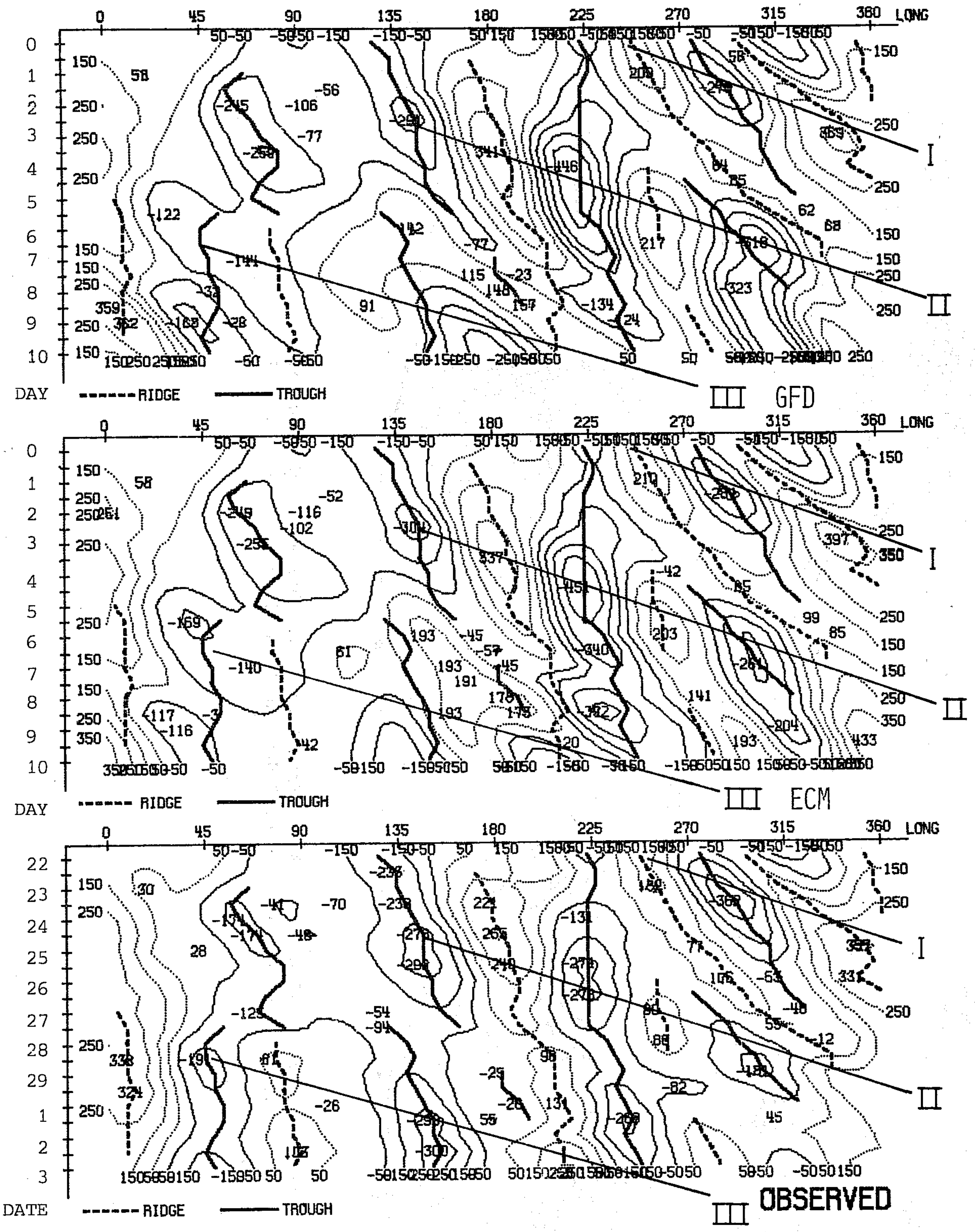


Fig.6.1.5 As Fig. 6.1.2 for forecasts from 22.2.76.

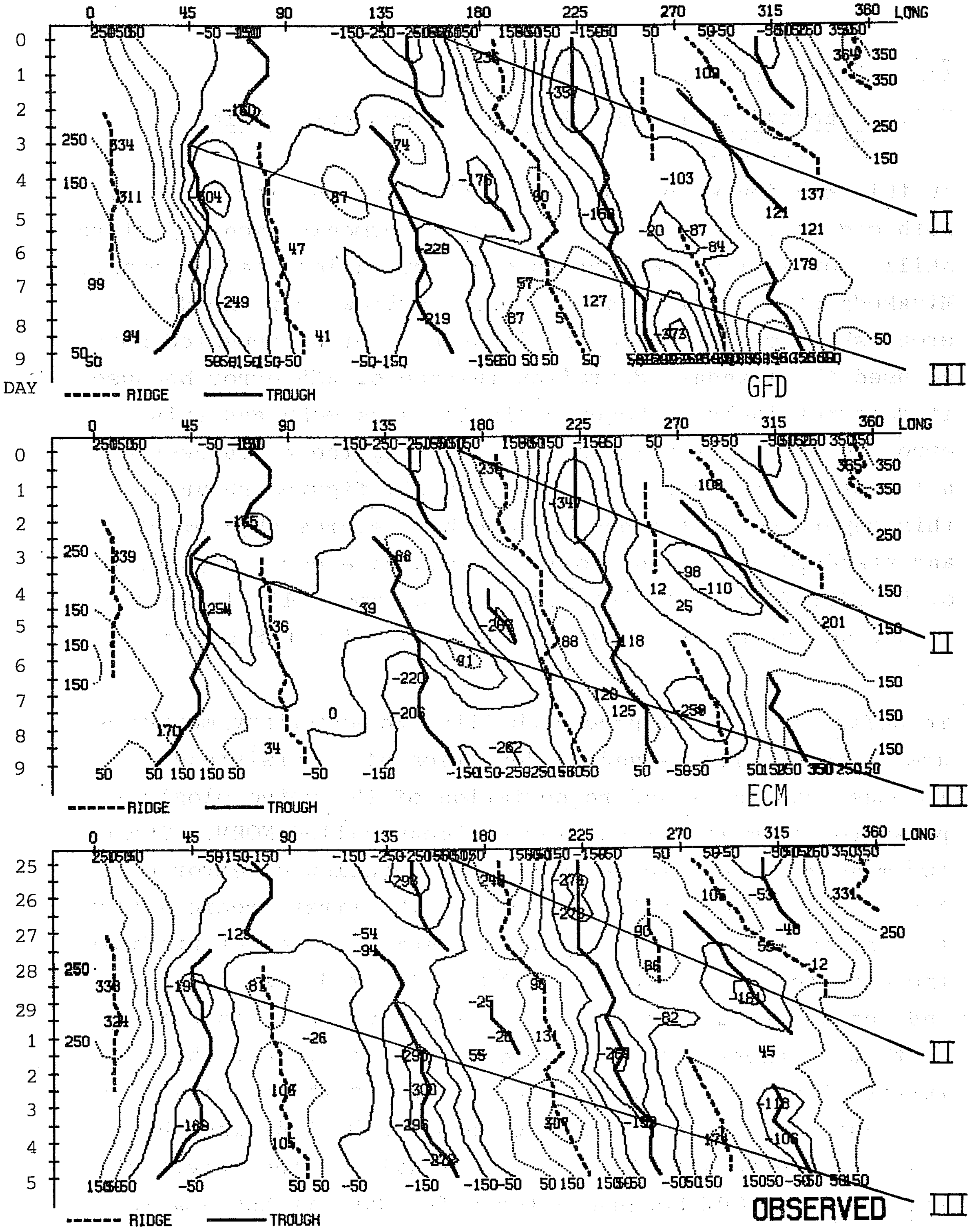


Fig. 6.1.6 As Fig. 6.1.2 for forecasts from 25.2.76

7. Large scale verification scores

7.1. Standard deviation (RMS error) and correlation coefficients

In this section we discuss the quality of the forecasts of both models as reflected by the most commonly used objective skill scores, i.e. RMS errors and correlation coefficients. Miyakoda et al (1972) gave a good overview over the different skill scores therefore it should not be repeated here. We used the standard deviation instead of RMS error because it does not include the error in the area mean and this error is of no interest when considering the usefulness of a forecast. It turned out that for all figures shown in this report no difference between both scores (RMS error and standard deviation) could be spotted except in panel a of Fig.7.1.7 where the GFD forecast became a little worse to the end of the forecast period when using RMS error.

To define a range of predictability two auxiliary measures are included, i.e. standard deviation of a persistence forecast and the standard deviation of the meteorological parameter from its climatological mean called NORM. These two measures indicate the standard deviation of error of two types of "no-skill forecasts"; the first predicts that the initial values will not change while the other predicts the climatological (normal) value. The model's forecast ends being useful when it gets a score which is inferior to that of those "no skill forecasts." Therefore the intersection of the standard deviation of the model's forecast with one of those two measures will define the range of predictability; mostly it will be the intersection with the NORM because after a few days it has smaller values than the standard deviation of the persistence forecast.

Another widely used skill score is the correlation coefficient. We used here the anomaly correlation coefficient, i.e. the deviation from a climatological mean, because we consider it more appropriate for medium range weather forecast than the correlation coefficient of change from the initial value, which is mostly used for short range weather forecast. To define a range of predictability several methods are in use, i.e. the intersection of the anomaly correlation coefficients with the 60%, 50%, or 0% level or with the anomaly correlation of the persistence forecast. Again we will use the most restrictive definition for predictability. We did this because we found that they gave best agreement with the synoptic evaluation in each individual case.

As these measures for forecast skill are not yet uniform in the literature one has to be careful when comparing the results of our experiments with those of other authors, who may use somewhat less stringent criteria.

We used the following formulation for the calculations of standard deviations and anomaly correlation coefficients:

The error of a quantity X is given by

$$\delta X(t) = X_p(t) - X_o(t)$$

where $X_p(t)$ and $X_o(t)$ are the predicted and observed (analyzed) values at the predicted time, t, respectively. The anomaly is defined by

$$\Delta X(t) = X(t) - X_n$$

where X_n is the climatological (normal) value for February. The climatological values were taken from Crutcher and Jenne (1970). We used two averaging operators, one for an area mean or vertical average:

$$\overline{X} = \frac{\sum_i g_i X_i}{\sum_i g_i}$$

where X_i is an arbitrary variable at the grid point i and

g_i is a weighting factor. As we had a regular latitude longitude grid we used the cosine of latitude as weighting factors. Vertical averages were, of course, weighted by pressure .

The second type of averaging is an ensemble mean for a sample of 7 February cases:

$$E(X) = \frac{1}{7} \sum_{n=1}^{7} X_n$$

where X_n is the value of skill score for a particular case.

$$\text{standard deviation} = E\left(\sqrt{\overline{(\delta X-\overline{\delta X})^2}}\right)$$

$$\text{persistence standard deviation} = E\left(\sqrt{\overline{\left\{X_o(t)-X_o(t_o)-[\overline{X_o(t)-X_o(t_o)}]\right\}^2}}\right)$$

$$\text{NORM} = \frac{1}{T} \sum_{t=1}^{T} \sqrt{\overline{\left\{\Delta X_o(t)-\overline{\Delta X_o(t)}\right\}^2}}$$

T means a period of 6 February months which were taken from the NMC operational analyses from 1965 to 1970.

$$\text{anomaly correlation} = E\left(\frac{\overline{\left(\Delta X_p(t)-\overline{\Delta X_p(t)}\right)\left(\Delta X_o(t)-\overline{\Delta X_o(t)}\right)}}{\sqrt{\overline{\left(\Delta X_p(t)-\overline{\Delta X_p(t)}\right)^2}}\cdot\sqrt{\overline{\left(\Delta X_o(t)-\overline{\Delta X_o(t)}\right)^2}}}\right)$$

$$\text{anomaly correlation of persistence} = E\left(\frac{\overline{\left(\Delta X_o(t_o)-\overline{\Delta X_o(t_o)}\right)\left(\Delta X_o(t)-\overline{\Delta X_o(t)}\right)}}{\sqrt{\overline{\left(\Delta X_o(t_o)-\Delta X_o(t_o)\right)^2}}\cdot\sqrt{\overline{\left(\Delta X_o(t)-\Delta X_o(t)\right)^2}}}\right)$$

To enable a separation into contributions by disturbances of different scales the calculations are based on Fourier-series. The transformation of the formulae given above into wavenumber domain is straightforward but not given here because they look quite complicated and would disguise the main points.

In Fig. 7.1.1 we see the time evolution of the vertically (1000 mb - 200 mb) and horizontally (20 ^{O}N - 82.5 ^{O}N) averaged skill scores of the height fields. For each day of forecast the range of variability is also indicated by vertical lines. The values obtained from the ECM model are given by solid lines and those from the GFD model by dashed lines. In addition thin lines in the panels for standard deviation (a-c) indicate the mean climatological standard deviation for February and in the panels for correlation coefficients (d-f) the 60% level. Heavy dots indicate the skill of a persistence forecast.

Up to day 7 the differences between the scores for the total field of both models (panel a and d) are marginally in favour of the ECM model. This is mainly due to the contribution by (the mainly baroclinic) waves with zonal wavenumbers 4 to 9 (panels b and e).

After day 7 the GFD model shows better scores, but this is normally a period when the forecasts have little valuable information left.

The range of variability (vertical lines) shows some difference between both forecasts at the long waves, especially in the correlation coefficients. The ECM model has a somewhat higher variability and is, therefore, somewhat less reliable regardless of its mean value, unless we can spot the unreliable cases from the initial synoptic situation. The variability of the scores is small at first but it increases considerably at about the 5th day of the forecast and this may already indicate a limit in reliable predictability.

From the standard deviation of the total fields (panel a) we get a 6 day period of predictability for both models. This period is not only valid for the vertical mean, which is shown here, but also for the 500 mb level while we have a shorter period of 5 days for the 1000 mb level.

Within the 7 cases we have quite a large variability. The lowest values (5 or 4½ days for the GFD or ECM model respectively) were for the case based on the data of the 22 February 1976 and the best values (9 or 8½ days for the GFD or ECM model respectively) were for the case based on the data of the 15 February 1976. The other cases were quite near the average.

If we consider the behaviour of the standard deviation in the long (1-3) and medium (4-9) zonal wavenumber groups, we get a slightly higher period of 6½ days for the long waves (panel c) and a slightly lower period of 5½ days for the medium waves (panel b).

It might be worth mentioning that our best case overall (15.2.76) had at 1000 mb the worst range of predictability (only 3½ days) for the medium waves (wavenumbers 4-9) for the height field in both forecasts. This could be seen in the standard deviation as well as in the correlation coefficients.

We now consider the anomaly correlation scores. Using the intersection of the correlation coefficient with the 50% or 60% level leads to a period of 6¼ or 5½ days respectively for the total fields (panel d) for both models and agrees quite well with the 6 day period we got from the standard deviation. The variability within the 7 cases is also quite similar to that found for the standard deviation. The 60% intersections are, on the whole, nearer to the values from the standard deviation-NORM intersection. This agreement is mainly due to the contribution by the long waves; the medium waves take only 4¼ or 4¾ days to reach the 60% or 50% level respectively for both models, which is about 1 day shorter than the time to reach the standard deviation-NORM intersection.

The difference in the skill scores between both forecasts are fairly small and one would like to know if they are below the noise level. A partial answer is provided by a comparison of the forecasts with a different set of analyses. We had available the operational analyses made by NMC and by Deutscher Wetterdienst. As the differences between both analyses should indicate the uncertainty of our knowledge of the true state of the atmosphere, a comparison of the forecast result with both analyses should indicate the uncertainty of the skill scores. Although the differences between both sets of analyses are considerable (20-30 m standard deviation) the values given in Fig. 7.1.1 are almost identical for both sets of analyses and, therefore, we may have some confidence in these results.

In Fig. 7.1.2 the vertical distribution of the correlation coefficients of the total height fields is displayed. Two relative maxima of error are indicated: one at the 300 mb level and the other one at the surface. Obviously the 500 mb level is the most reliable one with half a day longer predictability than at the surface. The error growth is strongest in the boundary layer.

Up to now we have only discussed means of the skill scores between latitudes 20 $^{\circ}$N and 82 $^{\circ}$N and it would be interesting to know if the results would be different by chosing other zonal belts. To investigate the question three belts were chosen: 32.5 $^{\circ}$N - 42.5 $^{\circ}$N, 45 $^{\circ}$N - 60°N, and 65 $^{\circ}$N - 75 $^{\circ}$N. The absolute values of standard deviation are quite different within these three belts but so are the climatological standard deviations (NORM). The intersections between the standard deviation and the NORM gave almost the same period of predictability with a slight advantage of half a day for the ECM-model at the belt 65 ^{o} - 75 ^{o}N and for the GFD-model at the belt 45^{o} - 65 ^{o}N compared to the 6-day period we got from Fig. 7.1.1.

The latitudinal variability of the anomaly correlation coefficients is shown in Fig. 7.1.3. They are vertically averaged between 1000 and 200 mb. There is a pronounced meridional structure after day 5 with relatively good skill at about 35 ^{o}N and 72.5 ^{o}N in the ECM-forecasts and worst skill at about 55 ^{o}N. The GFD-forecasts show a similar structure. If we take the 50% level for the definition of predictability we get a range between 5½ days to almost 7 days compared to 6¼ days for the average for the whole belt between 20^{o} and 82.5 ^{o}N.

This structure can be found in most cases under consideration but for each case the latitude of minimum correlation was different, as was the day when the structure became obvious. Fig. 7.1.4 shows one of the most clear cut examples based on the forecast from 15 February 1976. After 3 days there is a sharp drop of correlation to -13% concentrated in a narrow latitudinal band. This low correlation was not associated with low amplitudes, because the RMS error corroborates the feature. To get an insight into this problem we examined the Hovmöller trough-ridge diagrams for all cases at the latitudes with low correlation and at the next latitudes with relative maxima in the correlation to the north and south of it. The main finding was that the trough or ridge which causes the low correlation is usually almost stationary south of this latitude and progressive north of it. Both the stationarity and the progression are forecast correctly and it is at the transition latitudes that the forecasts are wrong although this failure may be for a short period only.

The transition latitude tends to occur further south in the forecasts than in the observations.

To explain this finding we considered aspects of atmospheric behaviour or forcing which have a similar meridional variation. One such aspect is the fact that south of 40 °N or north of 70 °N there are hardly any cyclones while 55 °N is near the latitude of strongest cyclone activity. If this fact is the reason for the meridional variation in the correlation coefficients we would expect to see it most clearly in the scores at the lowest levels and in the contribution by the transient waves. In fact we find that this meridional variation of the anomaly correlation coefficient is strongest at the 200 mb level and in the quasi standing waves (see Fig. 7.1.5). Therefore it is unlikely that the mean position of cyclone tracks is the cause of this meridional variation of correlation coefficient.

On the other hand the topography has a meridional variation whose structure is similar to that of the correlation coefficients. Fig. 7.1.6 shows the standard deviation of the topography along longitude lines for all latitudes for the northern hemisphere. At those latitudes where we found better skill in the forecast (i.e. 35 °N and 72.5 °N) a much higher forcing of the atmospheric motion by the mountains can be observed. The centres of the Himalayas, the Rockies and of the Greenland massif are situated at these latitudes, while at 55 °N the direct forcing by mountains is smallest.

This coincidence is not conclusive proof, but it suggests that, in the models, the mountains provide a forcing to the model's atmosphere which is nearer the truth than other large-scale forcings. Fig. 7.1.6 shows not only the variance of the topography which was used in the model (solid lines) but also the values extracted from a more recent tabulation by RAND (dashed lines prepared by Gates and Nelson (1975)). Obviously they are not the same especially north of 60 °N but the structure is similar, which is the main argument here.

The suggestion that this meridional variation of skill is connected with the mountain forcing is also supported by the fact that the variation is best seen in the contribution by the quasi stationary flow of the 200 mb level (see Fig. 7.1.5).

The fact that this feature can be seen better in the correlation coefficients than in the RMS errors indicates that the forcing by mountains gets the phases of the troughs and ridges correctly.

To get an insight into the amplitudes of the waves, we need other tools. Hovmöller's trough-ridge diagrams of the standing waves at 500 mb show that the amplitudes for the model's results at 35 ^{O}N are slightly higher, but no clear answer for the other latitudes under consideration can be found.

Figs. 7.1.4 and 7.1.5 show only the correlations from the ECM-model because both models had the same behaviour with only slight differences in the positions of maxima and minima of skill.

Fig. 7.1.7 shows the same skill scores as Fig. 7.1.1 but now for the 850 mb temperature fields. We know already from Fig. 7.1.2 that the errors grow most rapidly near the lower boundary and it may be expected that the scores for low level quantities would indicate shorter periods of predictability than those given in Fig. 7.1.1. The standard deviation intersection with the NORM of the total (850 mb temperature) field (panel a) gives a period of 5¼ days of predictability with a slight advantage for the ECM model. The corresponding correlation coefficient intersection with the 50% or 60% level (panel d) gives periods of 5½ or 4½ days respectively.

On the whole we find a better 850 mb temperature forecast by the ECM model in the medium waves, a fact that was less obvious in the skill scores of the height fields (Fig. 7.1.1). On the other hand the correlation coefficients of the long waves (panel f) show some advantages for the GFD model.

Again, we did a second comparison based on a different set of analyses (Operational analyses by Deutscher Wetterdienst). Here the choice of another set of analyses leads to greater differences. The advantage for the GFD model, which was found for the long waves, disappears and even reverses to a disadvantage, and the advantage for the ECM models temperature predictions in the medium waves becomes even more pronounced. Both changes lead to a clear advantage for the ECM model for the total field when using the German analyses for comparison. There is also a reduction in predictability of 850 mb temperature for the GFD model of about half a day.

The variability in the scores for temperature at the 850 mb level makes it difficult to give a judgment on the relative performance of the models in the long waves, but the advantage for the ECM model for the medium waves (4-9) seems to be real. One could have chosen another level with less uncertainty in the scores (e.g. 500 mb), but on the other hand it is the 850 mb temperature field which is more interesting for weather forecasting.

CONCLUSION

The large-scale skill scores show a period of predictability, according to our criteria, of about 5 to 6 days, they indicate a slight advantage for the ECM model in predicting the mostly baroclinic waves with wavenumbers 4 to 9 and a slight advantage of the GFD model for the long waves. The later statement is mainly based on the smaller variability in the skill of the GFD model. Its better performance in the scores after 7 days need not be taken into account,since this is already beyond the period of predictability.

There is some indication that the forcing of the long waves by mountains is handled better by the models than other large scale forcings.

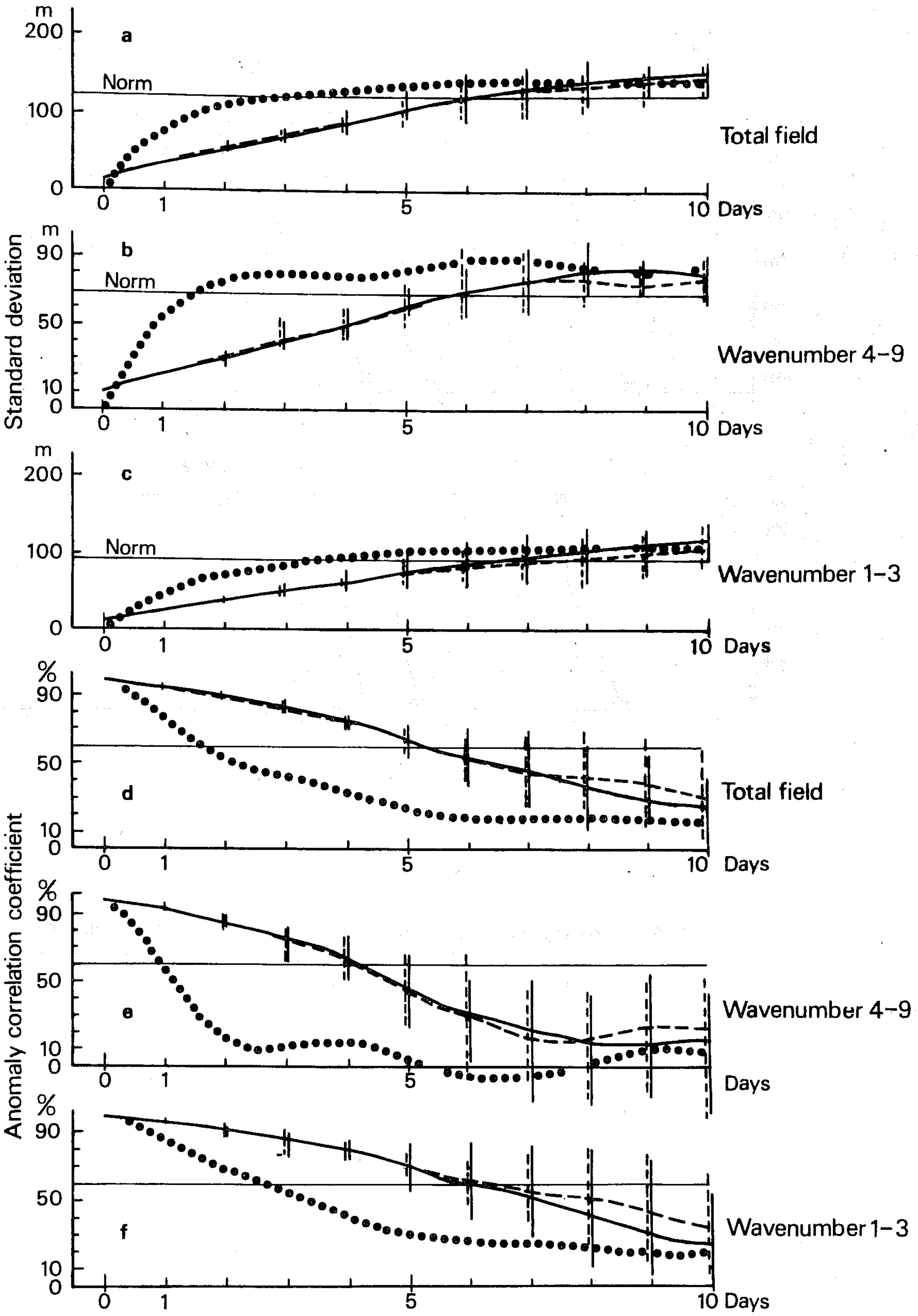


Fig. 7.1.1 Horizontal and vertical mean of standard deviations and anomaly correlation coefficients of the height fields for 7 February cases. Average for the area 20 °N to 82.5 °N and for the levels 1000 mb to 200 mb. Solid lines: ECM-forecast, dashed lines: GFD-forecast, dotted lines: persistence forecast, thin solid lines: NORM in panels for standard deviation and 60% for panels with anomaly correlations.

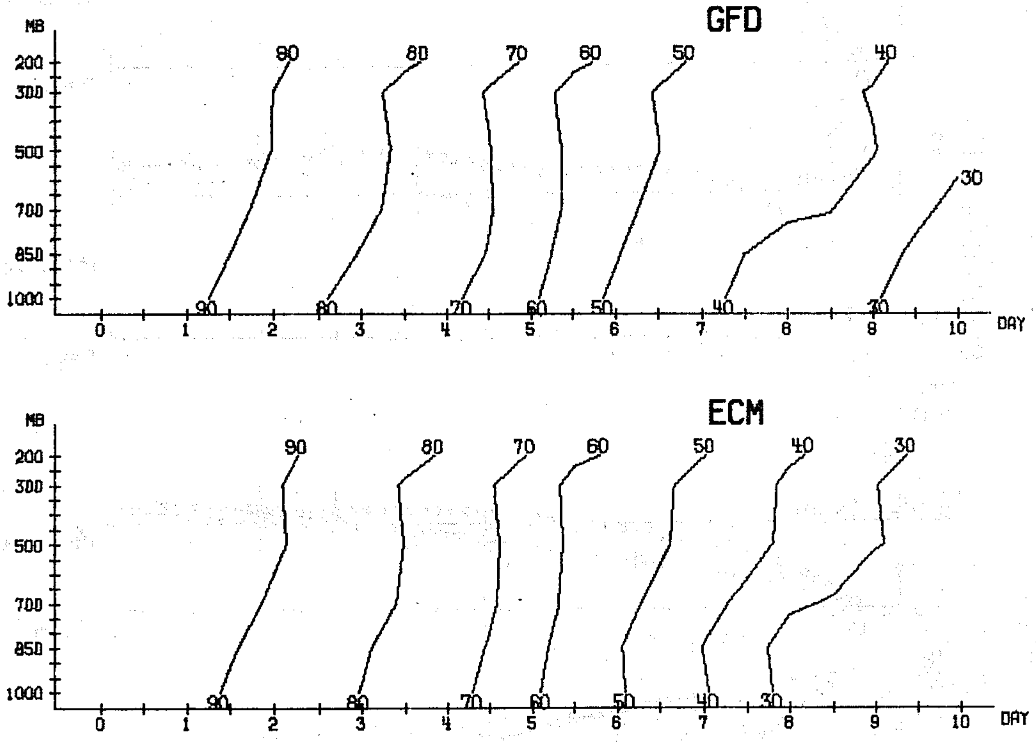


Fig.7.1.2 Anomaly correlation coefficient (%) of height field for all standard levels. Mean for 7 February cases. Averaged between 20 °N and 82.5 °N.

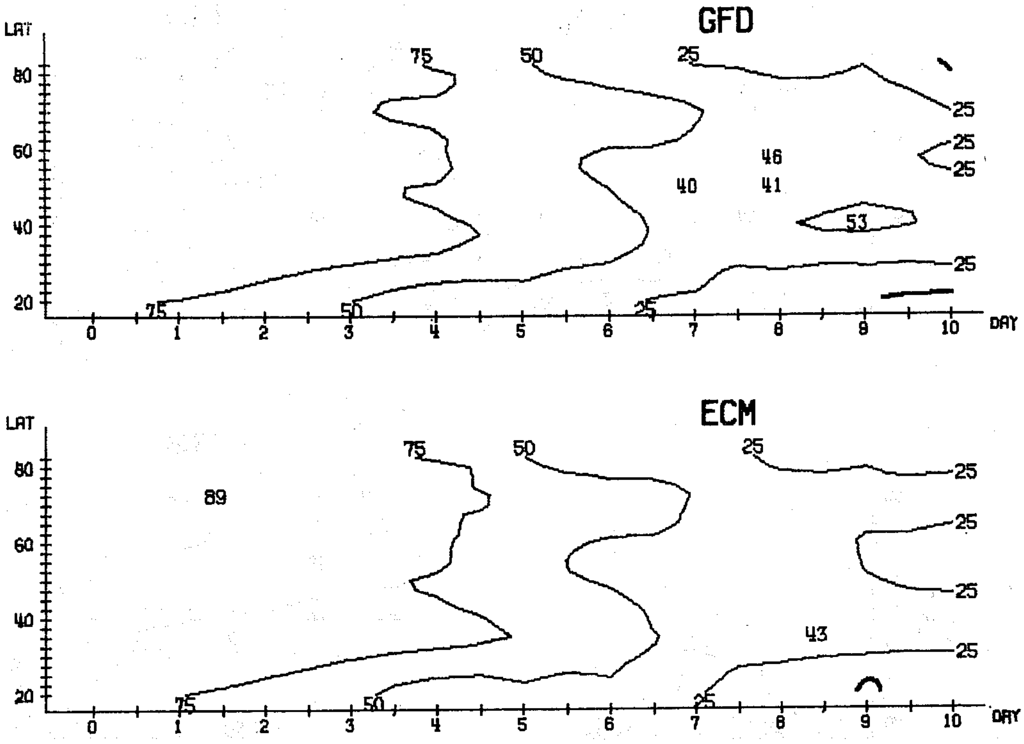


Fig.7.1.3 Anomaly correlation coefficients (%) of height field. Mean for 7 February cases and averaged between 1000 and 200 mb, as a function of latitude and time

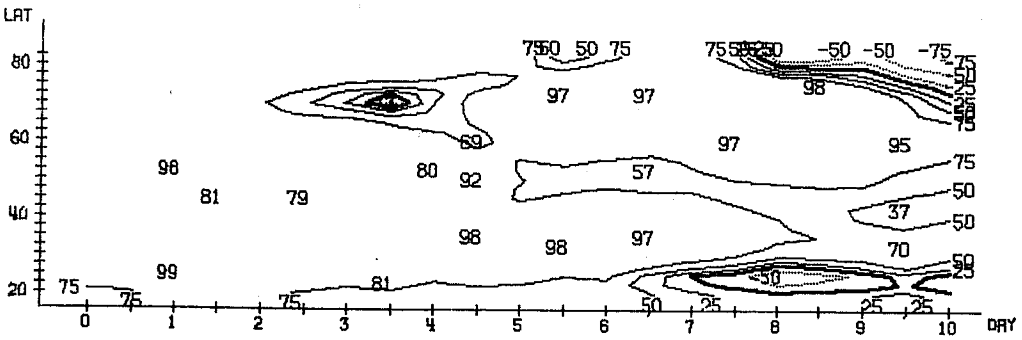


Fig.7.1.4 Anomaly correlation coefficients (%) of the 500 mb height field, contribution by wavenumber 1-3 only. ECM forecast based on the initial data of 15.2.76

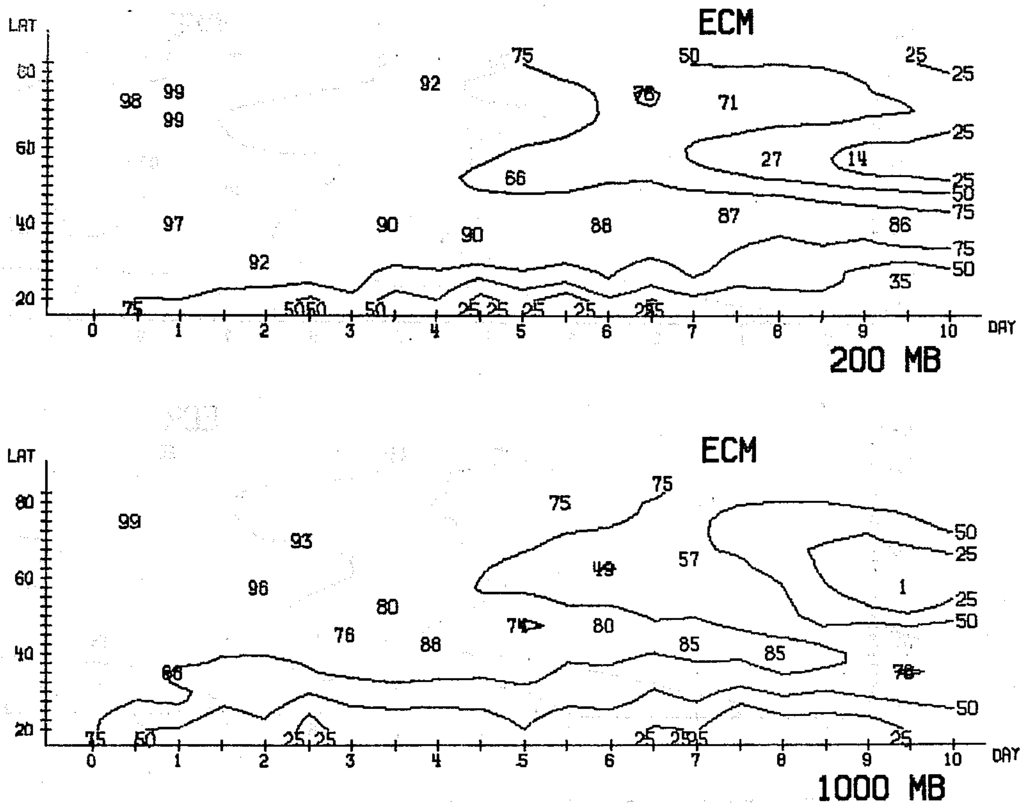


Fig.7.1.5 Anomaly correlation coefficients (%) of height field of the quasi-stationary waves (mean of 7 cases), as a function of latitude and time

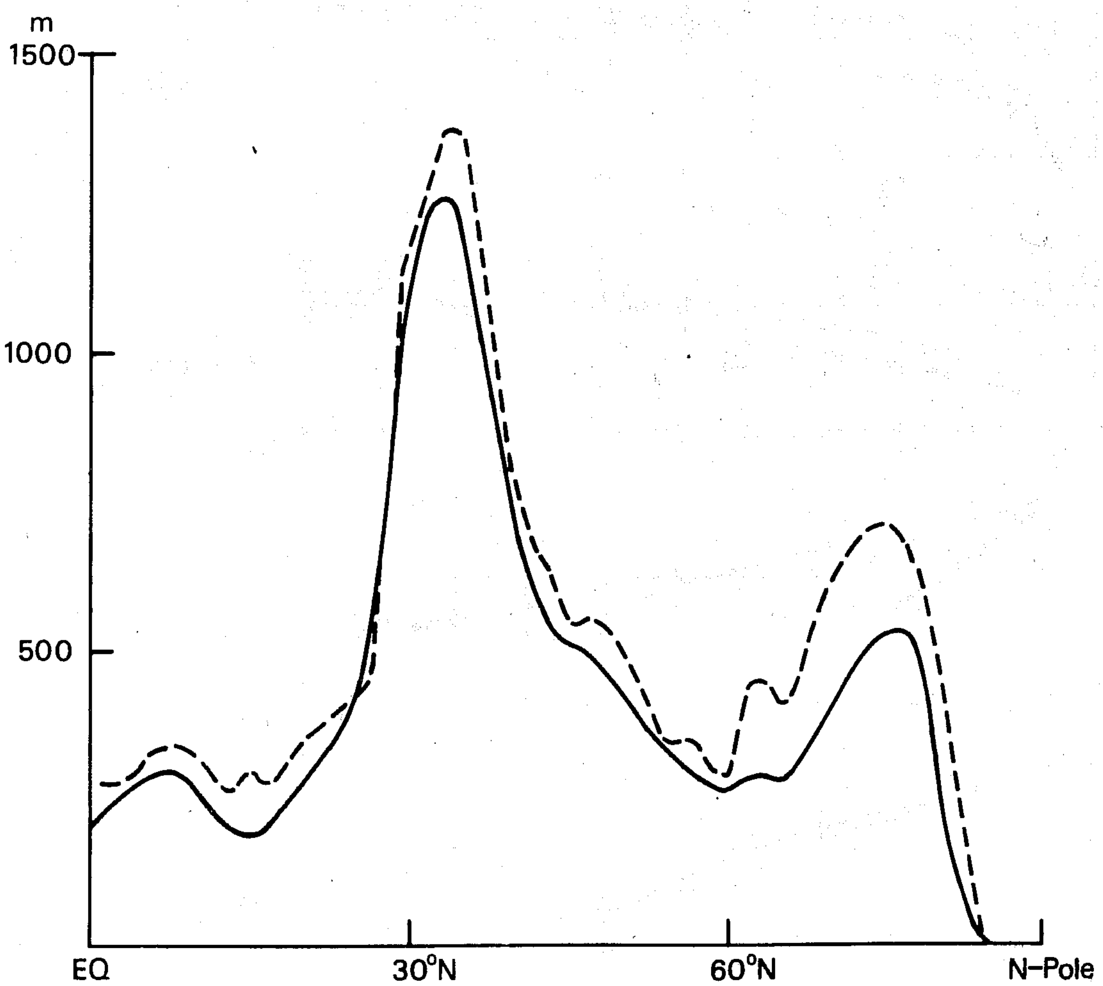


Fig. 7.1.6 Standard deviation of topography at each latitude. Solid line: Model's topography. Dashed line: RAND topography

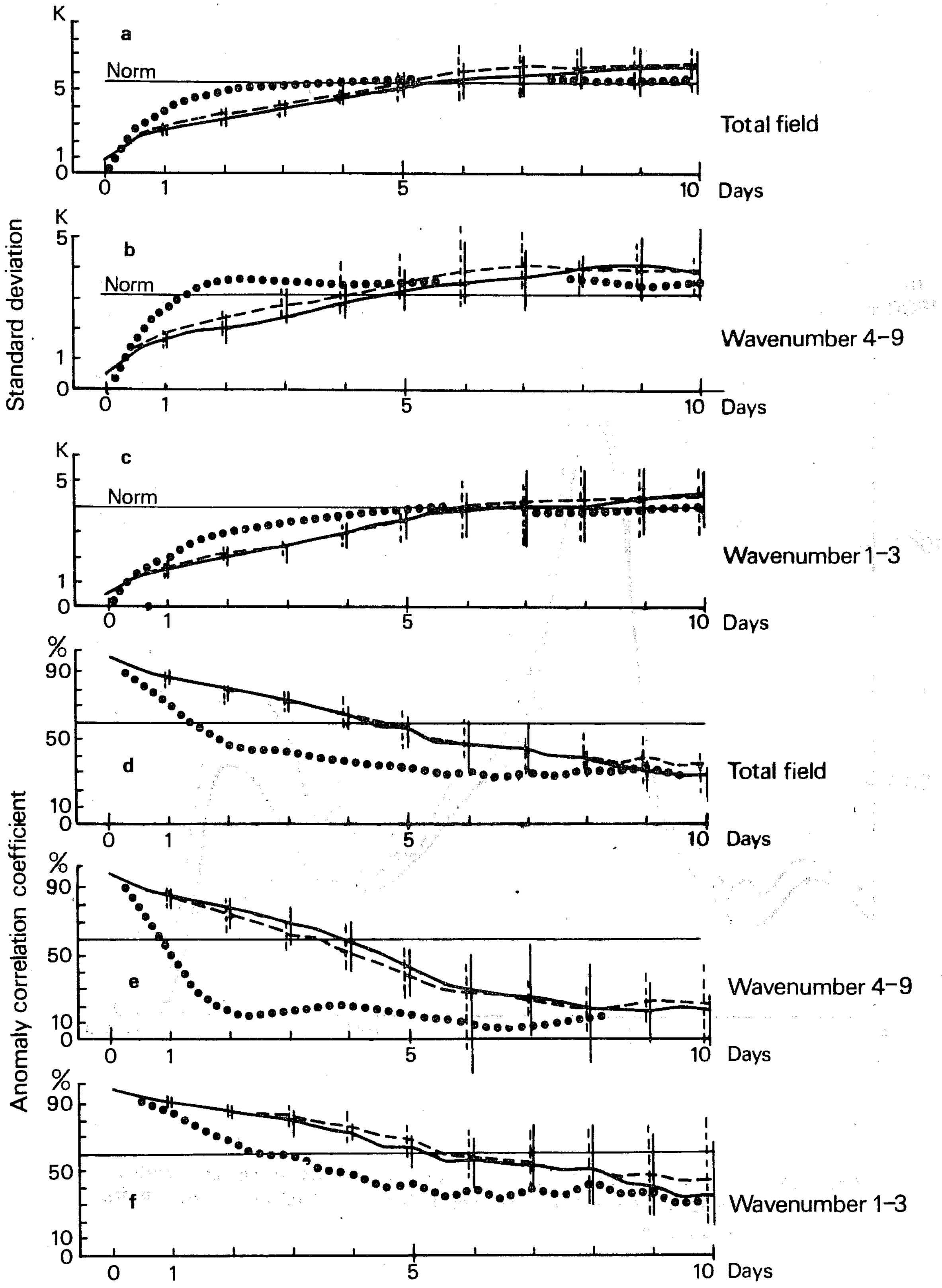


Fig.7.1.7 Horizontal mean of standard deviations and anomaly correlation coefficients of the 850 mb temperature field for 7 February cases. Average for the area 20°N to 82.5 °N. Solid lines: ECM-forecast, dashed lines: GFD-forecast, dotted lines: persistence forecast, thin solid lines: either NORM or 60%.

7.2 Objective measurement of differences between the models

As mentioned in the introduction one of the main motivations for this study was to estimate the relative performance of the two parameterization schemes. The indications from the results are that the forecast quality of the models is rather similar. This does not imply that the forecasts are the same and in this section we examine the rate at which the forecasts diverge from each other.

Fig. 7.2.1 shows the average over the 7 cases of the anomaly correlation coefficient for the ECM forecast, for the persistence forecasts and for the ECM forecasts correlated with the GFD forecasts. We see that the correlation between the models for all wavenumbers has reached 60% by day 10. For waves 4 - 9 the correlation coefficient reaches 60% by day 7 while for the long waves it is 65% at day 10.

Even in the wavenumber group 10 - 20 the differences between the models grow quite slowly. The correlation in this wavenumber band reaches 60% at day 4½.

Fig.7.2.2 shows the corresponding results for the standard deviation. The differences between the models reaches the NORM at day 9½ for the total field. The NORM is reached in the medium waves after day 8 and has not been reached by day 10 in the long waves.

This could be interpreted as meaning that one model is a useful predictor for the other up to 10 days.

This result may have wider implications if we regard the two sets of forecasts as a particular type of predictability experiment. Let us assume that one model is the real atmosphere and the other an almost ideal forecast system where the only errors arise from our ignorance of the parameterization scheme. Then the result indicates that the predictability time is 10 days. If this interpretation is valid then it would suggest that in the real world the predictability should be less than 10 days when we take into account important error sources such as initial data errors and finite differencing errors.

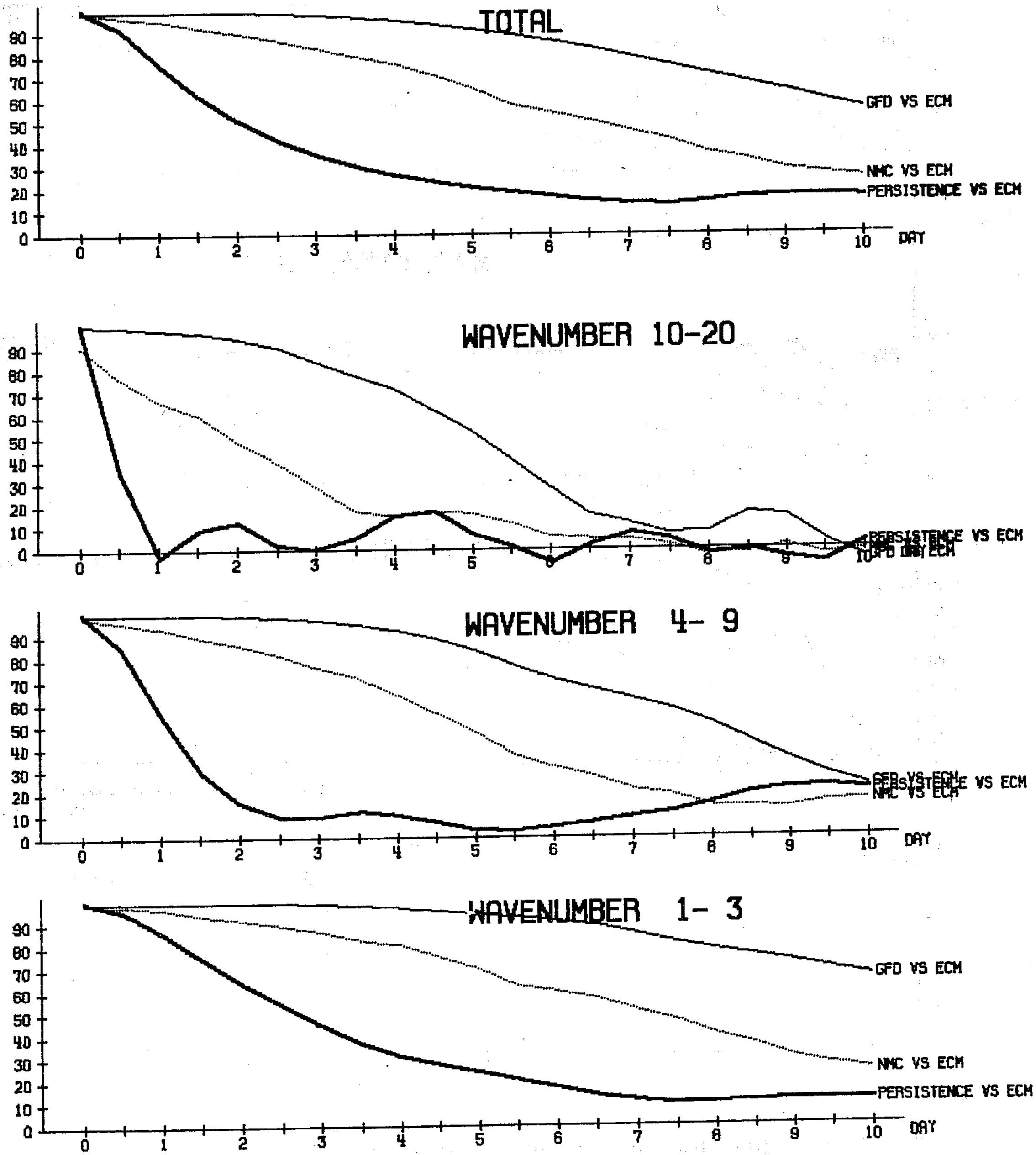


Fig.7.2.1 Anomaly correlation coefficients (%) of height. Mean for the troposphere 1000-200 mb and 20 ^{O}N - 82.5 ^{O}N. 7 February cases. The thick line is the result for the ECM model verified against the initial day, the thin line is the result for the GFD model verified againt the ECM model and the dotted line is for the ECM model verified against the valid observations

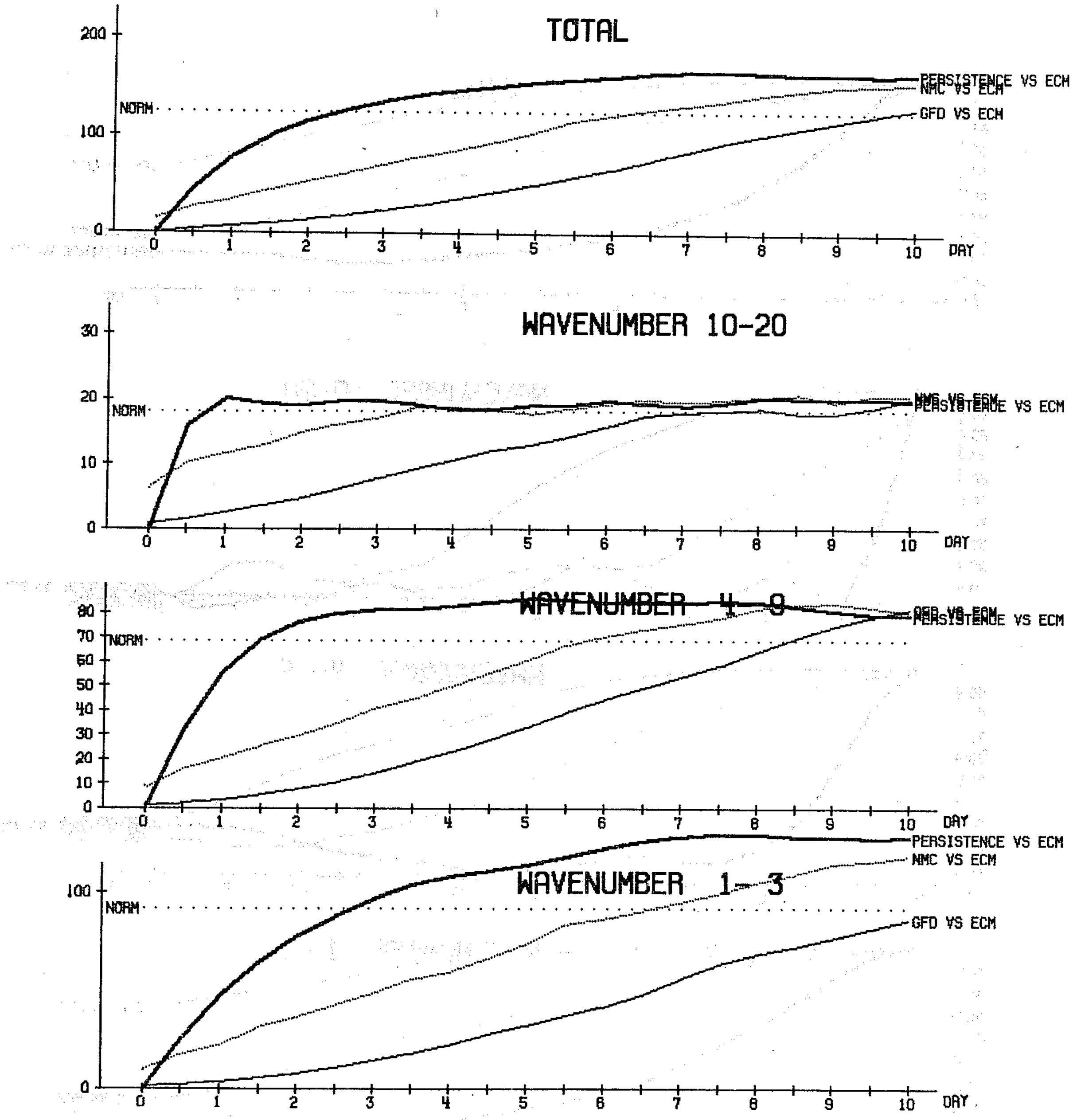


Fig.7.2.2 Standard deviation of height fields. Mean for the troposphere 1000-200 mb and 20 ^{o}N - 82.5 ^{o}N. 7 February cases. The thick line is the result for the ECM model verified against the initial day, the thin line is the result for the GFD model verified against the ECM model and the dotted line is for the ECM model verified against the valid observations

7.3 Objective cyclone verification

One of the main goals in weather prediction is the forecast of the positions of cyclones. In this section we want to see how our experiments behaved in this respect. To get objective measures we first had to make some definitions. A relative minimum of at least 20 m in the 1000 mb height field was used here as the definition of a cyclone. A position of a cyclone was forecast correctly if, at the time under consideration, the centres of the forecast cyclone and of the observed cyclone were less than 1000 Km apart. A forecast for a cyclone was also considered correct if within a 12 hour time interval the centres of the forecast and observed cyclones were less than 500 Km apart. To measure the skill in forecasting the cyclone positions a threat score was used.

$$\text{cyclone skill} = 100\% \; \frac{NC}{NO+NF-NC}$$

with:

NC = Number of correct cyclone positions
NO = Number of observed cyclones
NF = Number of forecast cyclones.

A skill of 100% will only be reached when the number of observed cyclones is equal to the number of forecast cyclones and when all cyclones are correctly forecast.

Fig. 7.3.1 shows the results from such a score in one forecast. The 100% level is not reached at the initial day due to interpolation to sigma levels, initialization, and interpolation or extrapolation back to pressure levels. The strongest differences occur over mountainous areas.

This case is typical in that the scores cluster in two groups, one with values more than 40% for the first 4 or 5 days and another one with mostly less than 30% at the second half of forecast period. Looking through all cases the level of 30% seems to be the most suitable level for separating good and bad forecasts. This level is reached normally after about 4½ days with one extreme deviation on the case which started on the 15 February 1976 where this level was reached already after 2½ or 3 days for the ECM- or GFD-model, respectively. In Section 7.1 this case was pointed out to be the best forecast for the long waves and the worst for medium waves at 1000 mb height fields according to correlation coefficients and standard deviation.

Let us interpret the results of this score in terms of number of correctly forecast cyclones. The normal number of cyclones, as defined above, is observed to be about 9 to 10. The forecasts tend to have one or two more cyclones than the analyses. If 5 out of 10 cyclones are correctly forecast a score of 33% will result. This would seem to be a reasonable criterion for skill.

For an overview over all cases we show the mean and the variability of cyclone skill in Fig. 7.3.2. The decrease of skill with time is almost linear up to day 7 when a final value of about 20% is reached. A 20% level means in terms of cyclones that 3 out of 9 cyclones were forecast correctly and it seems that this score will be reached easily as long as the main areas of cyclone activity are forecast correctly.

If we use the level of 30% as a measure for predictability, as suggested before, we will get a period of 5 or 5½ day for the ECM- or the GFD-forecast respectively, which is 2 days earlier than the time when the final level of about 20% is reached. The small advantage for the GFD-forecast is not only to be seen at that 30% level but also for the whole period between day 3 and day 5½ for the mean values and between day 2½ and day 5 for the minimum values (corresponding to the worst cases).

Summing up one can say that this skill score shows some advantage for the GFD-forecast but we do not have enough experience to judge if it is significant. The range of predictability is in accordance with the results of Section 7.1.

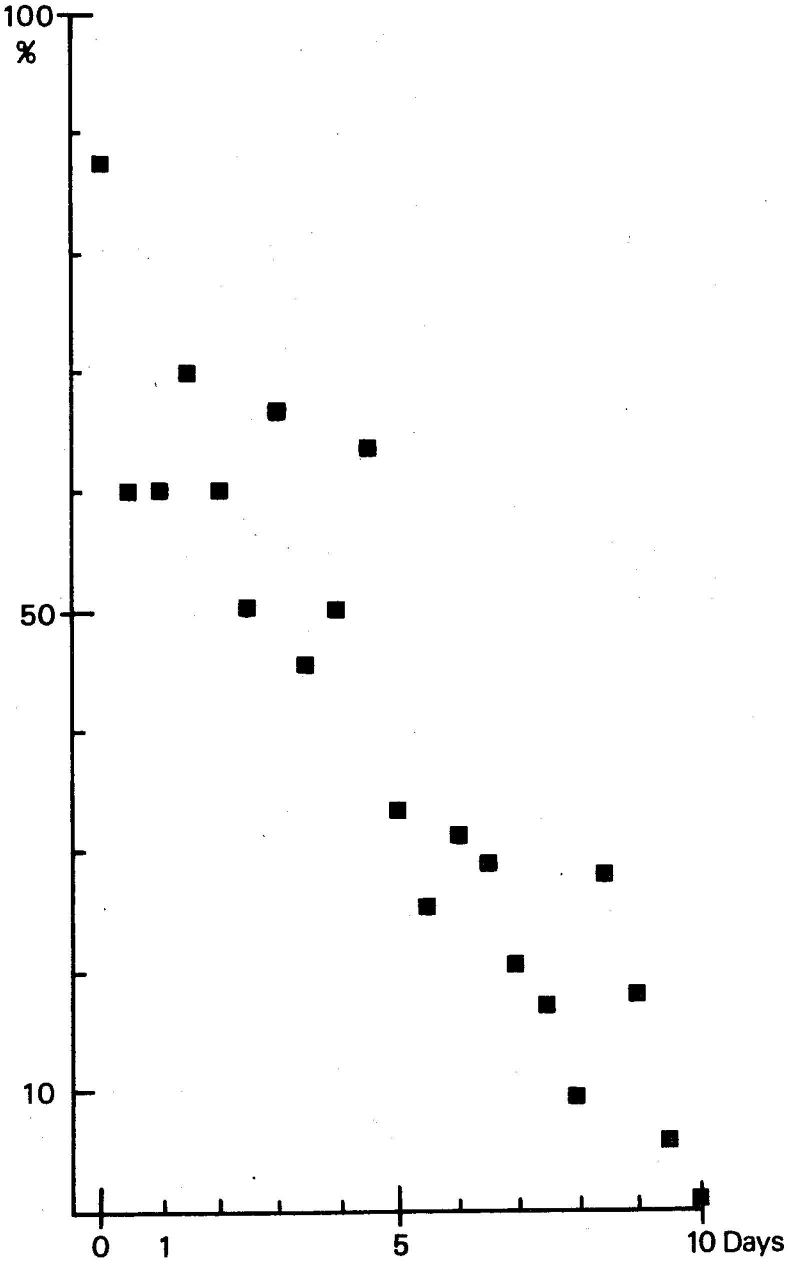


Fig. 7.3.1 Cyclone skill of a GFD-forecast based on data of the initial date 12.2.76.

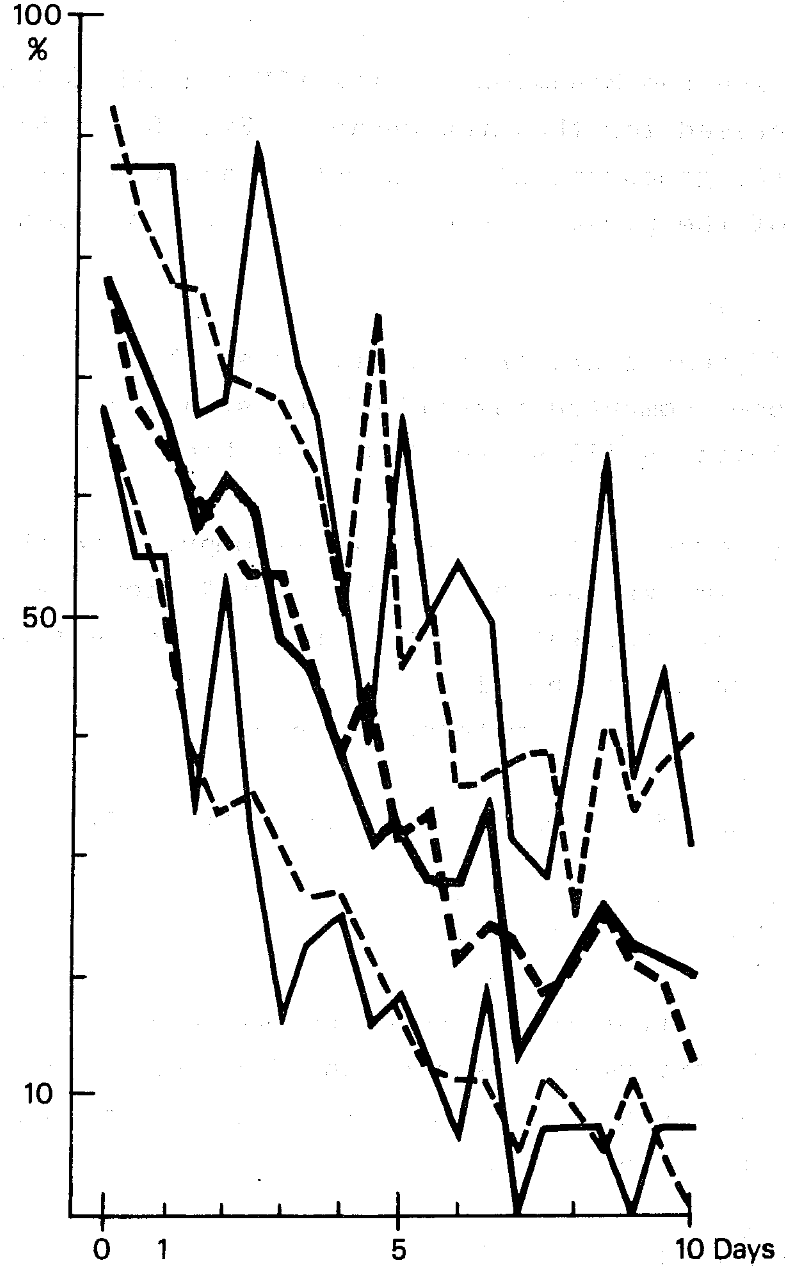


Fig.7.3.2 Cyclone skill, 7 cases.
For both experiments:

solid lines: GFD-forecasts
dashed lines: ECM-forecasts

bold lines: mean values
thin lines: minimum or maximum value

8. Skill scores over Europe

8.1 Introduction

It is well known that objective verification of model performance may give somewhat different results depending on the the area chosen for computing skill scores.

In this section the behaviour of the ECM and GFD models will be discussed for the area shown in Fig. 8.1.1 in respect to the geopotential, temperature and wind fields using some of the parameters described in the previous section.

The main differences are that in this case the root mean square has been computed instead of the standard deviation; and the following skill scores have also been used:

i) correlation of tendencies-computed in the same way as the anomaly correlation taking into account the changes from the initial value of the field instead of the deviation from the climatological mean.

ii) the S_1 score (Teweles and Wobus 1954)

$$S_1 = 100\, E\left(\overline{\frac{|\delta_G|}{|G_m|}}\right)$$

where δ_G is the error in the observed pressure gradient and G_m is the maximum of observed and forecast pressure gradients, and E denotes a sum over all experiments and the bar a sum over all points.

8.2 Geopotential

The time changes of the mean RMS error of the geopotential field in the troposphere is shown in Fig.8.2.1 for both models. The persistence and the level of the climatological variances (NORM) have been added to give reference values.

The ECM model shows some advantage over the GFD forecast, at least until 4 - 5 days when the NORM level has been reached. The range of variability (vertical bars) sharply increases at the same time indicating the lower reliability of the forecasts after this period.

The vertical distribution of the RMS error (Fig. 8.2.2) is very similar for both the models. The error grows most rapidly at the highest levels. In latitude the fastest error growth occurs in the belt between 58 N and 73 N on the 4th day of the forecast in the EC model (the GFD model is slightly worse in this respect) and then affecting all areas (Fig. 8.2.3). Considering the S_1 score (Fig. 8.2.4) the performances of the models are almost indistinguishable. The distribution of errors of the gradients of the geopotential field must therefore be very similar in the average for both models.

The correlation coefficients for anomalies and for tendencies can be compared in Figs. 8.2.5 and 8.2.6. The tendency correlation shows erratic behaviour at the outset due mainly to the smallness of the actual changes. It is satisfactory that the relative behaviour of ECM and GFD models is very similar whichever of these two different correlations is used. Miyakoda et al (1972) pointed out that for large ensembles and long forecasts the anomaly correlation coefficients tend to zero while the tendency correlation coefficients tend to ½, which explains the different asymptotic levels between both correlation coefficients.

The percentage range of variabilities of the correlation scores is as large as for the RMS scores but the vertical distribution of the correlation is more uniform (Fig. 8.2.7, 8.2.8).

If we take the 60% level of anomaly correlation to determine the range of predictability we see an advantage for the ECM-forecast for all levels. The best scores can be found for the lower troposphere, i.e. 700 and 850 mb, where we get a range of 5 or 3½ days for ECM or GFD forecasts respectively. For higher levels we get lower values going down to 3¾ or 3 days respectively at 200 mb. This behaviour is quite different to that of the hemispheric values in Section 7.1 where we had lower scores at 850 and 1000 mb compared to higher levels. We don't have guidelines for tendency correlation coefficients but the variability in Fig. 8.2.6 indicates already on day 3 high uncertainties.

8.3 Temperature

Most of the results found for the geopotential are also true for the temperature field. There is a slight increase of the time necessary to reach the NORM level (Fig. 8.3.1) while the vertical distribution of RMS (Fig. 8.3.2) is somewhat different from the corresponding parameter computed for the geopotential due to the error minimum at about 300 mb (tropopause) where the temperature variances are small. The correlation of anomalies (Fig. 8.3.3) show a reduction of the time period in achieving the 60% value to 3 - 4 days while the comparison between the correlation of tendencies for the temperature (Fig. 8.3.4) and for the geopotential (Fig. 8.2.6) do not show any major differences. No vertical cross-section of correlation is shown here but from Fig. 8.3.3 we can already find out that the forecasts at the 850 mb level are more reliable than at the 500 mb level.

8.4 Wind

Very few indications will be added here on the RMS of the wind vector (Fig. 8.4.1). The two models are almost identical in this regard. The advantage for the EC model in the vertical average is very small, and is not discernible at the 500 mb level. On the other hand the variabilities appear to be rather low giving the impression of a more uniform performance of both the models in respect to the wind field.

The vertical distribution of the error (Fig. 8.4.2) indicates that the maximum of error occurs on the 300 mb level which is the level of maximum variance.

8.5 Conclusion

It is very hard to draw a definitive conclusion about the relative behaviour of the two models examined over Europe. They appear to be very similar in the wind description and in the gradient distribution. Some small advantage for the ECM model can be recognized at least until 4 - 5 days in the height and temperature fields. The strong increase of variability in the scores after 4 days seems to suggest that consideration of the models' performance for a more extended period based only on these skill scores would be pointless. For the scores presented, our results seem to be at least comparable with results published by national weather services* for the same month and area which are presently available. However the smallness of our sample prevents a detailed comparison. In Section 9 we will see that Europe is one of the areas where the models have large systematic errors, so we may gain a misleading impression of the performance of the models by considering only the results from this region.

---

*WMO - Numerical Weather Prediction - Progress Report for 1976.

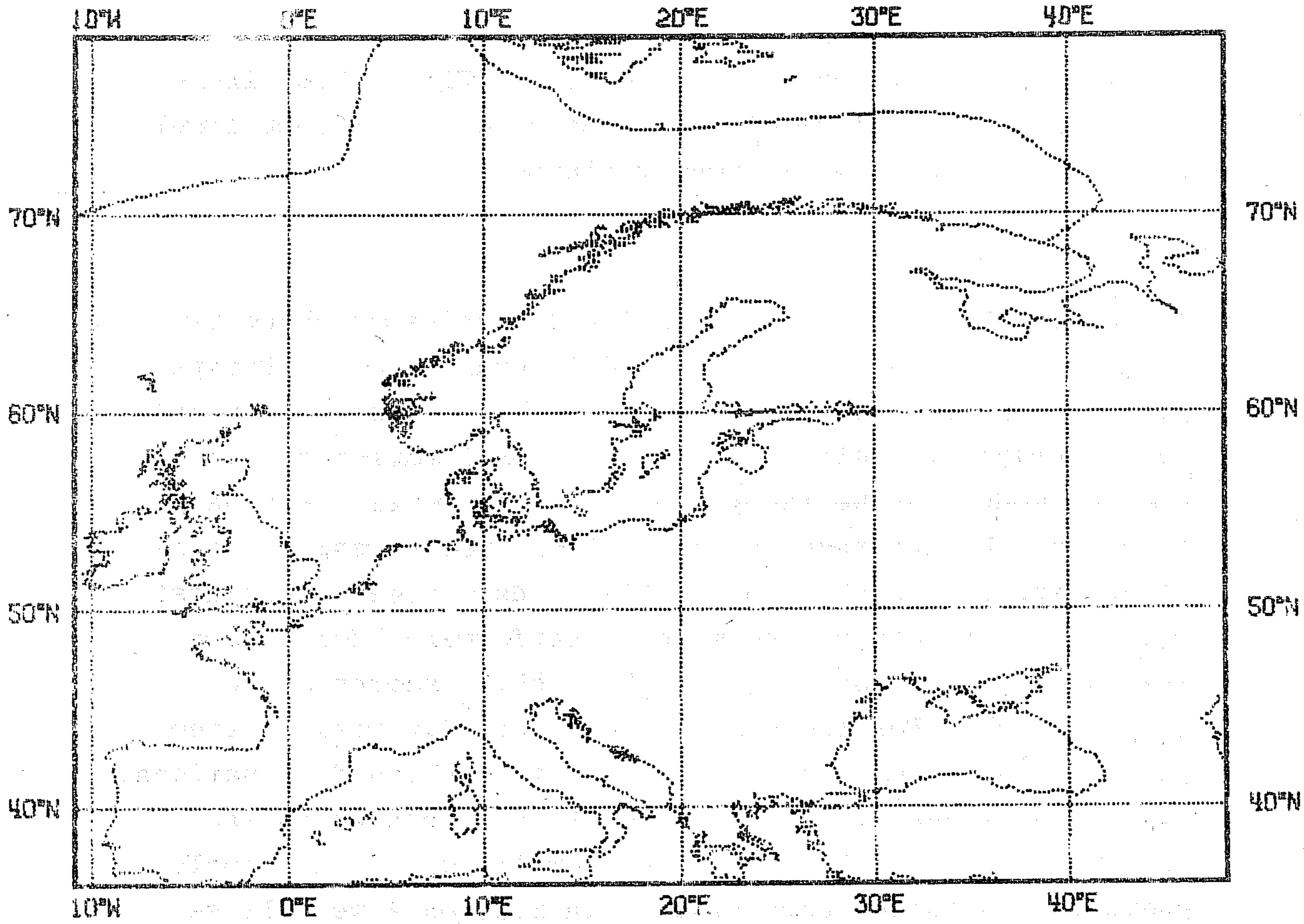


Fig.8.1.1 Verification area

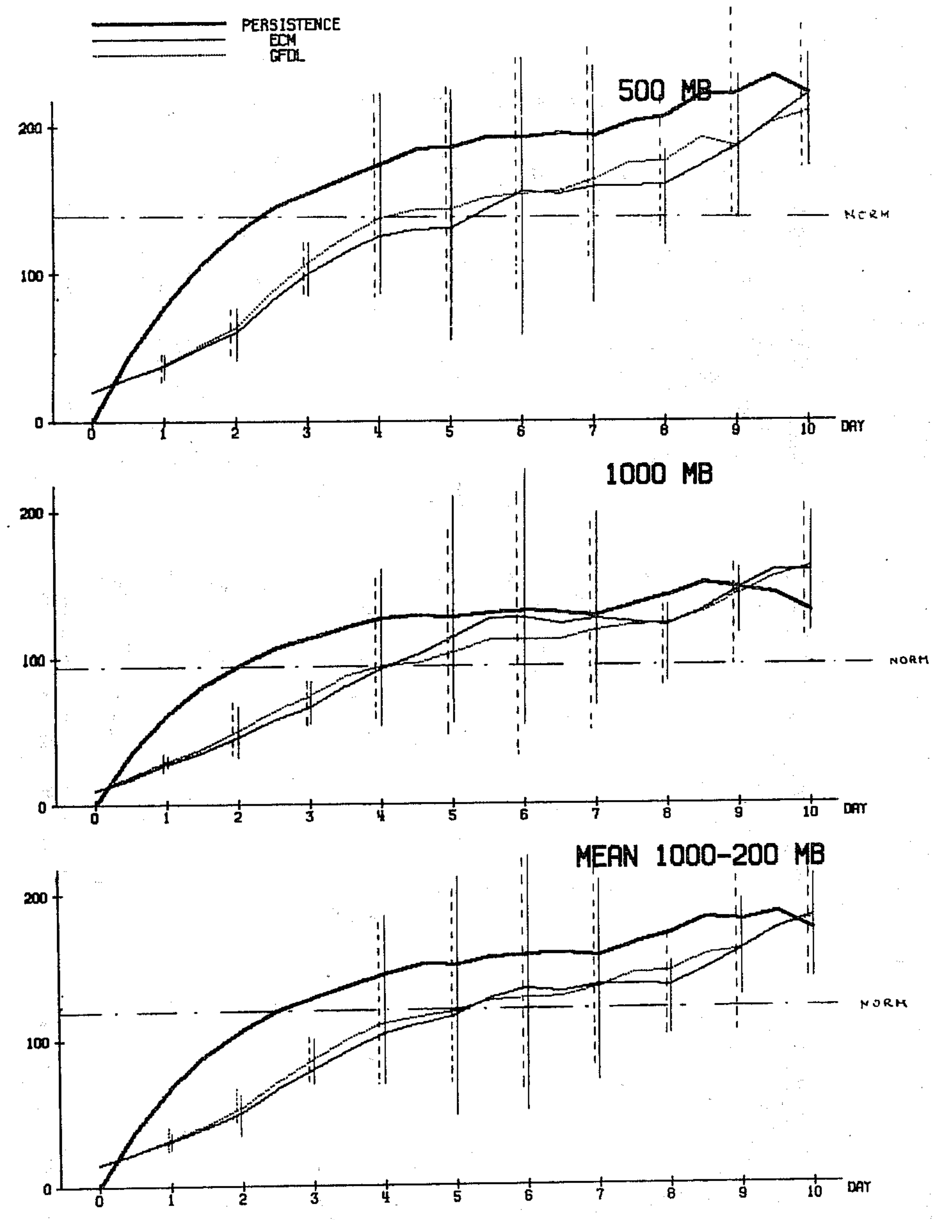


Fig.8.2.1 RMS of geopotential height (m) area mean

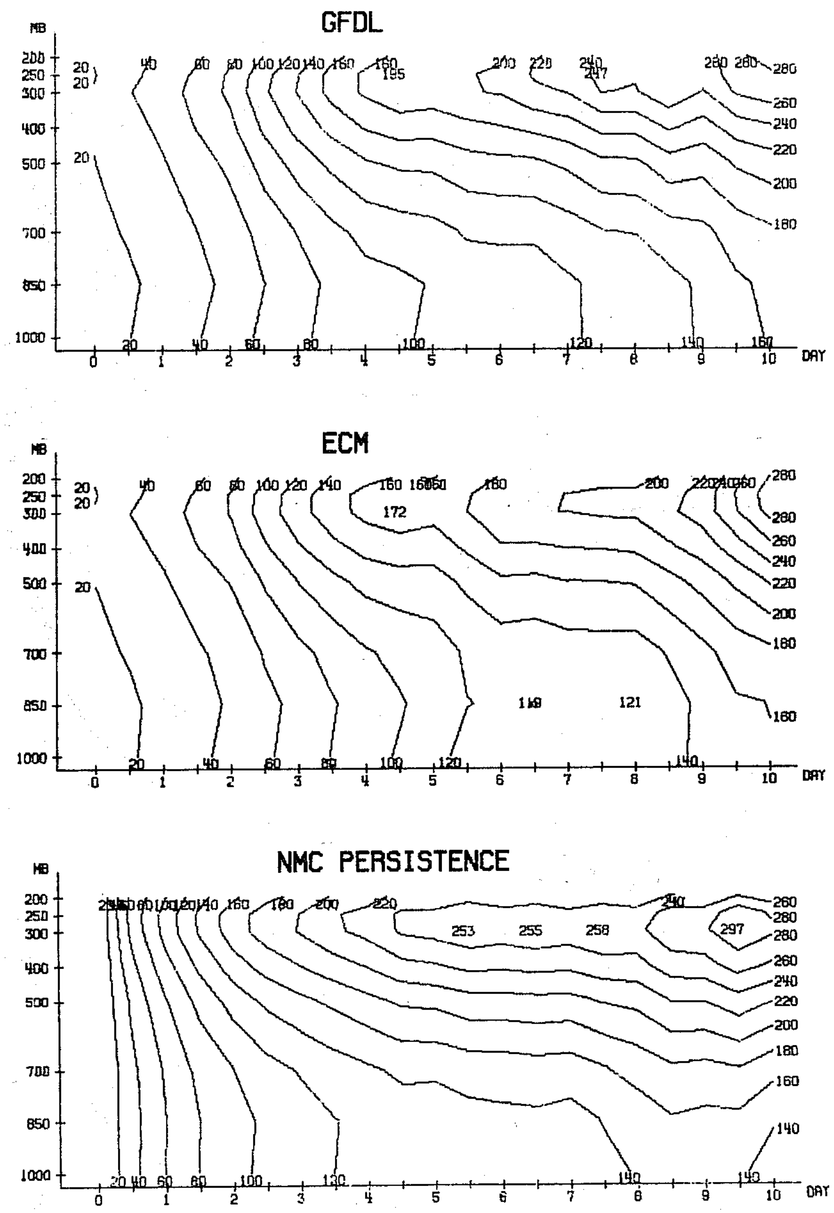


Fig.8.2.2 RMS of geopotential height (m) area mean

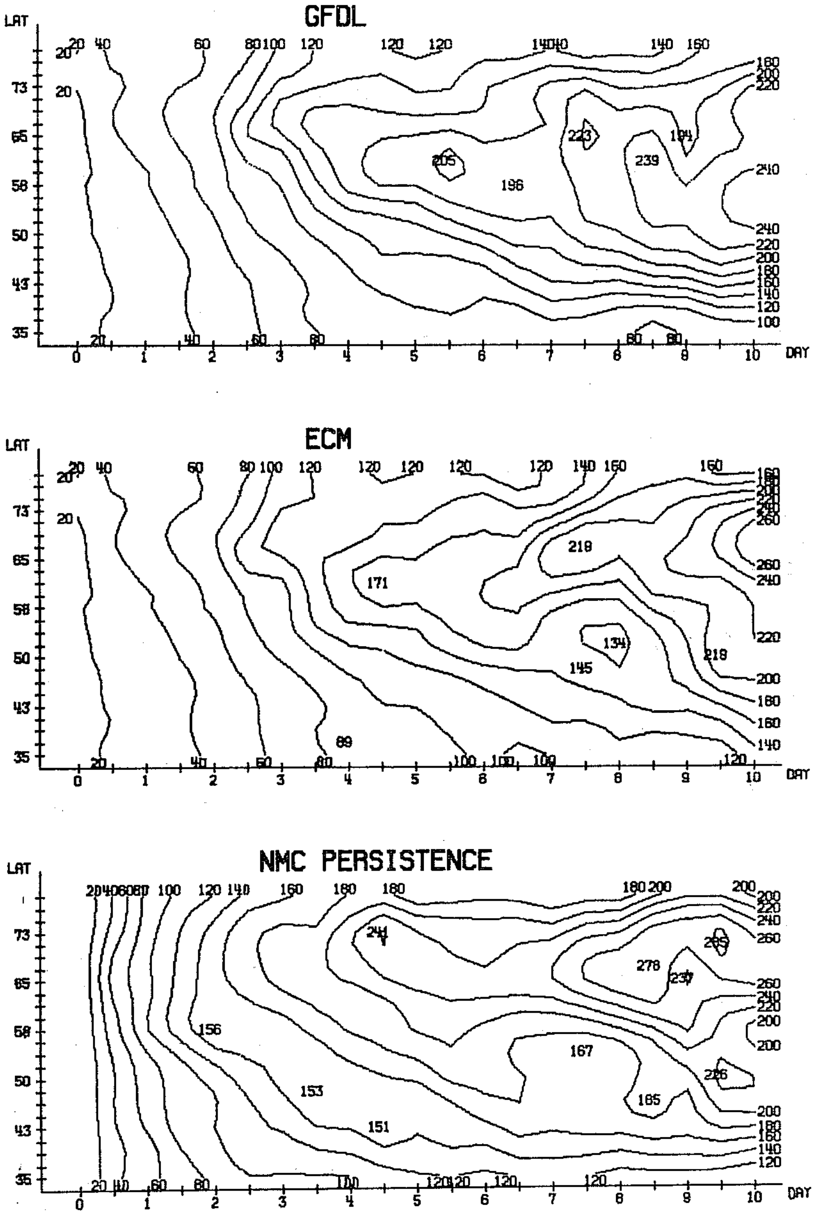


Fig.8.2.3 RMS of geopotential height (m)
mean between 1000-200 mb

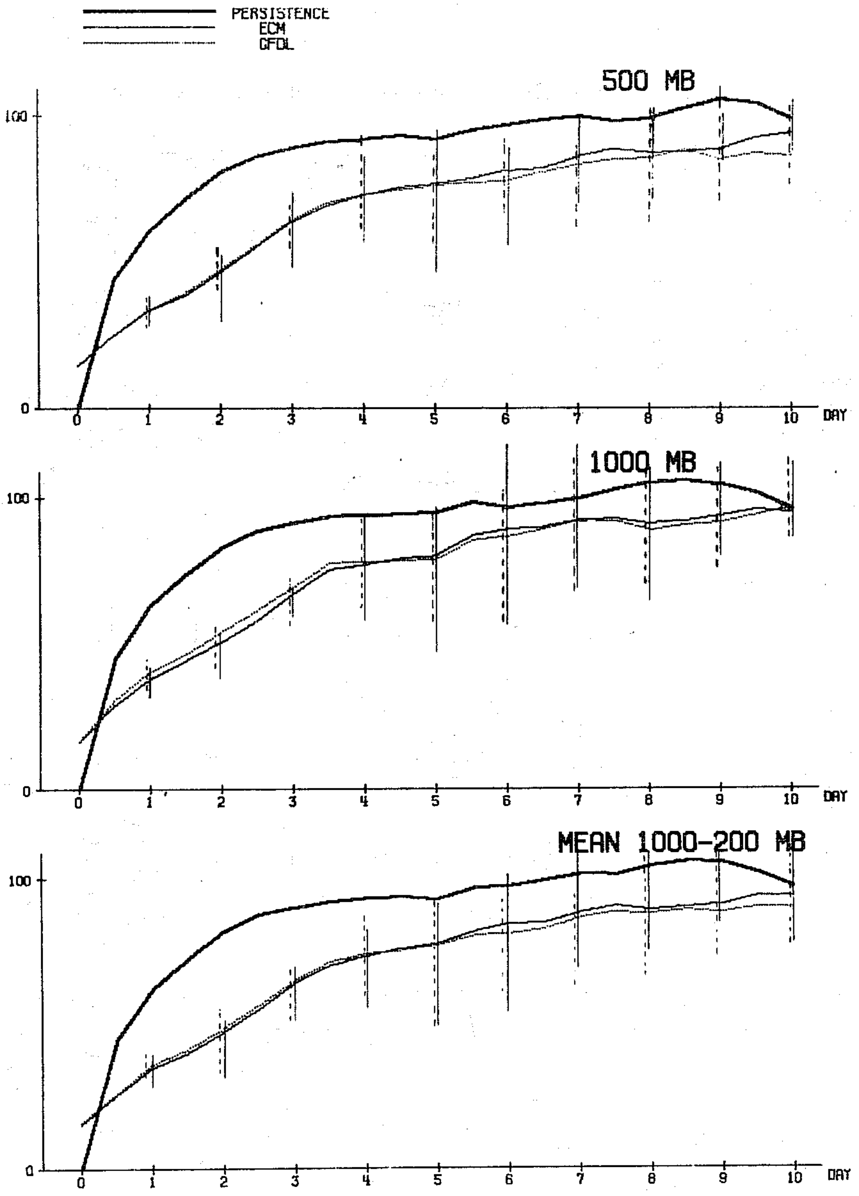


Fig.8.2.4 S_1 score of geopotential height

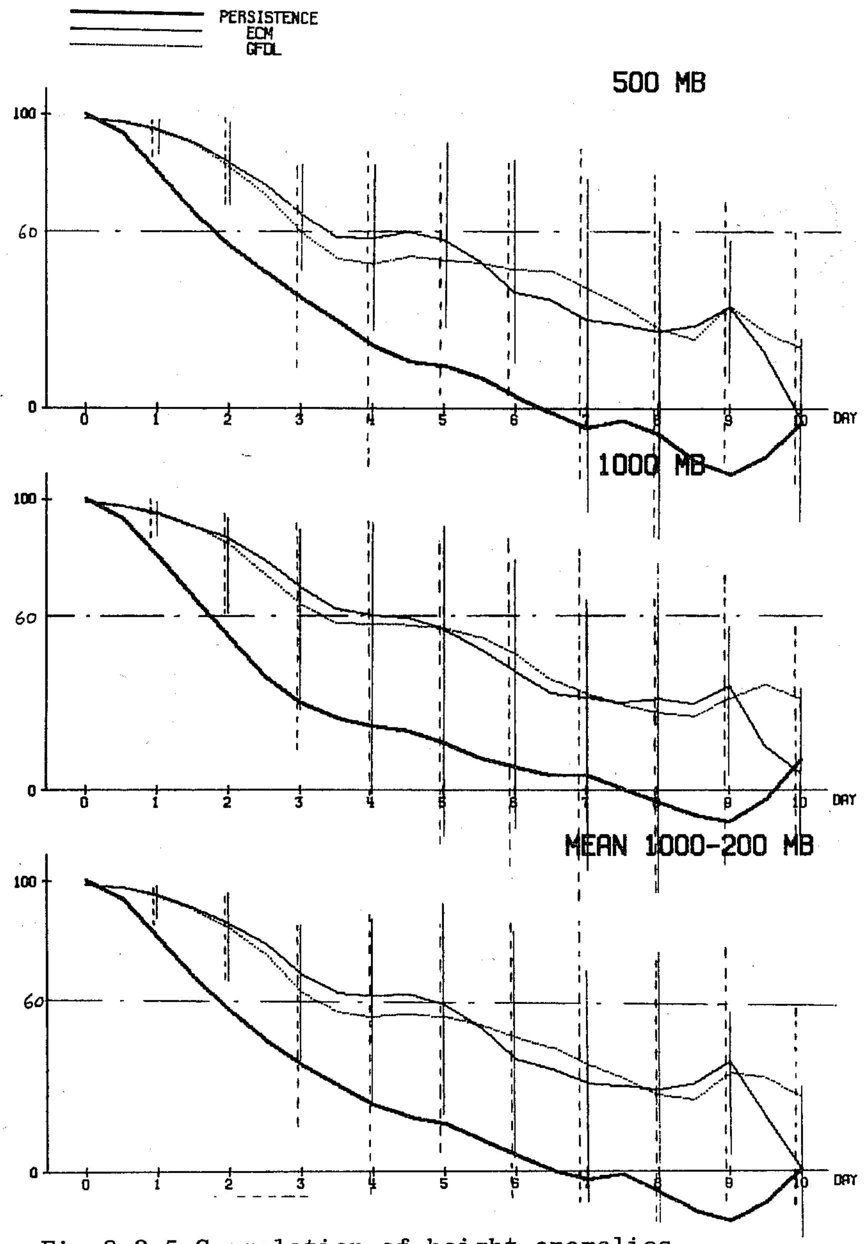


Fig.8.2.5 Correlation of height anomalies
area mean

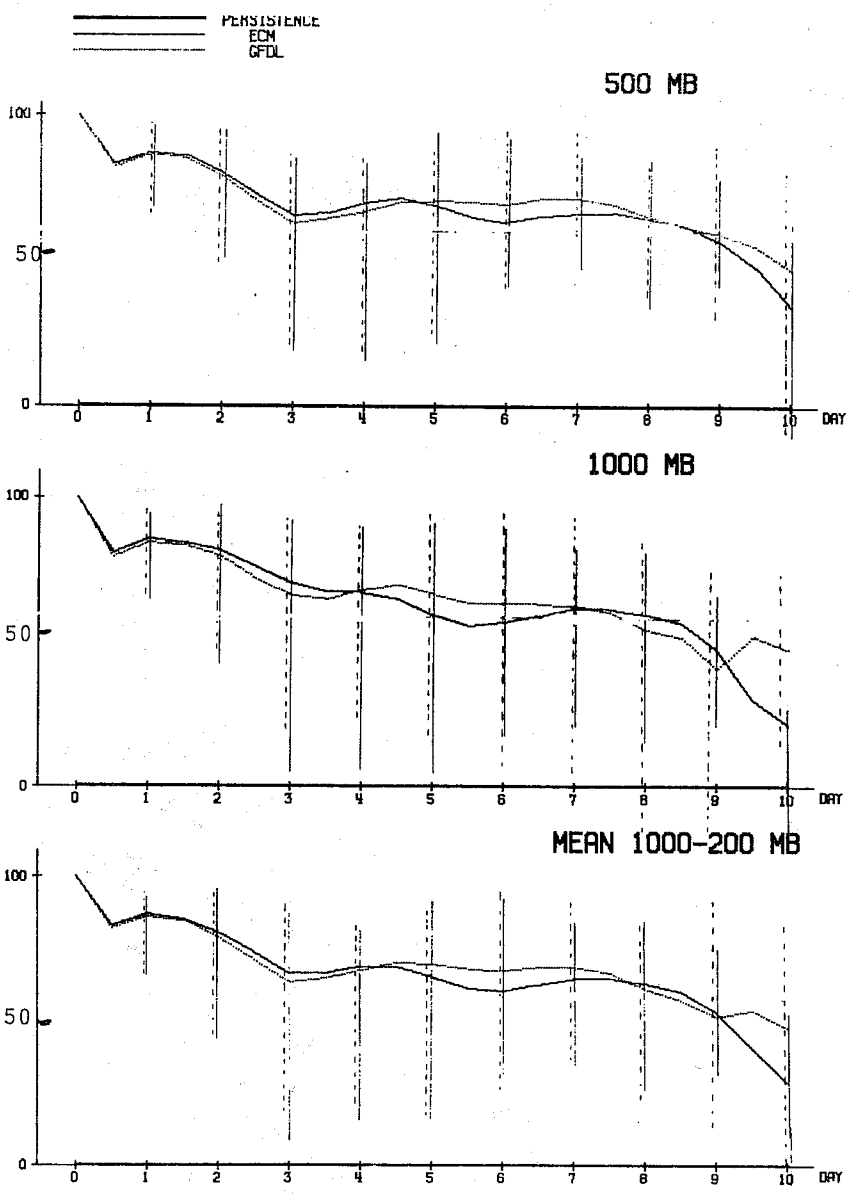


Fig.8.2.6 Correlation of height tendencies
area mean

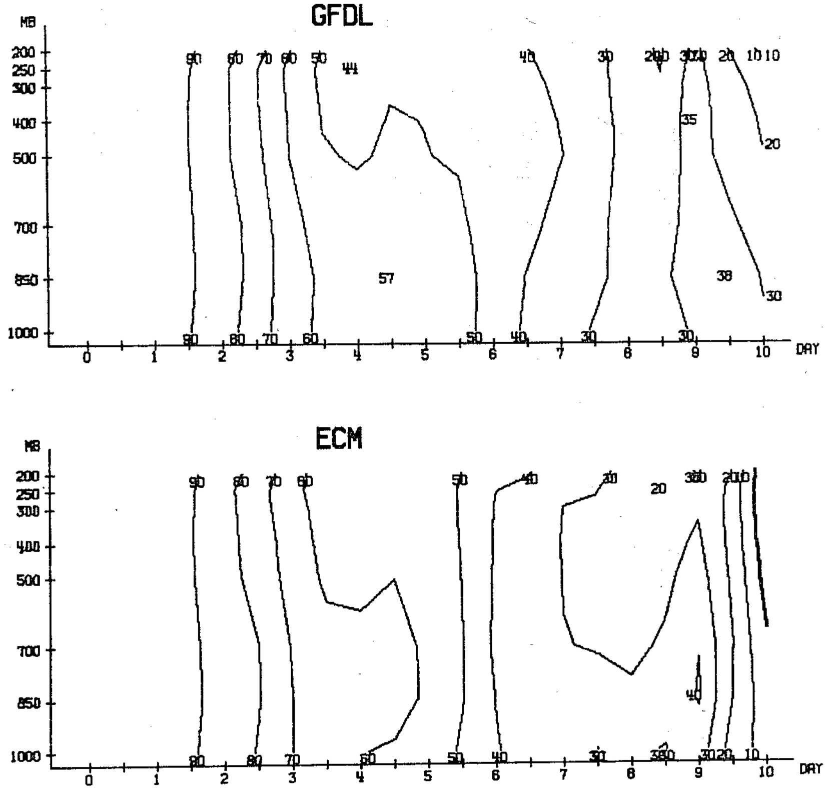


Fig.8.2.7 Correlation of height anomalies
area mean

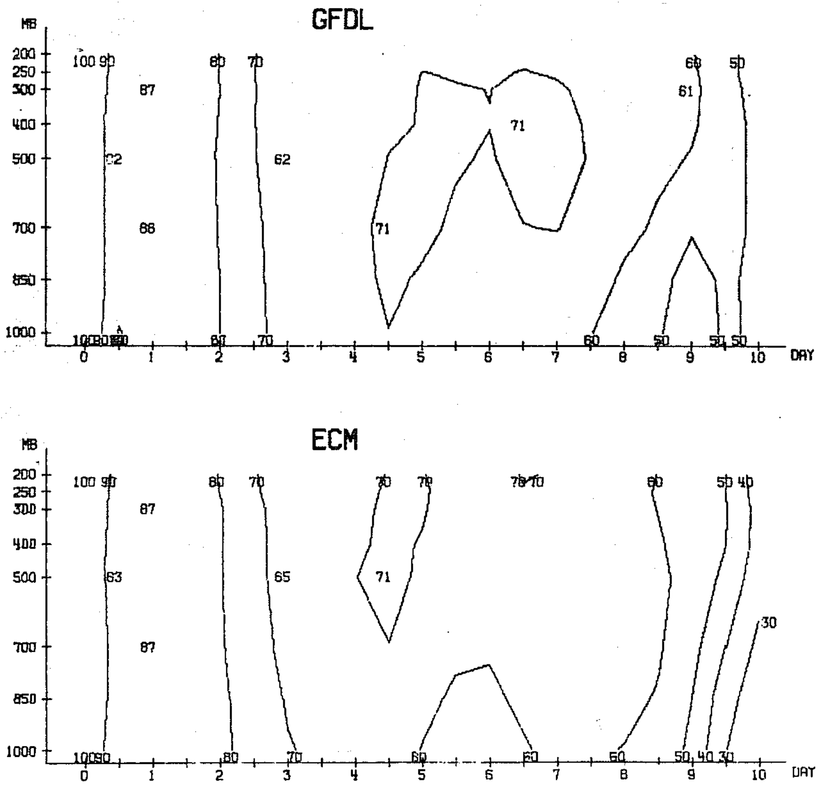


Fig.8.2.8 Correlation of height tendencies
area mean

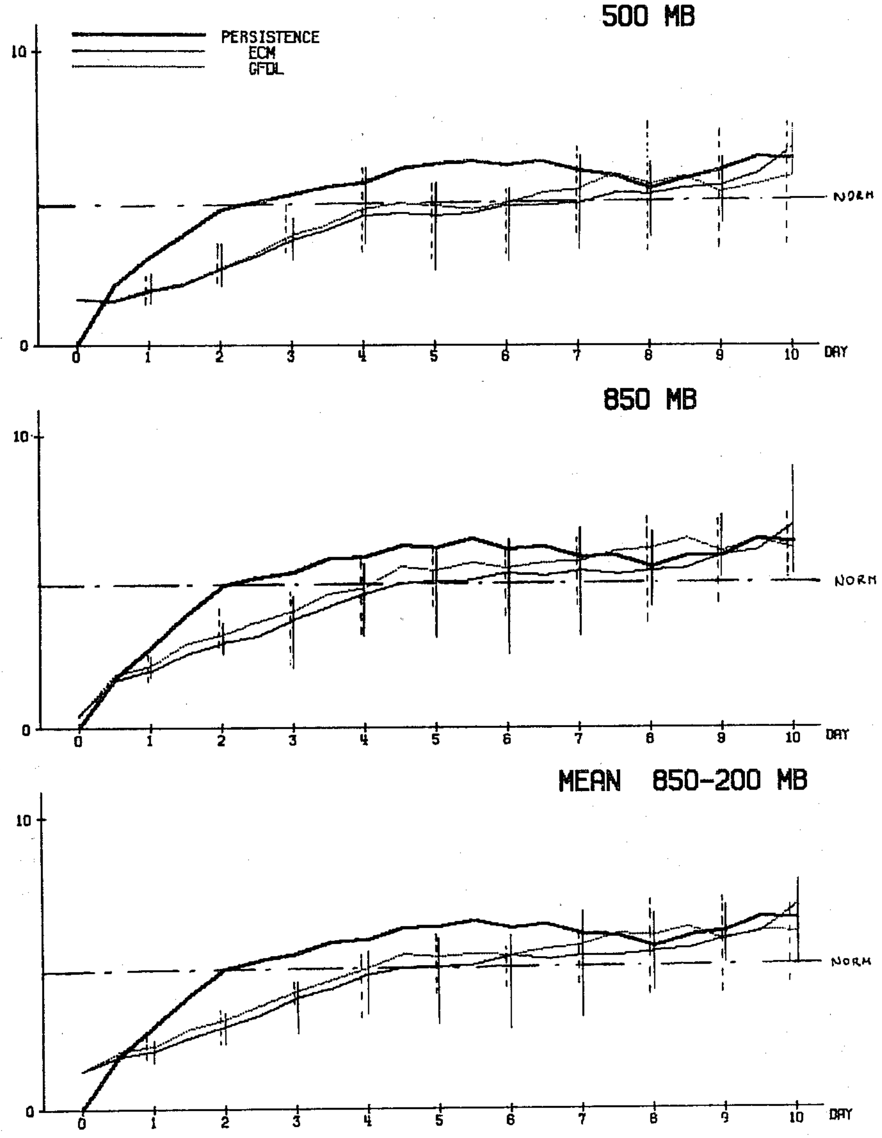


Fig.8.3.1 RMS of temperature (k)
area mean

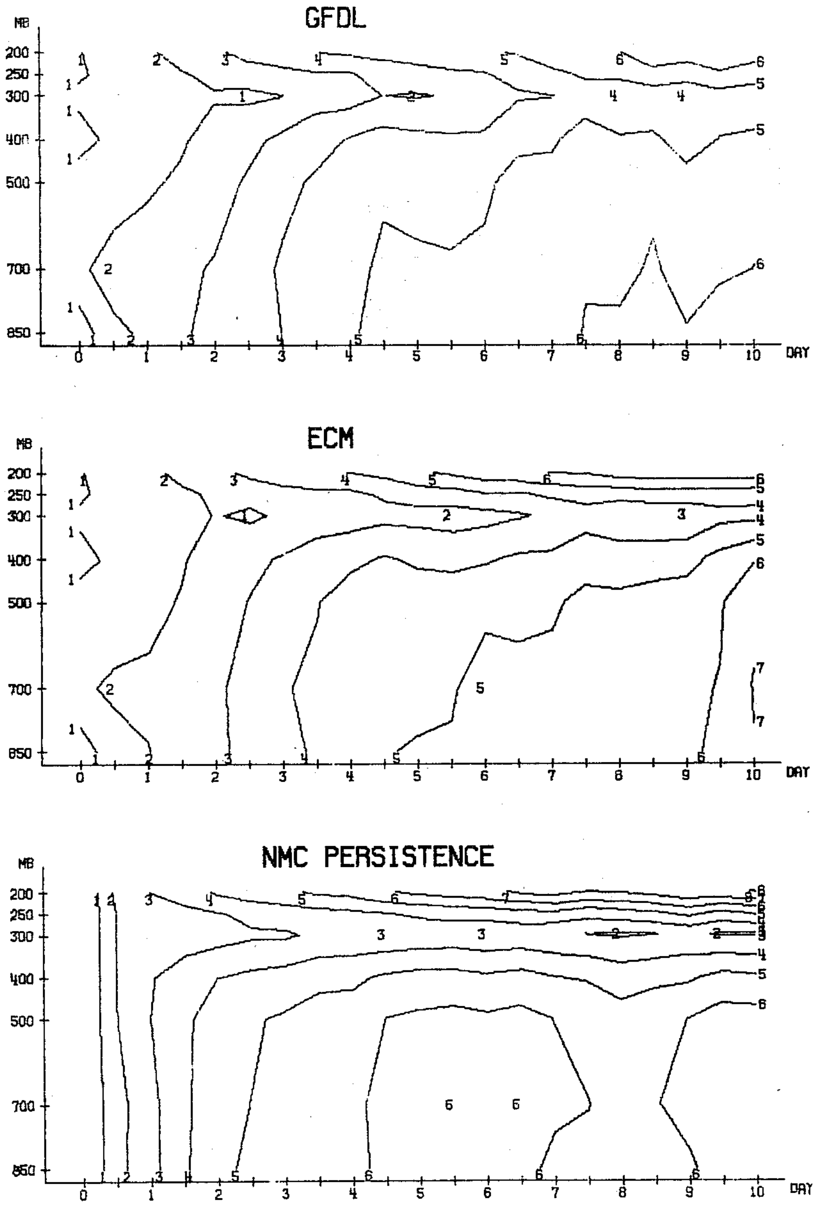


Fig.8.3.2 RMS of temperature (k)
area mean

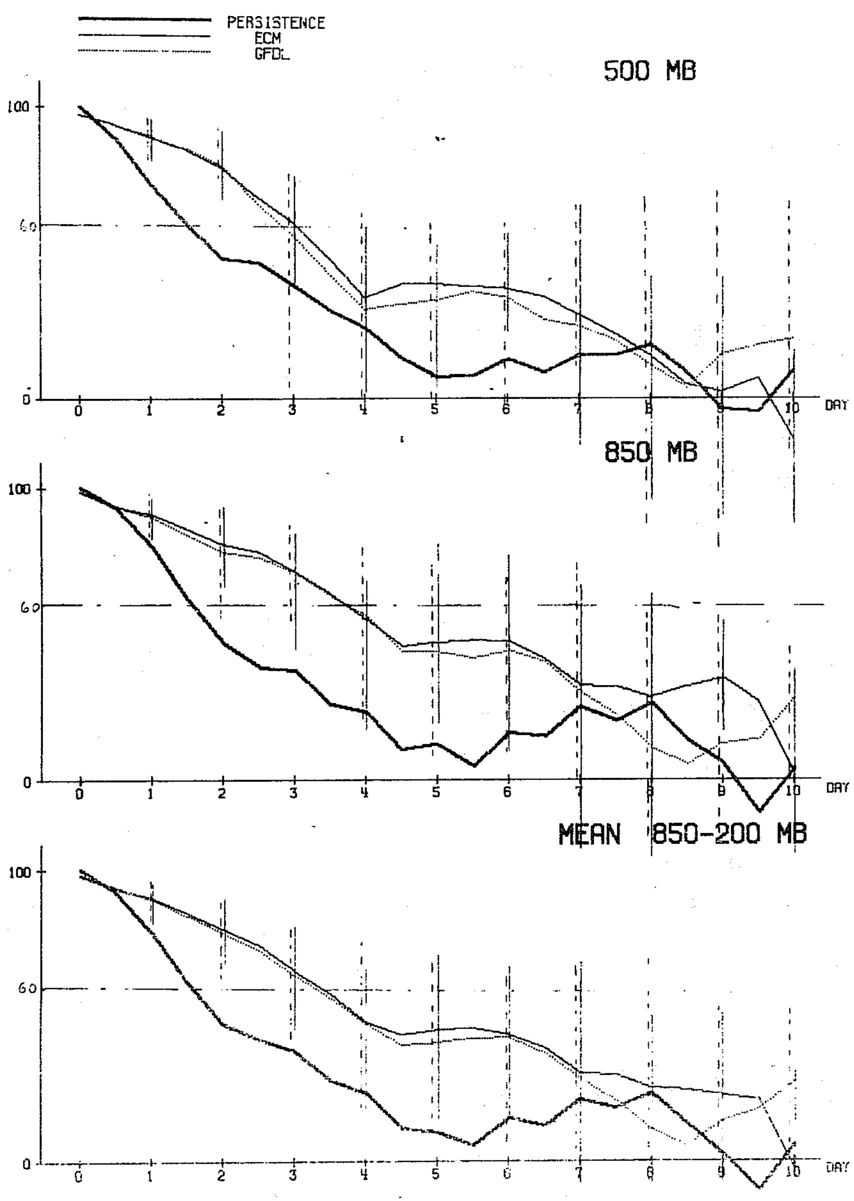


Fig.8.3.3 Correlation of temperature anomalies area mean

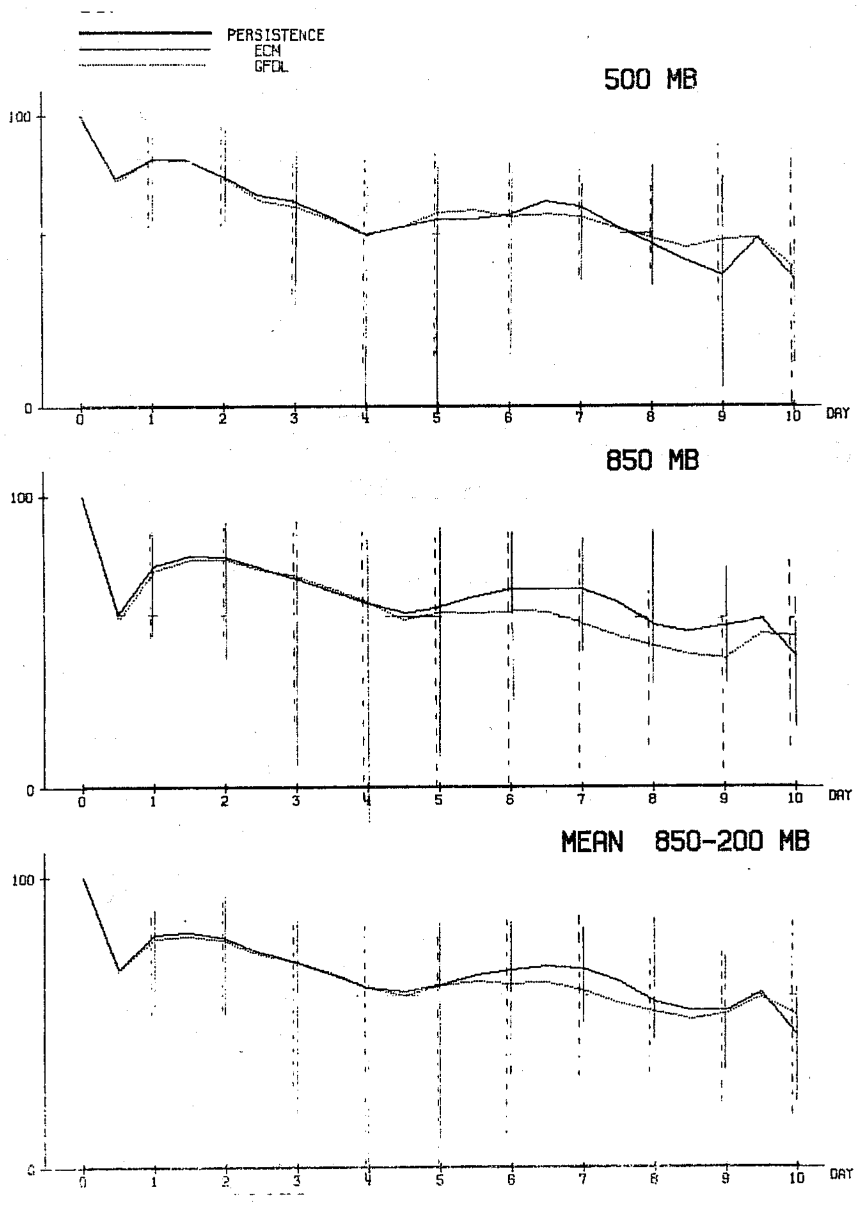


Fig.8.3.4 Correlation of temperature tendencies area mean

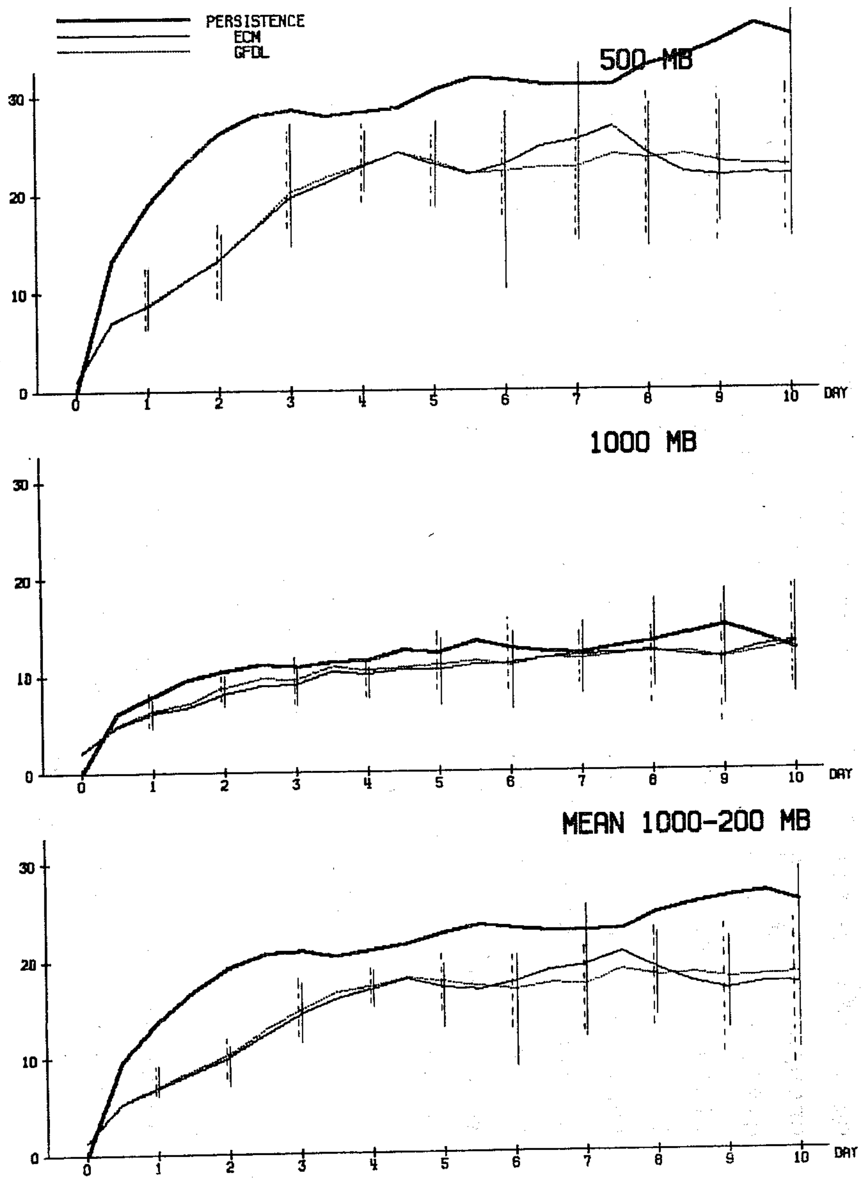


Fig.8.4.1 RMS of wind area mean (m/s)

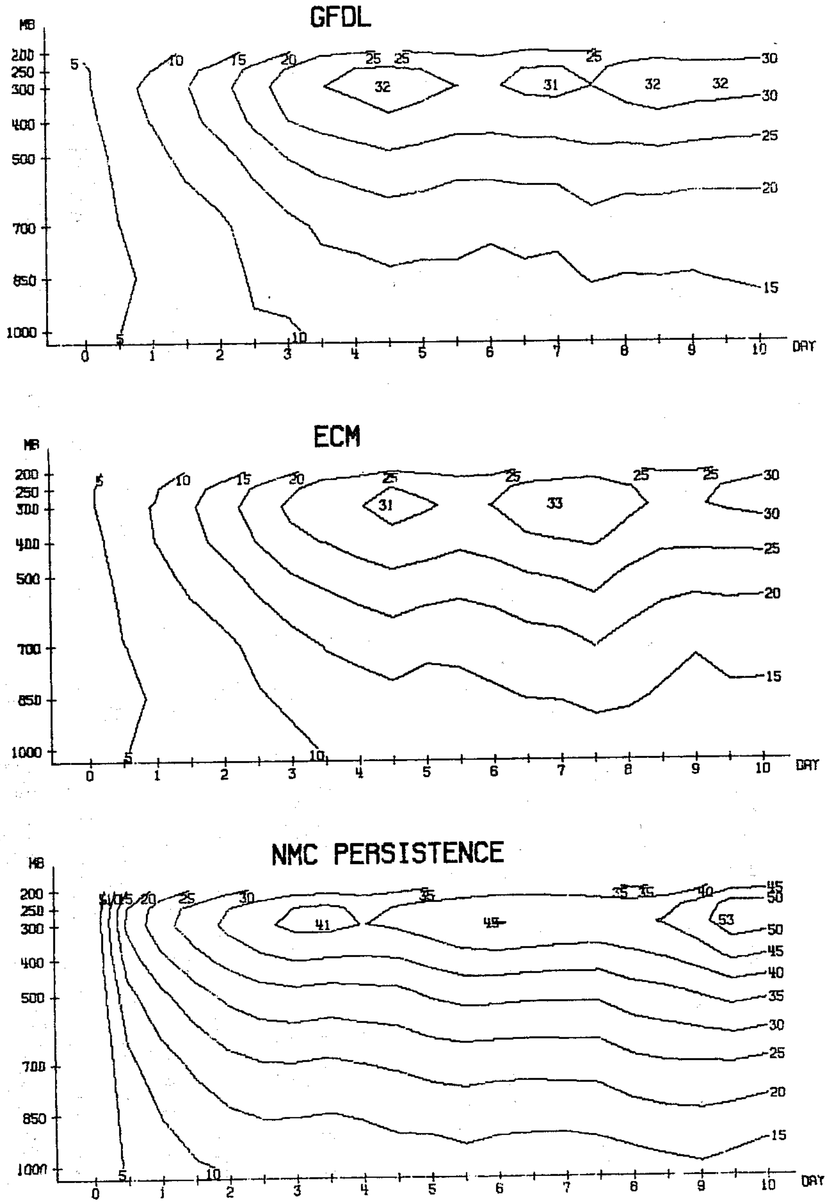


Fig.8.4.2 RMS of wind area mean (m/s)

9. Systematic errors

In this section we consider the systematic errors in the forecasts from several related points of view. In 9.1 we discuss the geographic distribution of the errors. We also discuss the time evolution of the long wave components of the errors. In 9.2 we discuss the evolution of the 500 mb height fields in terms of empirical orthogonal functions.

9.1. Geographical distribution and evolution of spherical harmonic amplitudes

The systematic errors for day N (N=1,2...) are the differences between two ensemble means, the ensemble mean of a set of day N forecasts and the ensemble mean of the verifying observations. For a large ensemble, the ensemble mean of the observations is climatology. We could, therefore, speak of the mean of a large ensemble of day N forecasts as the models day N climatology. Thus for large enough ensembles the systematic errors at day N are the differences between the day N climatology and the observed climatology. The trend of the day N climatology (for N = 1,2,3...), as it departs from the observed climatology therefore tells us something about the erroneous treatment of what should be the stationary component of the flow in the model. With a small ensemble of seven cases, such as we have here, this interpretation is not valid but it may be suggestive.

Fig. 9.1.1 shows the systematic errors (forecast minus observed) at day 2 in the 500 mb and 1000 mb height fields for both models. At both levels there is considerable similarity between the results for the two models. Large negative errors occur at both levels over the area between Iceland and Scotland, with another large and extensive area of negative error in the Eastern North Pacific.

These ensemble averaged error fields may be compared with those shown by Fawcett (1969) for the NMC primitive equation 36 hour forecasts for the months of January 1967 and January 1968. At 500 mb the major errors for the area shown by Fawcett are a large area east of Japan with positive errors between 90 and 120 m, an area near the west coast of America with negative errors between 60 and 90 m and an area in mid and eastern Atlantic with errors in the same range. His calculations are not directly comparable with ours but we believe it is significant that both sets of statistics show errors of the same sign in the eastern Pacific and eastern Atlantic.

Fawcett does not show systematic errors at 1000 mb for a winter month. However, we understand (L. Bengtsson, personal communication) that in the Atlantic there is a great deal of similarity between the distribution of our ensemble averaged 1000 mb error fields and those for the 24 hour forecasts made in the winters of 1967 - 68 and 1968-69 by the Swedish three level filtered model. Both show positive errors in the western Atlantic and negative errors over much of Western Europe. It is probably not fortuitous that these error fields from three quite different models are similar in their spatial structure. Relatively few maps of systematic error distribution have been published. It would be worthwhile to establish the extent to which these errors are common to all models.

Fig. 9.1.2 is similar to Fig. 9.1.1 except that it shows the results for both models at 500 and 1000 mb at day 7. The patterns for the two models are strikingly similar at each of the two levels. At both levels there are large negative errors over Western Europe and the North Western Pacific. Fig. 9.1.3, with the corresponding data for day 10, again shows remarkable similarity between the systematic errors for the two models. We note also that the error fields at day 7 and day 10 are very similar to each other in structure; with the error field at day 10 showing higher intensities. The error field is clearly dominated by errors in the largest scale. Fig. 9.1.4 shows the ensemble means of the 500 mb observations at day 10 and the ensemble means of the 500 mb day 10 forecasts from both models. The differences between these fields were shown in Fig. 9.1.3.. Also shown is the February 500 mb climatology. A comparison of the average of the observations with the climatology shows that our ensemble is at least large enough for the observations to capture the main features of the climatology. These are the major troughs over eastern Asia and eastern North America, the ridges over the Gulf of Alaska and Western Europe and the trough upstream from the Himalayas. The ensemble means of the day 10 forecasts on the other hand do not retain the climatological structure. The flow is too zonal; both the main ridges over Europe and Alaska have disappeared and have been replaced, to some extent, by troughs. The longest waves have disappeared and the dominant wavenumber is much shorter than it ought to be. The ensemble means of the forecasts differ quite a lot in their medium scale features. Nevertheless the major errors, as we have already seen in the error fields, are very similar and are dominated by the errors in the longest waves.

Fig. 9.1.5 shows the data corresponding to Fig. 9.1.4 for 1000 mb. At this level,as well as 500 mb,the ensemble is large enough for the observations to capture the main features of the climatology, namely, the lows in the North Pacific and North Atlantic, the subtropical oceanic anticyclones and the Siberian High. The features are of course more intense in the ensemble mean of the observations than in the climatological fields, just as they were at 500 mb. This is because of the small size of the ensemble. When we examine the ensemble averages of the day 10 forecasts it seems that the overall structure is better preserved than it was at 500 mb. The subtropical highs and the Siberian high have been preserved. The major errors are in the treatment of the two north oceanic lows. The main centres have migrated to the continental margins and there have been marked increases in intensity, particularly in the north Pacific system.

The growth of these errors is gradual over the forecast period and is consistent with the synoptic view that the lows are too deep and too far south as they approach the eastern sides of the oceans.

The error fields are of very large scale and it is possible that part of the error growth is due to a defect in the direct forcing of the longest waves. A study of the time development of the error may shed some light on this question.

The development in time of the error field structure may be seen by considering harmonic dials of the spherical harmonic components of the day N ensemble means. These show, on a polar diagram, the amplitude and phase of the indicated component in the day 1, day 2... day 10 ensemble means for the two models and the observations. Northern hemisphere data for 500 mb height were used, and symmetry about the equator was assumed.

The resulting fields were then expanded in spherical harmonics assuming symmetry about the equator. We show the dials for both models and for the observations.

On Fig. 9.1.6 we have the dials for the modes (1,1), (1,3), (1,5), (1,7); the dials for (2,2), (2,4), (2,6), (2,8) are shown in Fig. 9.1.7 while Fig. 9.1.8 shows the dials for (3,3), (3,5), (3,7), (3,9). These are the most important modes in the ensemble mean fields. A number of features are immediately evident. The wavenumber one components are reasonably well maintained by both models as regards both amplitude and phase. The only significant exception is the (1,3) component whose amplitude roughly doubles in the forecasts but whose phase is well treated. All four wavenumber two components of the ensemble mean fields have serious and progressively larger errors both in amplitude and phase. The growth of the systematic errors in these components seems roughly linear in time, at least for the last six days of the forecast. The similarity in the behaviour of the two models is very striking and suggests that the systematic errors are insensitive to the differences in the physics. The wavenumber three components all show systematic declines in amplitude.

A detailed investigation of the vertical and meridional structure of the systematic errors must await a later report. Nevertheless the smooth evolution of the error at 500 mb suggests that overdevelopment of baroclinic lows may not be the only contributor to the growth of the systematic error. An error in the direct linear forcing of the long waves would be expected to lead to error growth of the kind we see in the models.

Recently Lambert and Merrilees (1978) reported that a large part of the error in a forecast experiment was due to rapid propagation of long waves that should remain stationary. Desmarais and Derome (1978) in a related paper suggest that if the vertical resolution is inadequate the model will fail to recognize that certain modes are stationary forced modes and will treat them as transients. They found that the error was quite serious after four days in wavenumber one, which has the largest phase velocities for external Rossby modes.

Our data set did not include any examples of the kind shown by Lambert and Merrilees where a particular mode remained stationary in the observations and was propagating in the forecasts. But we did find that when the behaviour of a particular mode could be regarded as the sum of a stationary part and a transient part then the more serious errors were in the treatment of the stationary part. Moreover in the individual cases the errors tended to be larger in components with zonal wavenumbers 2 or 3 rather than wavenumber 1.

As an example, consider Fig. 9.1.9 which is an harmonic dial showing the amplitude and phase of the (3,7) component in the 500 mb geopotential for the observations and ECM forecast for ten days from February 22. Also shown is the amplitude and phase of the error in the ECM forecast. This example is the only case where we clearly see behaviour that appears similar to the findings of Lambert and Merrilees.

The observation and forecast curves show a behaviour that could be interpreted as oscillation about a mean state. The period of oscillation for the free external non-divergent Rossby mode whose structure is $P_n^M(\cos\theta)e^{im\lambda}$ is given by $\frac{n(n+1)}{2M}$ days. For the (3,7) mode this is $9\frac{1}{3}$ days which corresponds well with the period of the mode in the forecast. However, an inspection of the error shows that between days 3 and 10 it is predominantly stationary in character. On the other hand if one looks only at the error in the first four days one would say that it is mainly transient because of the initial growth in error amplitude.

Significance of the systematic error

Having shown the structure and evolution of the systematic errors we may ask how important these errors are. Fig. 9.1.10 shows the distribution of the root mean square (RMS) height errors at 500 mb and 1000 mb for day 10 for both models. The largest errors in an RMS sense clearly occur in the same areas as the largest systematic errors. There is no area where the systematic error is small and the RMS error is large. We may suppose then that the systematic error is an important component of the RMS error.

To discuss this question further we consider the standard deviation of vertically averaged height errors. We have already seen the ensemble average of the standard deviation of errors for this field in Fig. 7.1.1. Fig. 9.1.11 shows the variation with time of the standard deviation of the vertically averaged difference between the models' ensemble mean fields and the corresponding observed ensemble mean fields. We also show the curve for persistence of the day 0 ensemble which would be zero for a large ensemble. We write the ensemble mean standard deviation of error e(t) as

$$e(t) = (e_s^2(t) + e_n^2(t))^{\frac{1}{2}}$$

where $e_s(t)$ is due to the systematic error and $e_n(t)$ is due to the non-systematic error.

Let us suppose that we could improve our forecasts to the point where the systematic error is eliminated. Standard deviation of the error is then $e_n(t)$. What is the increase in the range of useful predictability assuming the systematic error and the non-systematic error are independent?

Let us model e(t) by a linear expression in time

$$e(t) = \frac{N}{T_1} t$$

where N is the "norm" standard deviation and T_1 is the time taken for e to reach the level N. Similarly we suppose that $e_n(t)$, the error of a revised model, which has no systematic error may be written

$$e_n(t) = \frac{N}{T_2} t$$

We assume that T_1, T_2 are the limits of predictability in the two situations

Then

$$\frac{T_2}{T_1} = \frac{e(t)}{e_n(t)} = \frac{e(t)}{\sqrt{e^2-e_s{}^2}} = \frac{1}{\sqrt{1-\left(\frac{e_s}{e}\right)^2}}$$

A comparison of Figs. 7.1.1 and 9.1.11 shows that for most times $\frac{e_s}{e}$ takes a value between .5 and .6. In Table 9.1 we present values of $(1-x^2)^{-\frac{1}{2}}$ for $0 \leqslant x \leqslant 1$

TABLE 9.1

x :	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$(1-x^2)^{-\frac{1}{2}}$	1.	1.01	1.02	1.05	1.09	1.15	1.25	1.43	1.67	2.27	∞

From Table 9.1 it follows that with these assumptions the elimination of the systematic errors would lead to a 20% gain in useful predictability.

We may ask what are the prospects for eliminating the systematic errors? On the one hand we have the question of the correct treatment of the mechanical interaction between large mountain masses and the atmosphere and the treatment of the thermal effect of land sea contrast on the atmosphere. Both of these effects are of large horizontal scale and have often been successfully studied by linear analysis (Charney and Eliassen 1949, Smagorinsky 1953). The behaviour seen in the harmonic dials of the ensemble mean fields looks rather linear and therefore amenable to analysis.

This encourages us to undertake linear studies of the phenomenon from the point of view of thermal or mechanical forcing of the stationary flow.

On the other hand the systematic errors are largest on the eastern sides of the Atlantic and Pacific Oceans. These are areas where there is a marked tendency for lows to weaken and fill, (Lau 1978). We have noticed a marked tendency for surface features in the models to overdevelop. If lows in the forecasts intensify rather than fill on the eastern sides of the oceans, then we would expect to see a systematic negative error in these areas. The problem of the overdevelopment of surface features is under active study and there is reason to expect that a satisfactory solution may be found.

The two points just discussed concern the non axisymmetric part of the flow. There is also an error in the axisymmetric part of the flow, i.e. the incorrect acceleration of the zonal mean flow at all levels. For geostrophic balance this requires an enhanced pole-equator pressure difference which is reflected in the fact that the systematic errors are largely negative north of 20 N.

The systematic error in the zonal flow is discussed further in Section 11.

To summarize, certain aspects of the systematic errors appear to be amenable to analysis. If the errors could be removed altogether then we would expect a gain of 20% in useful predictability. Since the systematic errors are in the largest scales of motion, which it is our goal to forecast, the study of these errors must have high priority.

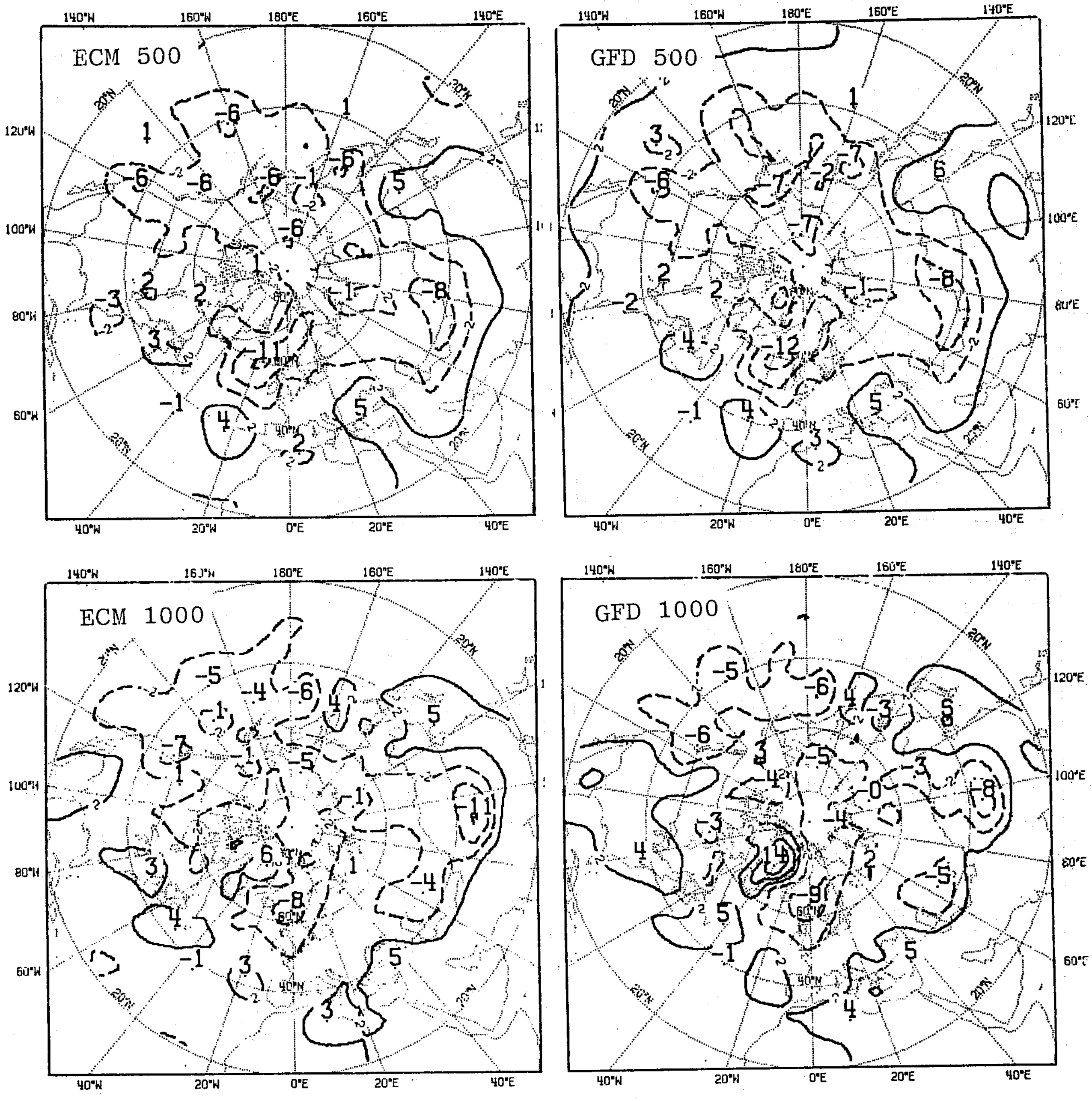


Fig. 9.1.1. Systematic error field (ensemble average of forecast minus observed) at 500 mb (top) and 1000 mb (bottom) for ECM model (left) and GFD model (right) for day 2 hour 0. The interval is 4 dam starting at ± 2 dam. Contours for positive values are solid, those for negative values are dashed.

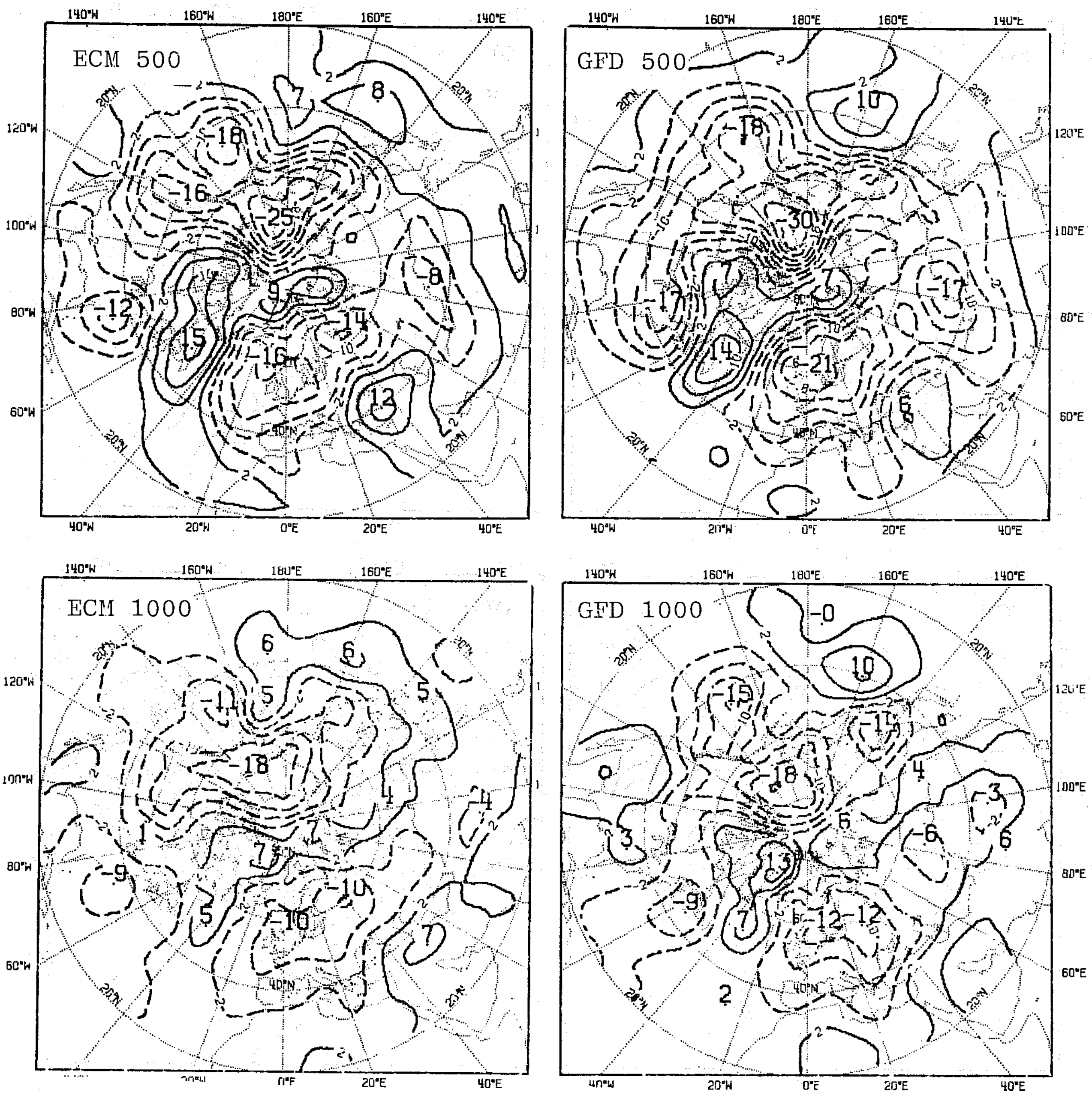


Fig. 9.1.2 As Fig. 9.1.2 for day 7 hour 0.

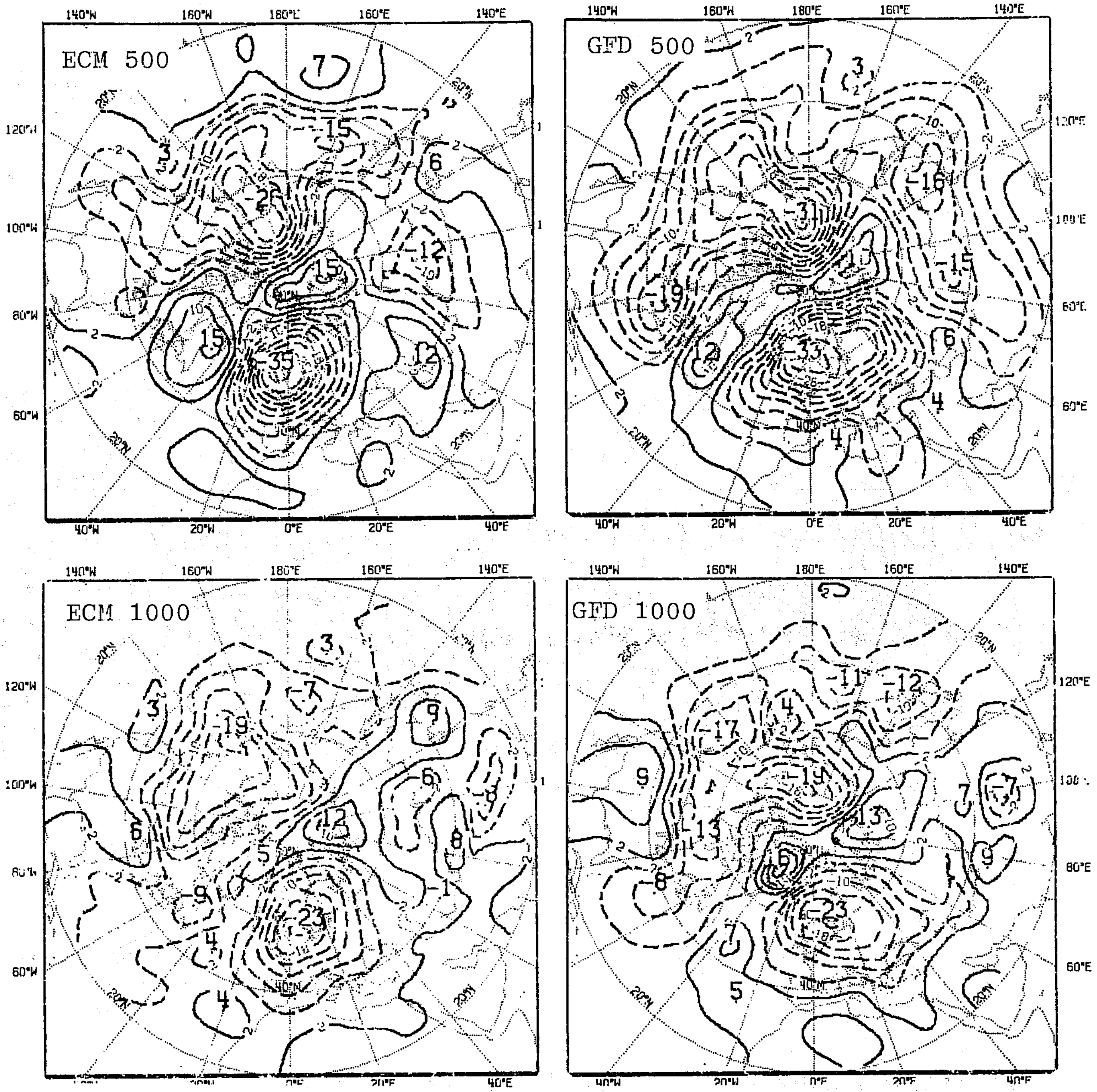


Fig. 9.1.3 As Fig. 9.1.1 for day 10 hour 0.

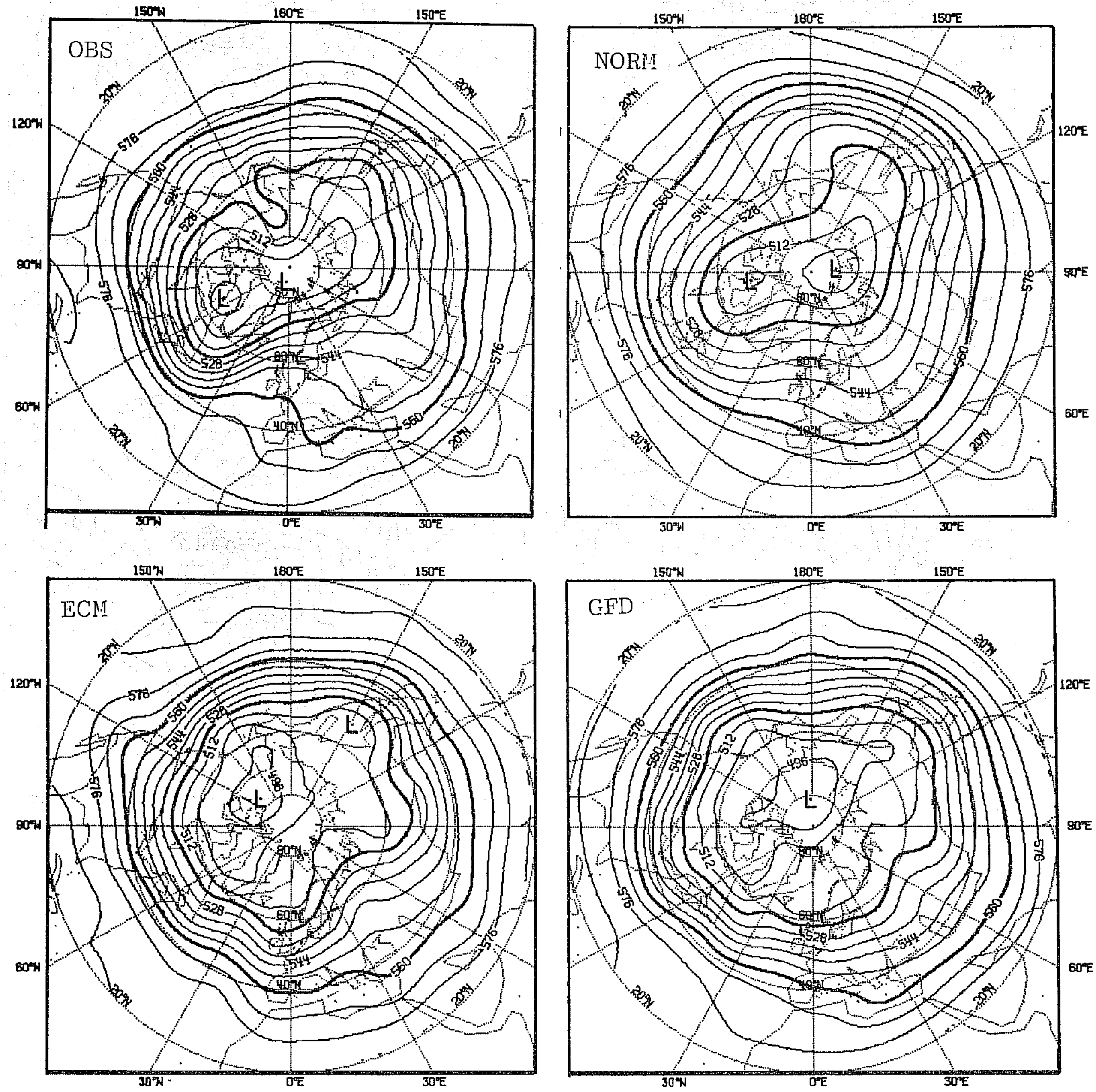


Fig. 9.1.4 Ensemble average of the 500 mb height field at day 10 hour 0 for observations (top left) the ECM model (bottom left) and the GFD model (bottom right). Also shown (top right) is the February normal for this field. The contour interval is 8 dam.

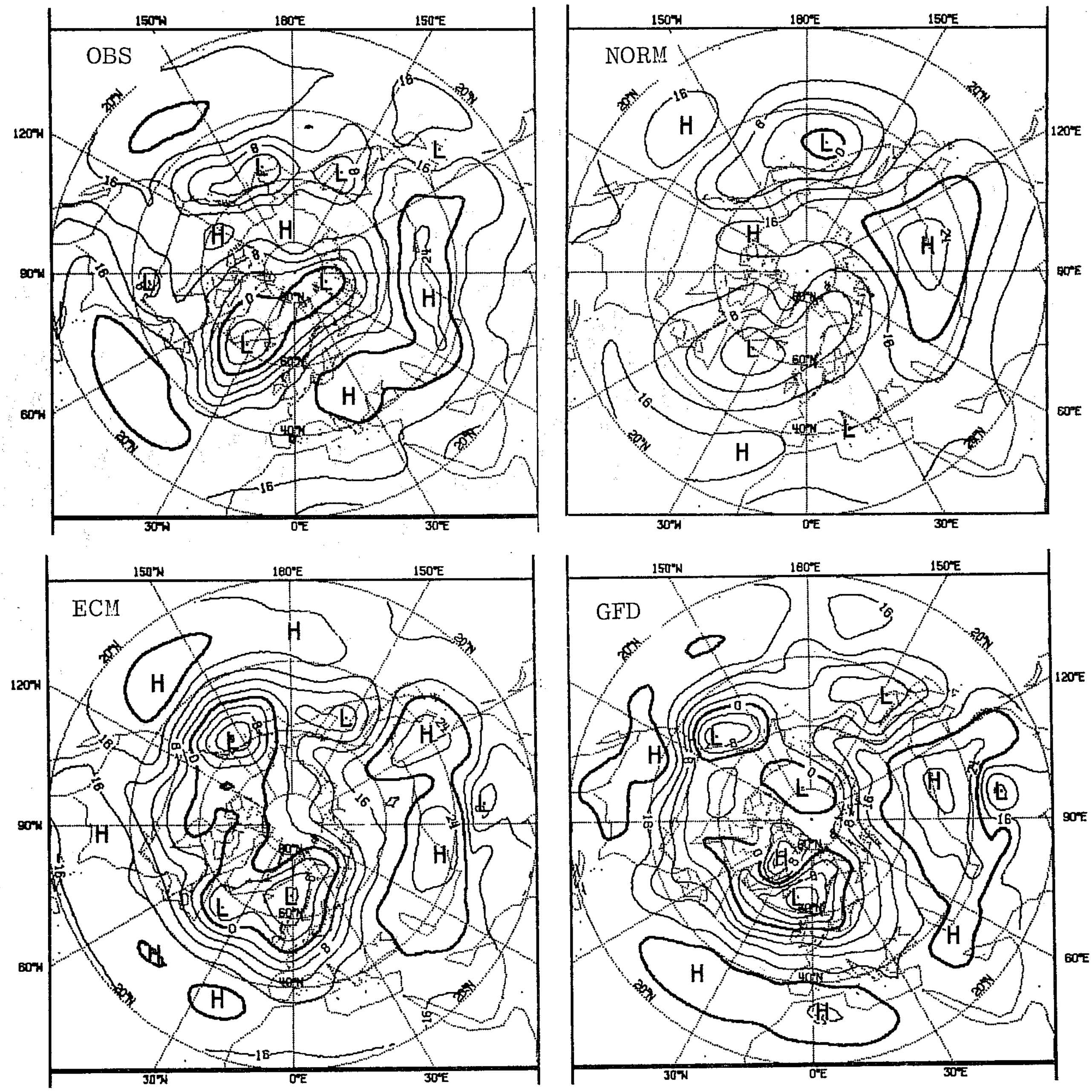


Fig. 9.1.5 As Fig. 9.1.4 for 1000 mb. The interval is 4 dam.

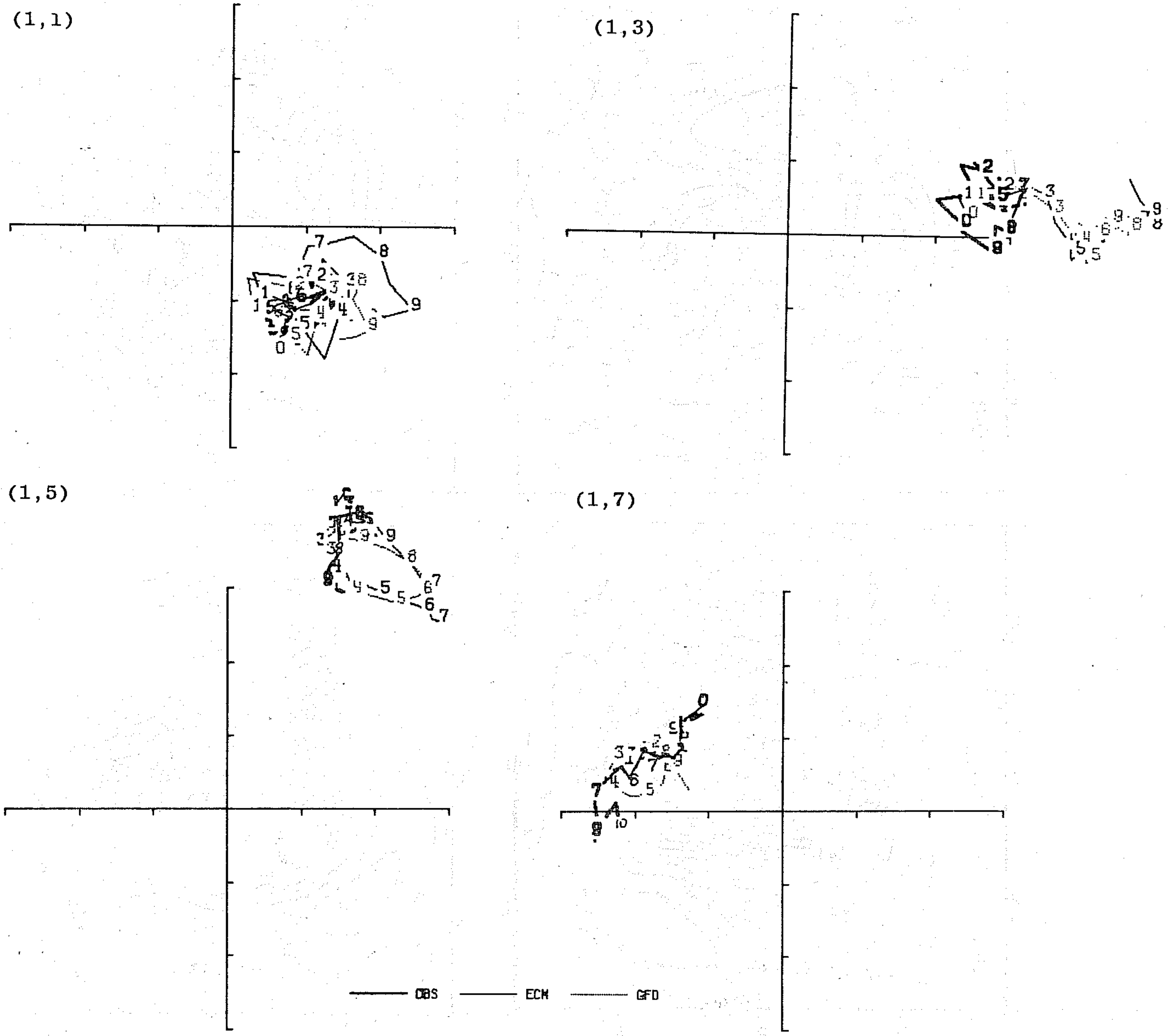


Fig. 9.1.6 Wavenumber 1 harmonic dials for ensemble average northern hemisphere 500 mb geopotential for observation (heavy line) ECM model (thin line) and GFD model (dotted line). The components shown are (1,1),(1,3),(1,5),(1,7). The figures show the valid time (e.g. 9 means day 9 hour 0) The interval between the ticks is 10 m.

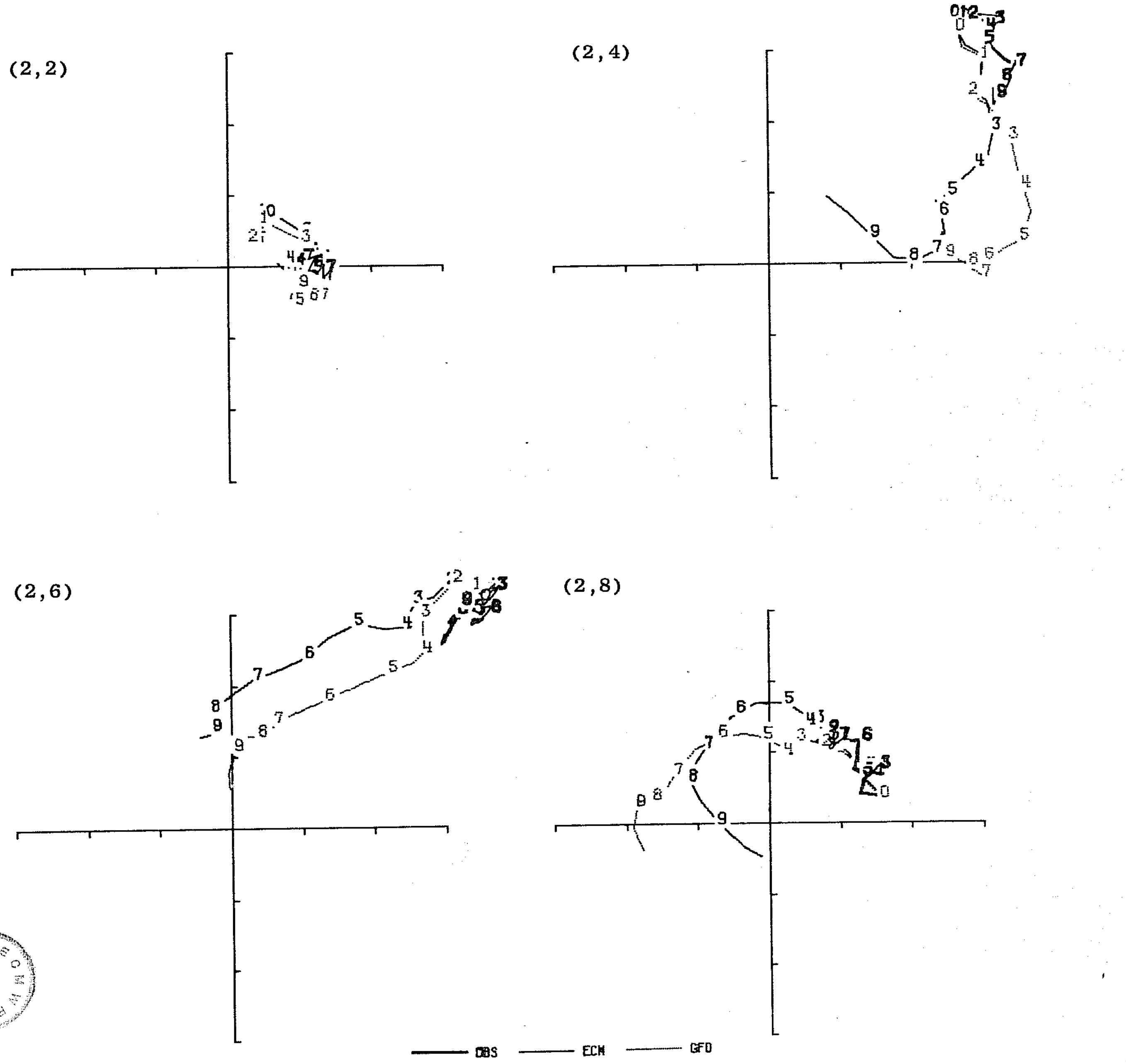


Fig. 9.1.7 As Fig. 9.1.6 for the wavenumber 2 components (2,2),(2,4),(2,6),(2,8).

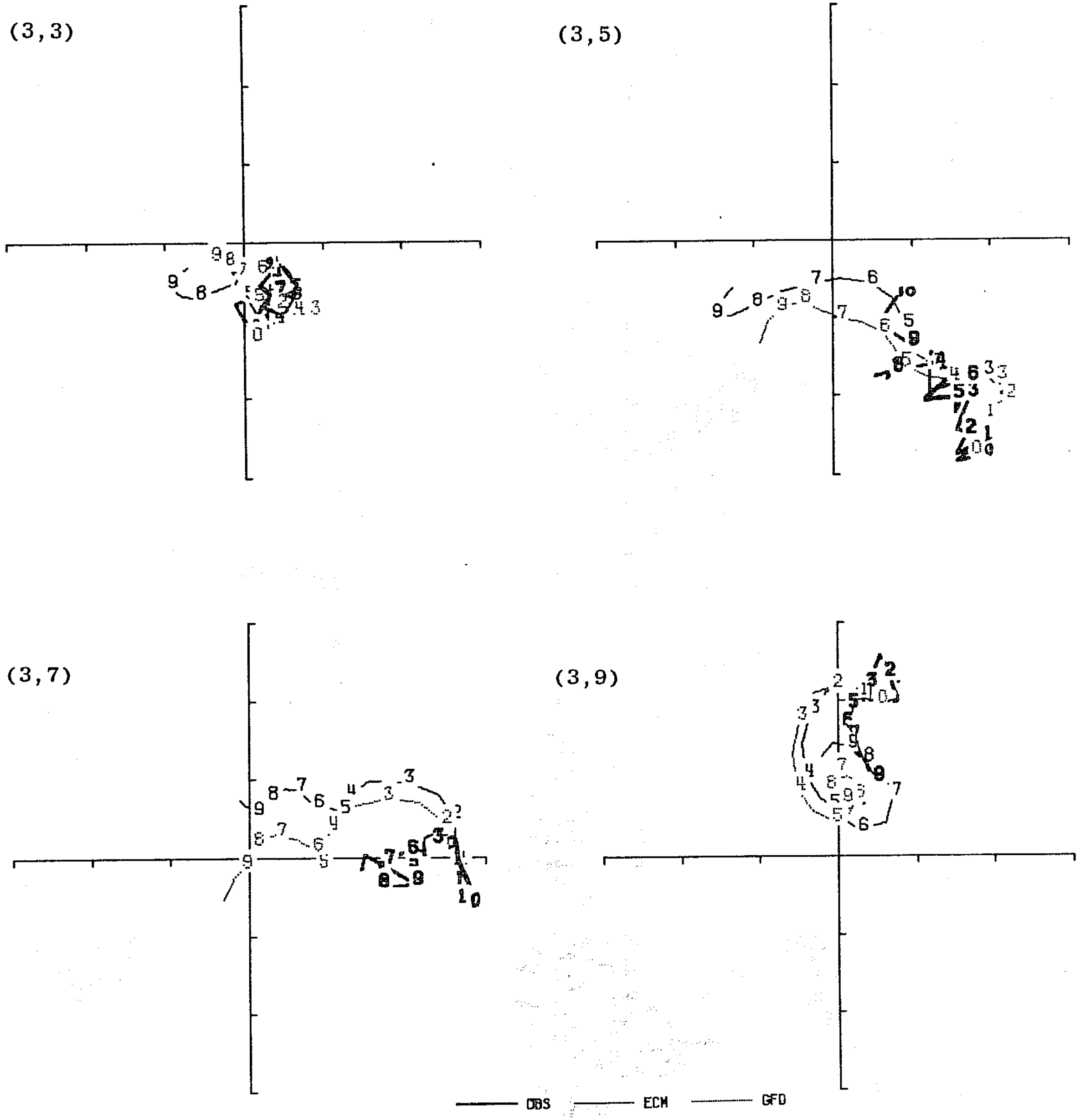


Fig. 9.1.8 As Fig. 9.1.6 for the wavenumber 3 components (3,3),(3,5),(3,7),(3,9).

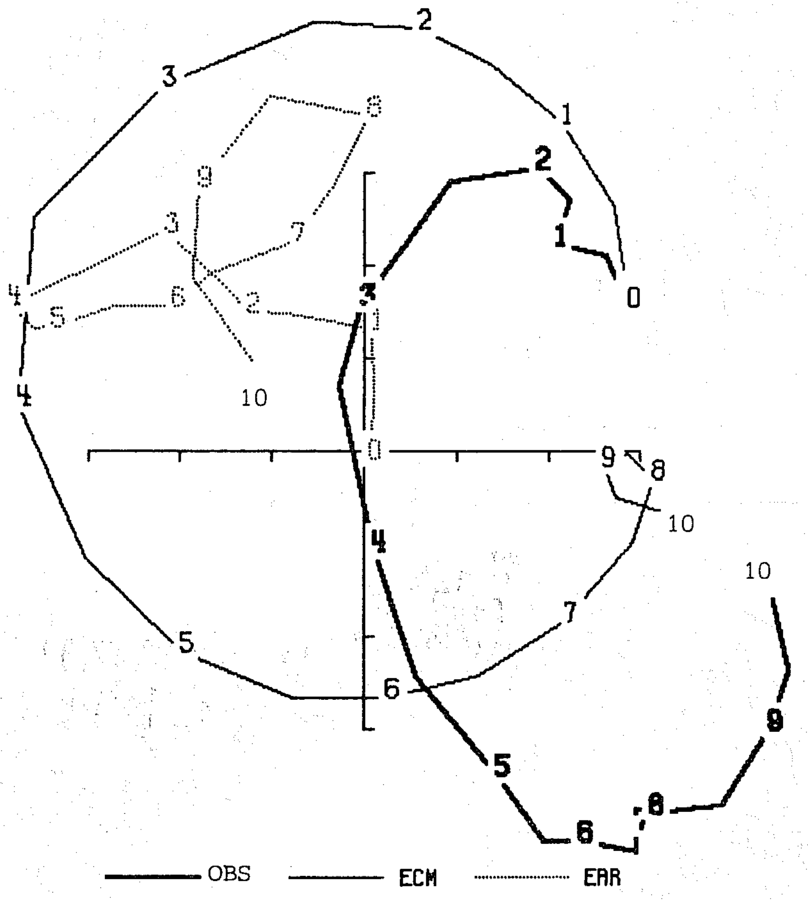


Fig.9.1.9 Harmonic dials for (3,7) component of 500 mb height field in observations (———), ECM forecast (......) and error field (------) for forecast from 22 Feb 1976. The numbers show the time in days from the initial time.

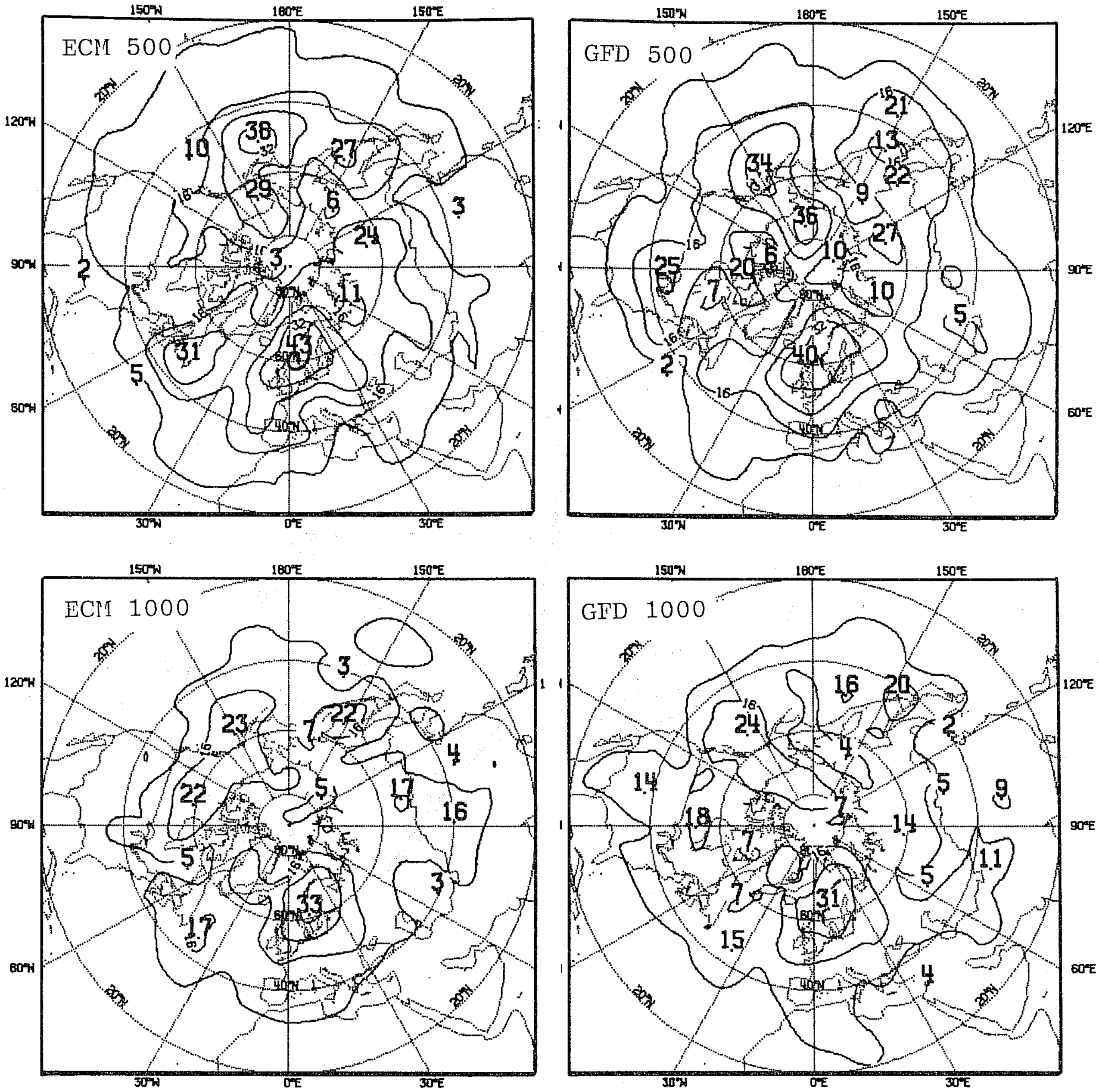


Fig. 9.1.10 Root mean square errors for the ensembles of forecasts at day 10 for 500 mb (top) and 1000 mb (bottom), and for the ECM model (left) and the GFD model (right). The interval is 8 dam.

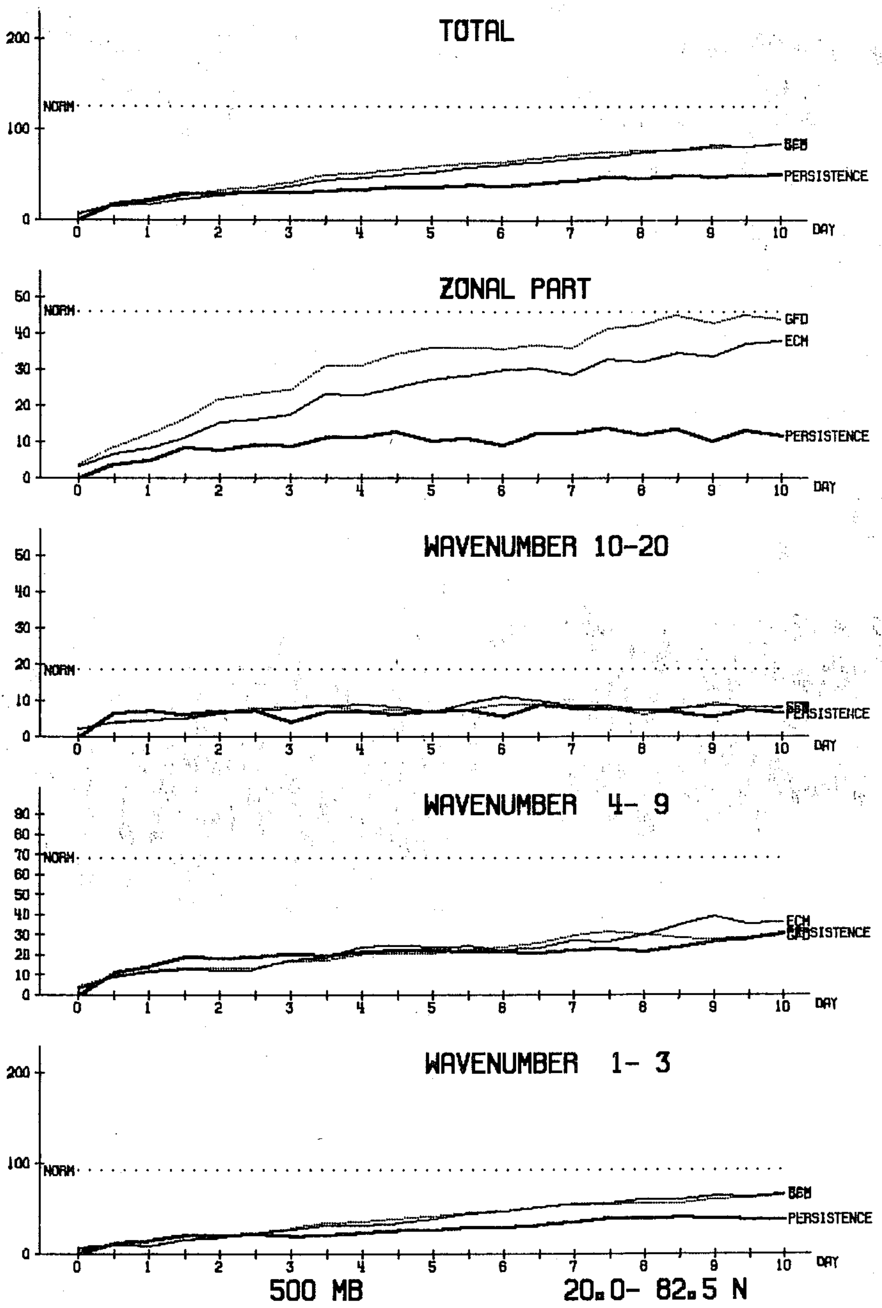


Fig. 9.1.11 Time evolution of the standard deviation of height error at 500 mb in the ensemble mean of the forecasts, as compared with the ensemble mean of the observations, for the total field (top) and for the indicated wavenumber groups. The ECM model results are shown with the thin line, those for the GFD model with the dotted line. The corresponding curves for persistence (thick line) would be practically zero for a sufficiently large ensemble. Also shown (with label NORM) is the level of the climatic standard deviation.

9.2. Evolution of forecast errors in terms of empirical orthogonal functions

In this section we study the errors in the 500 mb height field forecasts by expanding them into empirical orthogonal function (eof) series. The eof's are based on 20 years 500 mb height analysis material (calculated in the University of Helsinki, Karhila and Rinne, 1977); the first five modes $f_n(\lambda,\phi)$ and the 20 year time mean field $zm(\lambda,\phi)$ are shown in Fig. 9.2.1. Expanding the 500 mb height field in these functions gives:

$$z(\lambda,\phi,t) = zm(\lambda,\phi) + \sum_{n=1}^{175} c_n(t).f_n(\lambda,\phi) + \varepsilon(\lambda,\phi,t)$$

where ε is a small residual.

The forecast error for each mode n is the deviation between observed and forecast coefficients $c_n(t)$.

Fig. 9.2.2 shows the history of the first five coefficients in a typical winter case (15 February 1976). The thick line gives the observed evolution in NMC analyses, the thin line is ECM model and dotted line GFD model forecast. Up to day 3 or 4 the forecasts stay well together and then deviate away from the observed values. For the modes 1 and 2 the forecasts also deviate from each other with increasing forecast time, while the higher modes tend to have rather similar evolution. This tendency was more or less the same in all 7 winter cases and suggests that the two gravest eof modes are sensitive to the parameterization scheme. Some tests with the same physics but different integration schemes and resolutions showed that the lowest modes were not sensitive to these aspects of the model.

Fig. 9.2.3 shows the ensemble mean coefficients for 7 February cases. The largest systematic forecast errors fall on modes 1, 3 and 10, all of them being produced more negative than observed. The thick dashed line shows here the ensemble mean in analyses from Deutscher Wetterdienst (DWD). The two analyses differ most in the two lowest modes. In the initial state of the forecast the deviation to DWD analysis is consistently smaller than to NMC analysis in spite of the fact that the initial state is based on NMC analysis followed by interpolation to σ-levels, normal mode initialization and, for the diagnostic calculations, interpolation back to pressure levels. In the ECM model this deviation remains the same while the GFD model systematically deviates further away from the observed values. The too large negative coefficient for c_1 will tend to make the 500 mb heights lower (colder) over the eastern continents where the pattern of f_1 (Fig. 9.2.1) has its maxima. This can be seen in the systematic error maps for days 7 and 10 in Figs. 9.1.2 and 9.1.3, especially for the GFD model.

Unfortunately, the patterns of the modes 1, 3 and 10 do not provide us with the reasons for the relative failure in forecasting such structures in the atmosphere. However, we can estimate what effect they have. If we have a perfect forecast for c_1 the RMS error for 500 mb north of 20 $^{\text{O}}$N decreases about 5% for the ECM model day 7 forecast. If we use persistence for the slowly varying c_1 the RMS error for day 7 decreases about 4% for the GFD model, averaged over the 7 cases. These estimates are based on NMC analyses and would be smaller if DWD analyses were used.

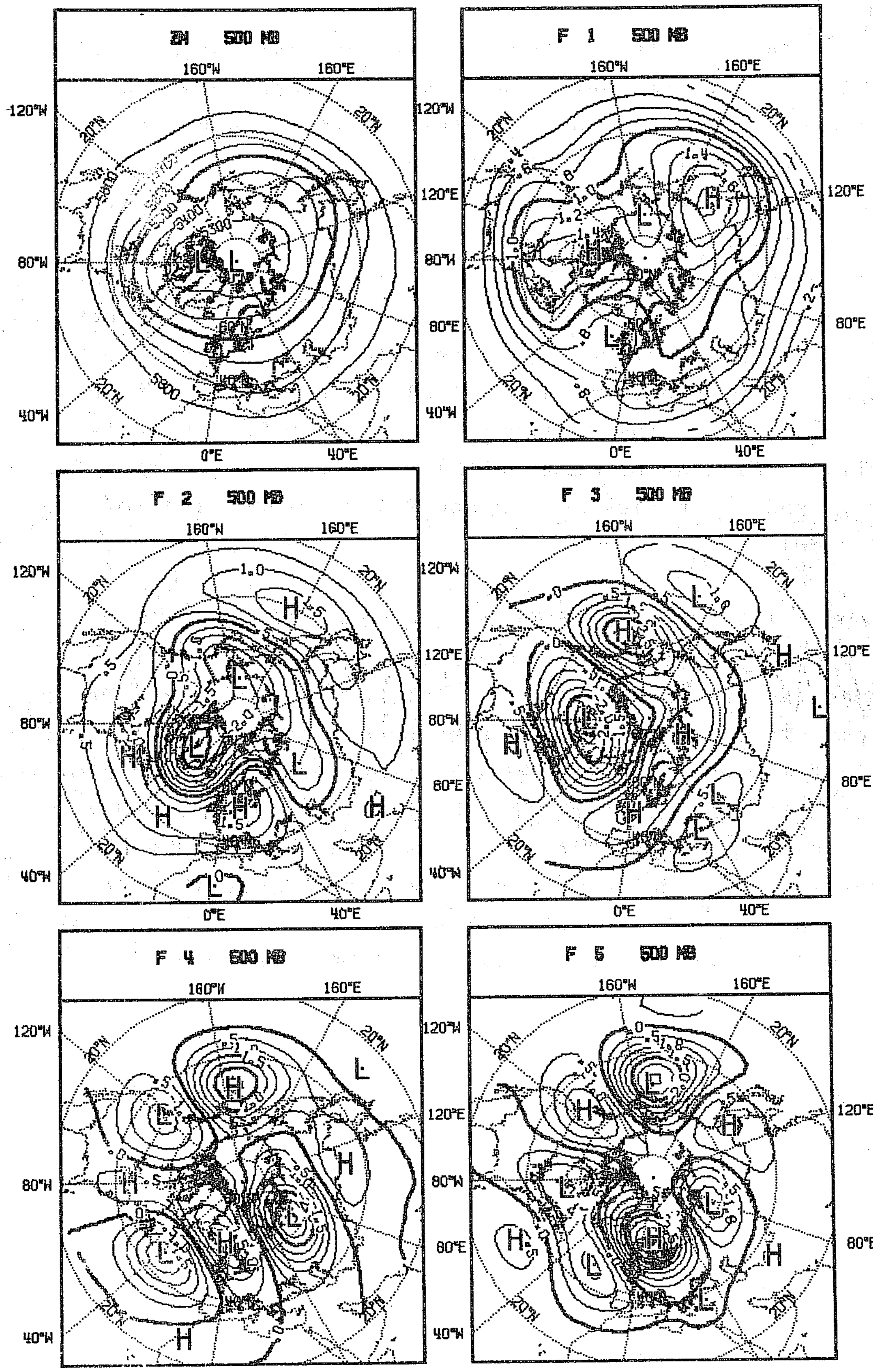


Fig.9.2.1 The time-mean field zm and first five empirical orthogonal functions for 500 mb geopotential height field. (After Karhila and Rinne, 1977).

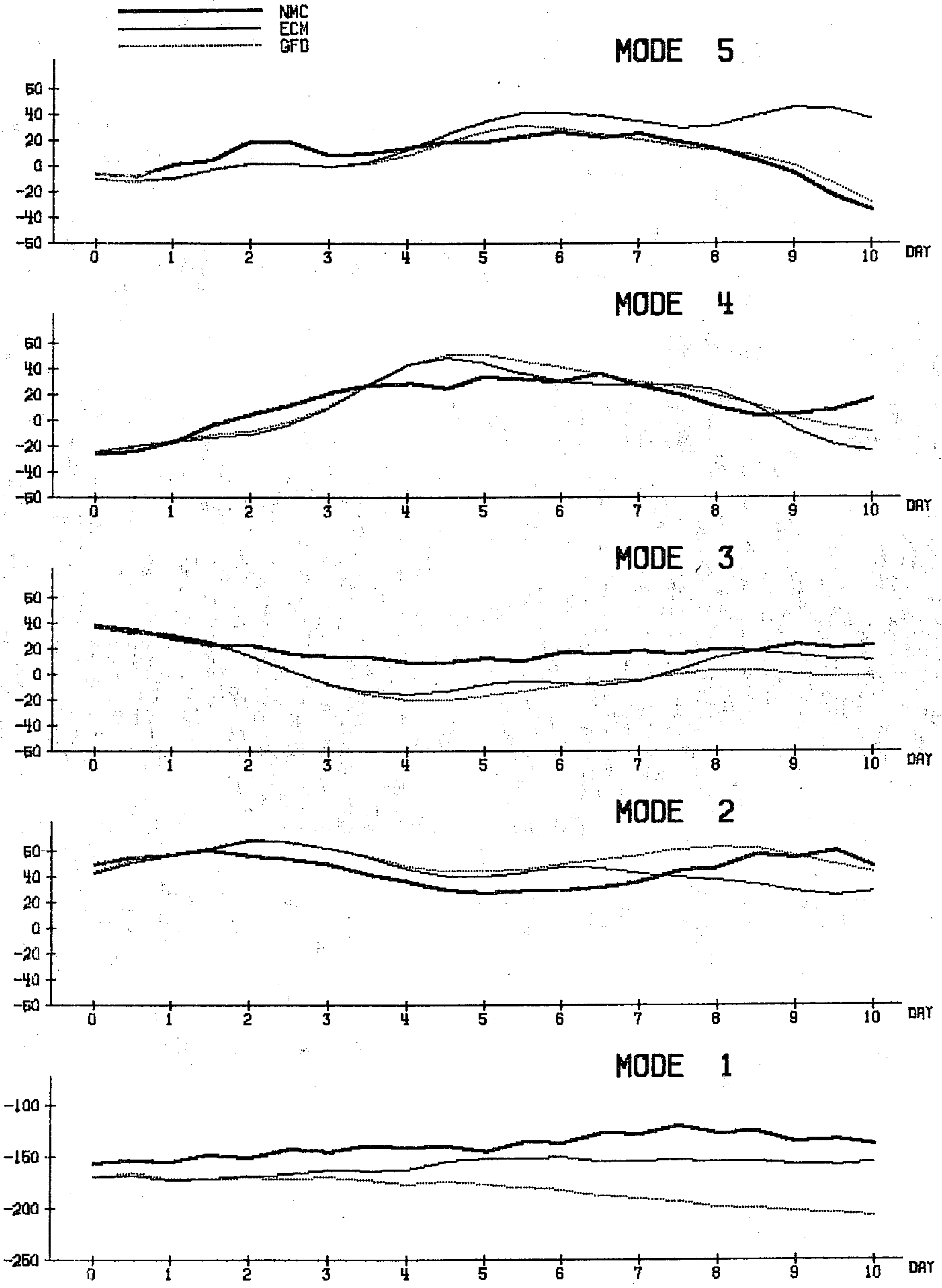


Fig.9.2.2 The evolution of modes 1 - 5 in a typical winter case starting from 15 February 1976.
Unit is m

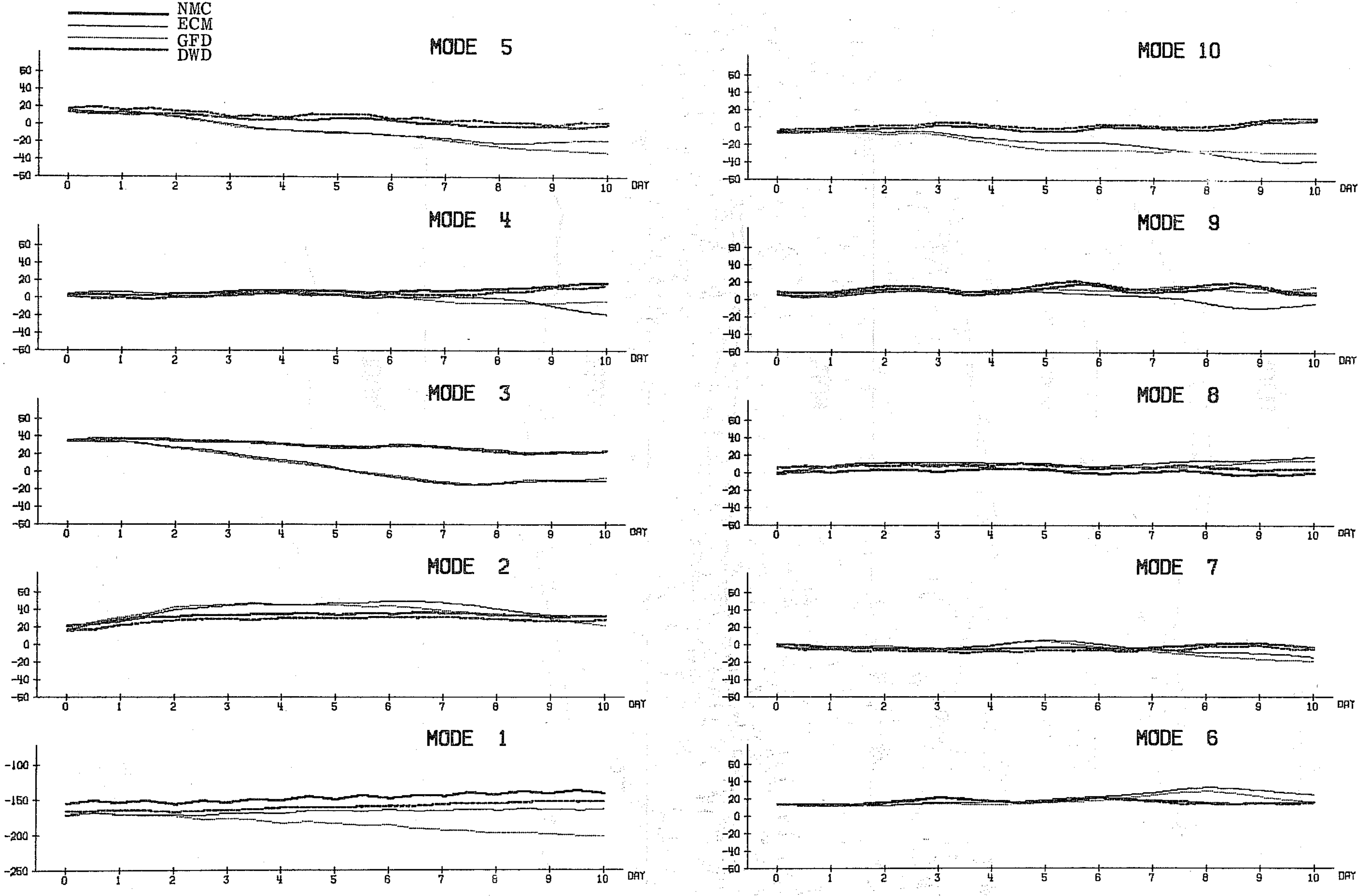


Fig.9.2.3 The ensemble mean of modes 1 - 10 in 7 February cases

10. Differences in the forcing due to the parameterization schemes

In this section the time changes of the flow due to physical processes are discussed. The discussion is based on the tendencies due to physical processes computed in the model during the time integration.

We show only time averaged zonal and global mean tendencies for the forecasts from 22 February 1976. The results are qualitatively confirmed by the other experiments performed so far, from the point of view of global budgets at least. There are some slight inconsistencies that the reader may notice between figures given at different places in this section. They arise because some elements of our diagnostic package have different time frequencies (every time step, every second time step...).

10.1 Temperature

During the 10-day forecast period the mean atmospheric temperature was kept nearly constant in both runs, as the contributions of the different physical processes to the temperature changes balanced reasonably well (Table 10.1).

TABLE 10.1

Global budget for sensible heat

Mean tendencies of temperature due to radiation, surface fluxes, latent heat release and energy conversion (potential energy to kinetic energy) - units are Watts/m^2.

	ECM-run	GFD-run
radiation	-96.8	-109.5
latent heat release	82.5	105.3
surface flux	12.9	5.7
energy conversion	-2.6	-3.6
net	-4.0	-2.1

If exact these figures would lead to a cooling of 0.34 °C in the ECM run and of 0.18 °C in the GFD run. In reality the atmosphere is cooled by 0.47 °C in the ECM-run and by 0.06 °C in the GFD-run. This gives us a measure of the uncertainties and of the global net temperature tendencies. (Fig.10.1.1) shows heating in the boundary layer and in the lower troposphere (ECM-run only) and cooling in the higher troposphere and stratosphere (ECM-run only).

The most striking differences between the two runs are the vertical distribution of heating and cooling and the magnitude of the heating/cooling rate. The ECM-physics gives heating throughout the lower troposphere and cooling at higher levels, whereas in the GFD-run the atmosphere is only heated in two shallow regions in the boundary-layer, where the heating is considerably more intense than in the ECM model, and at the top level. The extreme values of heating are systematically smaller in the ECM-run except for the large cooling rate at the top level. These differences are mainly due to radiative heating and latent heat release, which are the major sources for sensible heat (Table 10.1).

The heating by latent heat release (Fig. 10.1.2) is larger in the GFD-run except for the upper troposphere, where large heating occurs in the ECM-run due to penetrative moist convection (Kuo-scheme) but not in the GFD-run. This is best seen from the latitude height distribution of convective heating (Fig.10.1.3), which in the ECM-run for tropical regions shows a maximum in the upper troposphere (deep convection) and a second weaker maximum at the lower troposphere (shallow convection).

The large-scale condensational heating is largest in the lowest layer ($h = 30$ m) in the GFD-run and is largest in the third layer (900 mb) in the ECM-run (Fig. 10.1.2 and Fig. 10.1.3). The large cooling rate at the lowest level in the ECM-run is due to the evaporation of rain. The vertical distribution of large-scale condensational heating seems very unrealistic in both runs, especially in the GFD-run. The global mean heating by latent heat release is smaller in the ECM-run by about 20% (Table 10.1) as is the total rainfall during the 10-day forecast period (Table 10.5). The precipitation in the ECM-run (28.5 mm/10 days) is closer to the climatological value 2.75 mm/day, than the GFD run with 36.4 mm/10 days.

Fig. 10.1.4 shows the vertical distribution of radiative cooling and also the heating by all other physical processes (latent heat release + surface heat fluxes + energy conversion). The radiative cooling rates of both runs are similar in the upper half of the troposphere. They differ in the lower half of the atmosphere and at the 2 top levels (Fig. 10.1.4 and Fig. 10.1.5). Comparison of the radiative heating rates computed in the model with those computed by Dopplick (1972) using climatological data shows considerable differences, especially in the lower troposphere (see Fig. 10.1.5). Considering the net heating (Fig. 10.1.1) it is evident that in both models the radiative cooling predominates over the other processes above the 600 mb-level whereas the sum of condensational heating and surface fluxes predominate below that level. Nevertheless Fig. 10.1.4 indicates a high degree of "local" compensation between radiation and other physical processes especially in the ECM-run. This is obvious in the lowest layer where the cooling due to rain evaporation leads to a weaker radiative cooling.

Considering the surface heat flux we found that the heat flux in the ECM-run was about twice as large as that in the GFD-run (Table 10.1.1). However, even the larger value in the ECM-run is considerably smaller than the climatological value of 27 Watts/m^2. This difference is probably due to the unrealistically large condensational heating in the boundary layer, which damps potential temperature gradients between the lowest layer and the surface. This has been confirmed in a test-run where no condensational heating was allowed to take place below 900 mb. The surface heat fluxes observed in this test run were twice as large (21.7 Watts/m^2) but still smaller than the climatological value.

Figs. 10.1.6 and 10.1.7 describe the variation of the vertically integrated terms of the sensible heat budget with latitude (with a distinction between land and sea and average). The reader should note the changes of scale from one diagram to the other. The sudden jump to zero of the land or sea curves corresponds to latitude where only land or sea is present.

The main differences here appear in the extra-tropical parts of the southern hemisphere and seem to be primarily due to unrealistically low cooling rates in the GFD radiation scheme.

One should notice in the diagrams for the sum of all diabatic processes that the land sea differences are, at their peak, of the same order of magnitude as the Pole-Equator contrasts. But a very significant difference is that these peaks are at mid-latitude for the ECM-run and at the equator for the GFD-run.

Although different in the global mean, as seen previously, both the large scale precipitation and moist convective effects on temperature have very similar features in both runs with regard to both the land-sea contrasts and latitudinal variations.

The atmosphere is driven by non-adiabatic heating such that zonal available potential energy and also eddy available potential energy are generated. The major contributions come from differential radiative and condensational heating. Besides the horizontal variation of heating, the vertical variation of heating also determines the amount of potential energy being generated. As was shown before, both runs differ a lot with regard to the vertical profile of heating, especially in the lower troposphere (Fig.10.1.1 and 10.1.2). Unfortunately, the generation of available potential energy by diabatic heating was not calculated for these experiments. However, Fig.10.1.8 shows instead the conversion from zonal to eddy available potential energy for wavenumber 4 to 9.

The meridional distribution agrees well between both integrations and the observations. However, the vertical profile (right part in Fig.10.1.8) is different, showing unrealistically large conversion rates in the lower troposphere in the GFD-run and a more realistic distribution in the ECM-run, although the conversion rate in the ECM-run is too large compared to the observed rate. The very large conversion rate near the surface in the GFD-runs is probably related to the strong diabatic heating by condensation observed there.

Considering the net heating the two physics schemes showed considerable differences (Fig.10.1.9). The largest differences were observed in the low latitude higher troposphere, where the ECM-scheme gives net heating and the GFD-scheme gives net cooling. The net heating observed in the ECM-run is probably causing the poleward shift of the subtropical jet (Section 11). Large differences are also noticeable in the boundary layer, showing more heating in the GFD-run than in the ECM-run.

10.2 Kinetic energy

During the 10-day period the total kinetic energy increased in both runs, and more so in the GFD-run than in the ECM-run. The budget for the global mean of kinetic energy (Table 10.2) shows that both the generation of kinetic energy by energy-conversion and the dissipation are larger by about 35% in the GFD-run. The larger energy conversion in the GFD-run may be related to larger generation of zonal available potential energy. A difference noted was that the generation of available potential energy in the GFD-model appeared at lower heights probably because of the concentration of latent heat release at the very low layers. As no reliable data on the mean

TABLE 10.2

Global mean rates of dissipation and of conversion from potential to kinetic energy (Watts/m^2) for the 10-day period for the ECM-run and for the GFD-run.

	ECM-run	GFD-run
Dissipation	2.38	3.17
Energy conversion	2.64	3.60
Net change of kinetic energy	0.26	0.43

global dissipation and energy conversion are available, the model values cannot be compared with real values. Values for northern hemisphere estimated by various authors (see Newell et al, p.73) range from 2.2 to 10 Watts/m^2. Fig.10.2.1 shows the contributions of the different levels to the dissipation both by vertical diffusion and by horizontal diffusion. The main differences between both runs are: the dissipation near the surface is much larger in the GFD-run which again indicates that the noise generated there in association with local release of latent heat is more intense. Above the 800 mb-level the sum of vertical and horizontal diffusion in the ECM-run yields about the same amount of dissipation as the horizontal diffusion in the GFD-run by itself. The overall dissipation by horizontal diffusion took about half of the total dissipation in the GFD-run, but less than a quarter in the ECM-run.

Table 10.3 shows how the total dissipation differs over sea and land. As can be seen, the dissipation rates in both runs are nearly the same over land but are very different over sea, where far more energy is dissipated in the GFD-run.

In a similar way Table 10.4 shows how the dissipation rates by vertical mixing have their largest differences in the extra-tropical part of the southern hemisphere.

All of this can be found with more details in Fig. 10.2.2 where the distribution with latitude over both land and sea for both types of dissipation are plotted. The extreme values around the North Pole in the horizontal dissipation are so far unexplained.

TABLE 10.3

Global means of dissipation rates (Watts/m^2) by vertical diffusion and by horizontal diffusion over land and over sea.

		Land	Sea	Land and Sea
VDIFF	ECM	2.81	1.57	1.97
	GFD	2.38	1.99	2.12
HDIFF	ECM	0.49	0.41	0.44
	GFD	1.03	0.87	0.93
TOTAL	ECM	3.30	1.98	2.41
	GFD	3.47	2.86	3.05

TABLE 10.4

Geographic distribution of the dissipation rates (Watts/m^2) by vertical diffusion

	ECM	GFD
Global	1.93	2.09
N. hemisphere	2.93	2.95
S. hemisphere	0.93	1.23
North of 30 °N	4.26	4.27
30 °N - 30 °S	1.00	1.12
South of 30 °S	1.45	1.83

10.3 Moisture

Due to its large meridional variation latent heat release is an important factor in driving the atmosphere. Therefore the sources and sinks of moisture are of considerable interest. Table 10.5 contains the global budgets for moisture. As is seen, in the ECM-run the model atmosphere receives less moisture by surface fluxes and also less moisture is condensed and precipitated than in the GFD-run. Gain and loss are fairly well balanced in the ECM-run, whereas precipitation outweighs the evaporation in the GFD-run, by which the total moisture content decreased by almost 20%. This is probably due to net cooling in the lower troposphere (Fig. 10.1.1) which affects the storage of moisture. The different net changes of moisture in both runs may also depend on the different assumptions concerning the criterion for condensation - in the GFD-run condensation occurs whenever 80% is exceeded, whereas a height dependent criterion is used in the ECM physics decreasing from the surface value 100% to an asymptotic value 80% at higher levels.

TABLE 10.5

Global moisture budget. Mean rate of accession and loss of moisture due to surface fluxes and precipitation (mm H_2O/10 day) for the 10-day period

	ECM		GFD	
Surface fluxes	27.5		31.8	
Large-scale precipitation	16.6	28.5	18.6	36.4
Convective precipitation	11.9		17.8	
Net changes of atmospheric moisture	-1.0		-4.6	

Table 10.6 enables us to see that this imbalance of the GFD moisture budget (relative to the ECM one) is global rather than being geographically localised. There is however a better overall agreement between both models for the individual components in the Northern Hemisphere.

TABLE 10.6

Geographical distribution of the moisture budget. Same units as for Table 10.5.

	Surface fluxes ECM/GFD	Large-scale precipitation ECM/GFD	Convective precipitation ECM/GFD	Net changes of atmospheric moisture ECM/GFD
Global	27.5/31.8	16.6/18.6	11.9/17.8	-1.0/-4.6
N.Hemisphere	31.6/33.2	17.3/17.4	11.5/16.1	+2.8/-0.3
S.Hemisphere	23.3/30.4	15.9/19.8	12.3/19.5	-4.9/-8.9
Nth of 30 °N	21.1/20.3	22.2/19.0	4.9/9.2	-6.0/-7.9
30 °N-30 °S	36.7/42.6	12.7/17.4	19.9/26.1	4.1/-0.9
Sth of 30 °S	15.5/21.6	18.6/20.7	3.0/9.7	-6.1/-8.8

This is confirmed by the results shown on Fig. 10.3.1 where the variation with latitude and the land sea contrasts are plotted for the moisture budget. It is interesting to note that the land sea contrast (except for the tropics in the GFD run) is relatively far less in the sum of both terms than in the individual ones. In the atmosphere there is an advection of water vapour from sea to land which is compensated by the rivers' flow to sea. Therefore differences between land and sea should be expected in the global budget at least in the mid latitudes. Their absence indicates that the life time for the atmospheric water vapour is too short. This observation is consonant with the remarks previously made about precipitation in the boundary layer.

The meridional distribution of net moistening agrees well with real data values (Lorenz, 1967), where precipitation dominates in mid-latitudes and near the equator and evaporation dominates in the subtropics. (See Figs. 10.3.2 and 10.3.3).

10.4 Summary

The main points highlighted in this section are the following.

- Both runs seem to overpredict latent heat release in the boundary layer, especially the GFD run; this is one of the reasons for an underestimation of the sensible heat flux at the surface in both runs.
- There is a good balance in the global heating sources and this balance, essentially between radiation and latent heat release, has a more detailed vertical structure in the ECM-run in the lower troposphere.
- The kinetic energy budgets are satisfactory in both runs and the differences in the intensity of the energy transformations are due to different physical forcing in the generation of available potential energy.
- The global humidity budget of the ECM-run is more satisfactory than that of the GFD-run although the geographical distributions of the anomalies are very similar. As mentioned above this difference (exaggeration of the hydrological cycle in the GFD-run) is probably the source of a more active energy transfer in the GFD-run.

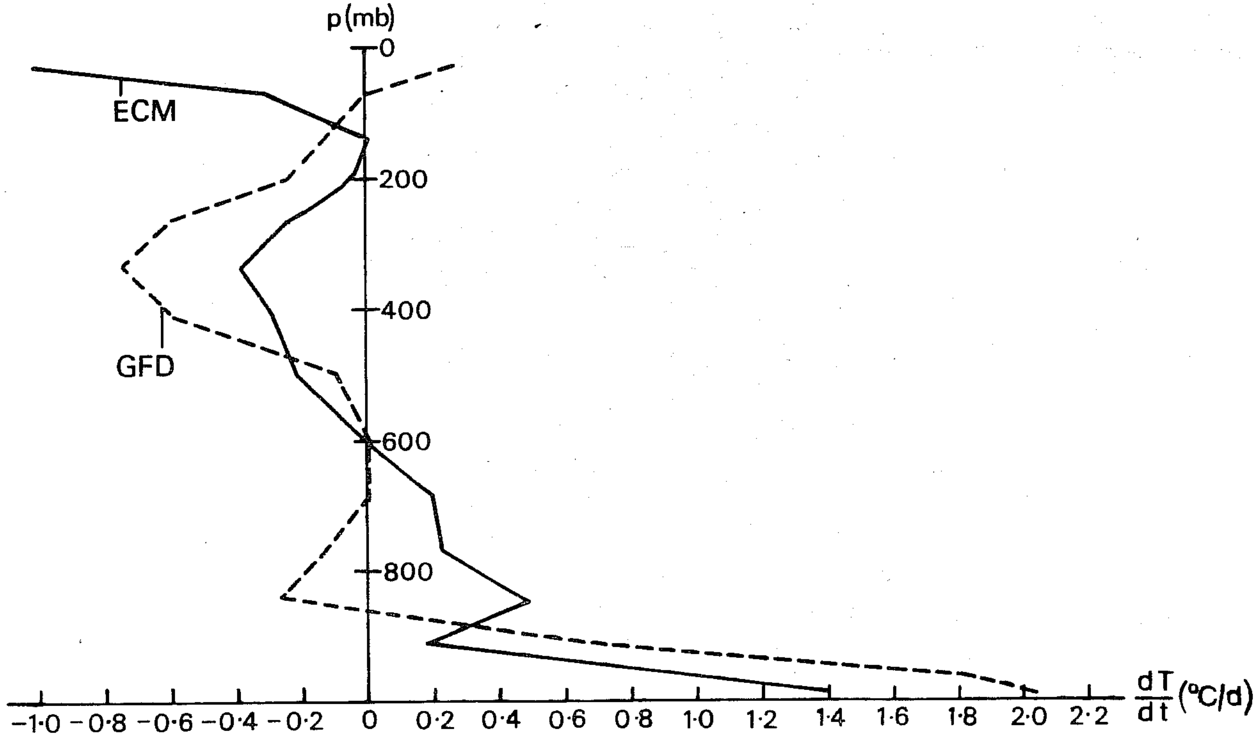


Fig.10.1.1 Vertical distribution of global mean of net diabatic heating (^{o}C/day) for the 10 day forecast period, starting from 22.2.76, for ECM-run and GFD-run.

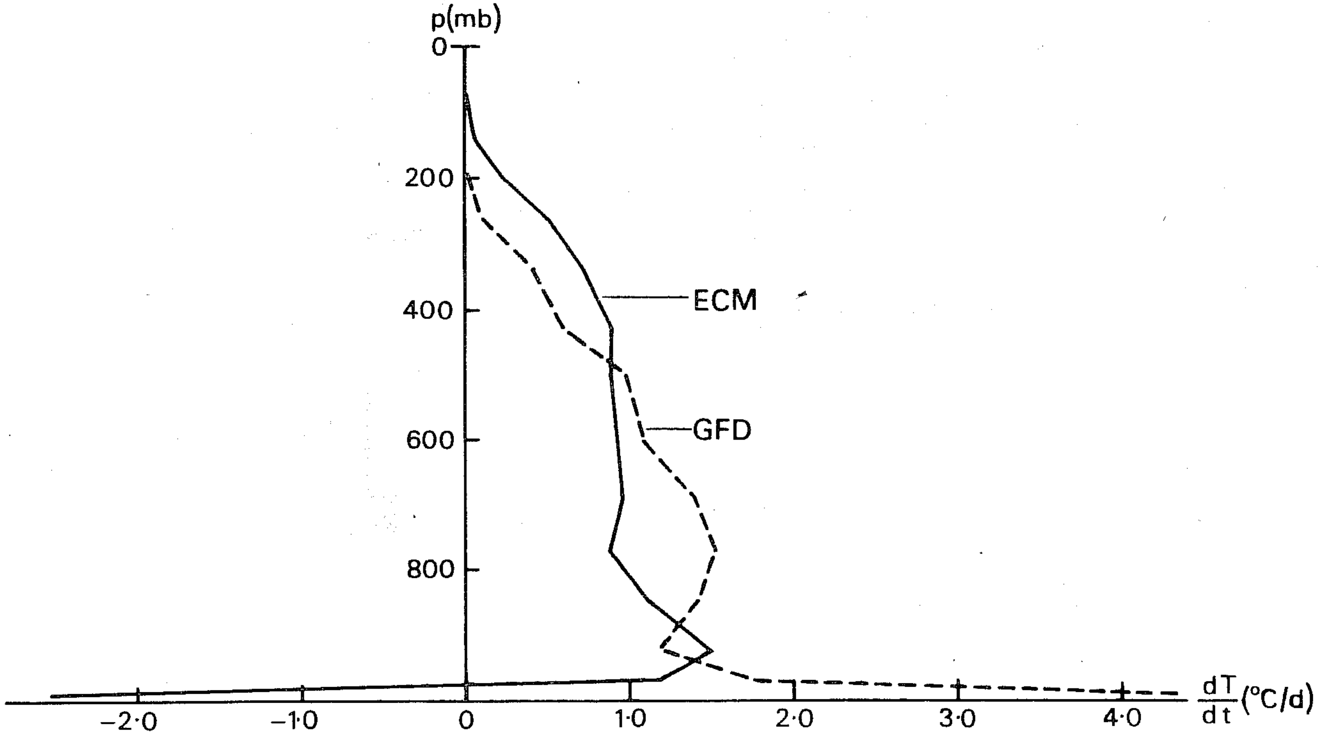


Fig.10.1.2 Vertical distribution of global mean heating (°C/day) by latent heat release for the 10 day forecast period for ECM-run and GFD-run.

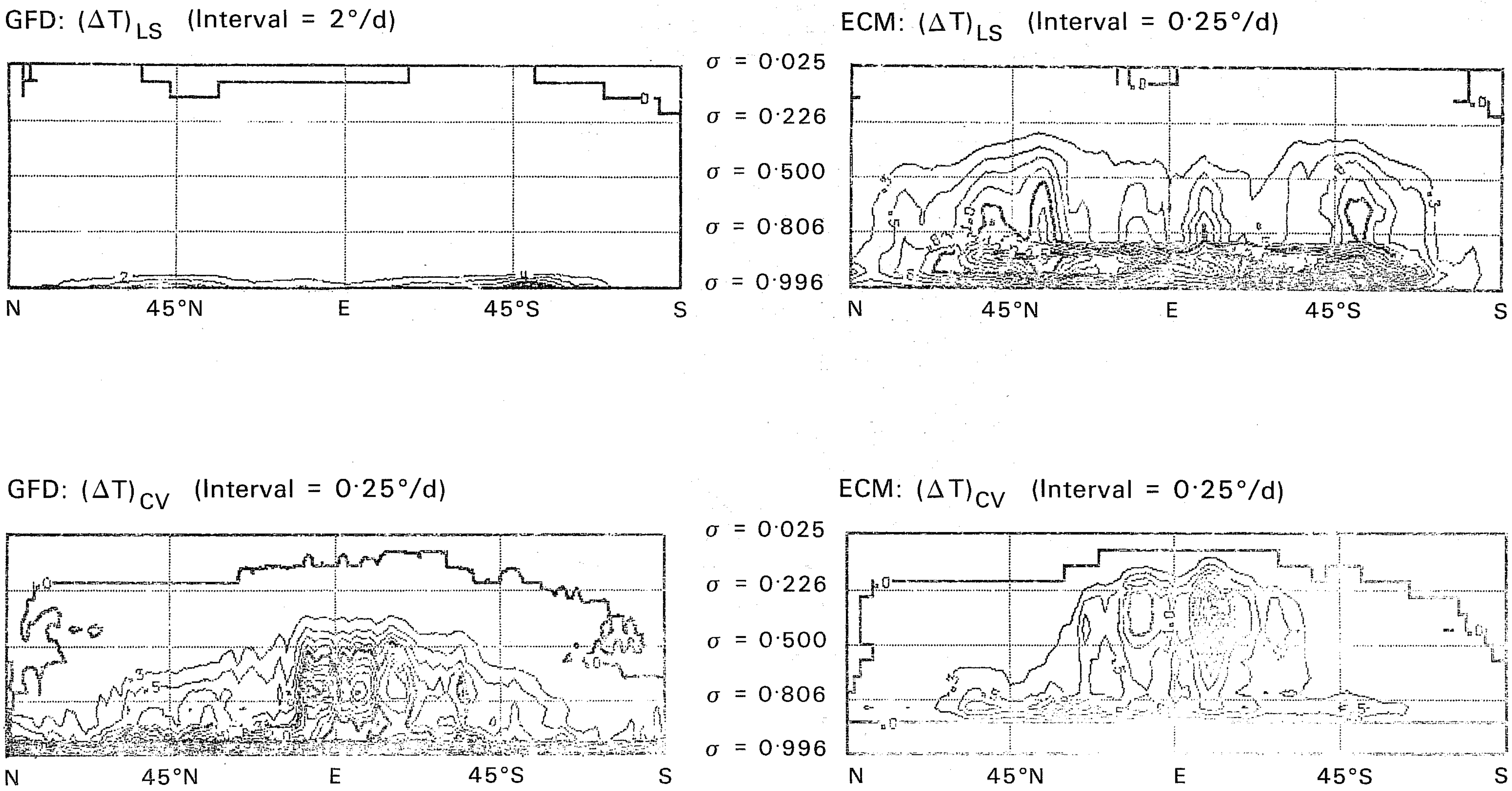


Fig.10.1.3 Latitude-height distribution of zonal means of heating due to large-scale condensation $(\Delta T)_{LS}$ and due to moist convection $(\Delta T)_{CV}$ for the 10 day forecast period for the run with GFD-physics (left) and for the run with ECM-physics (right). Notice the different intervals for $(\Delta T)_{LS}$ for GFD-run and for ECM-run.

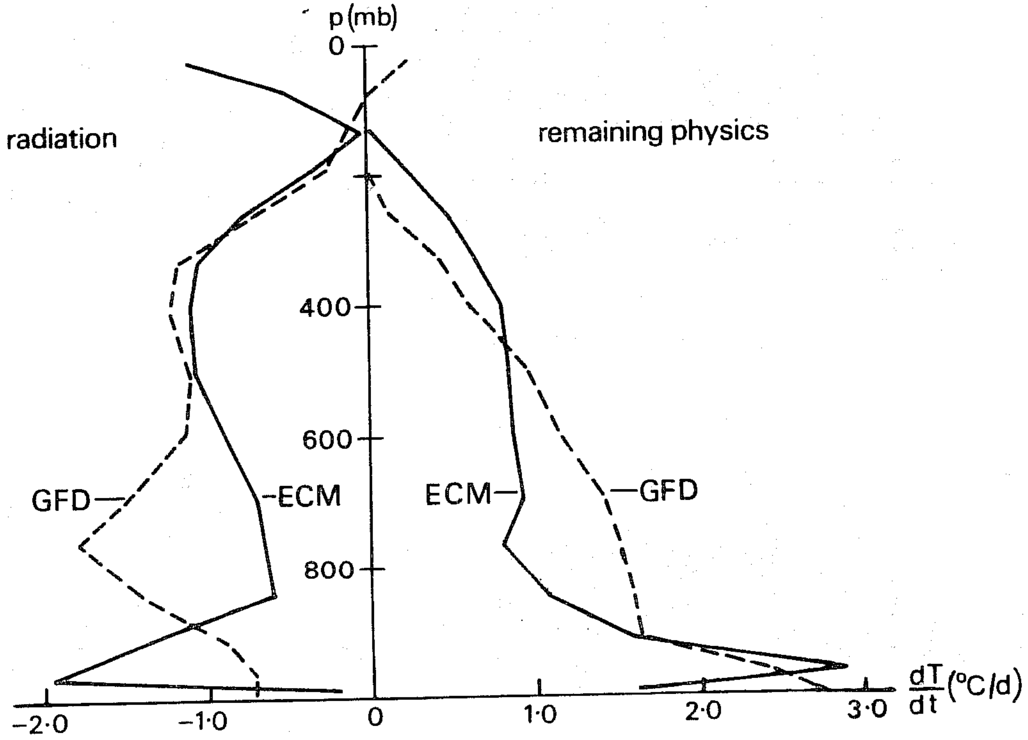
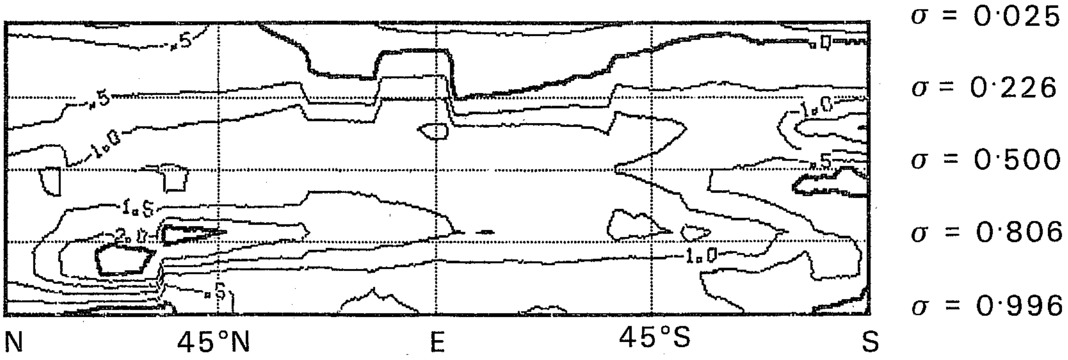
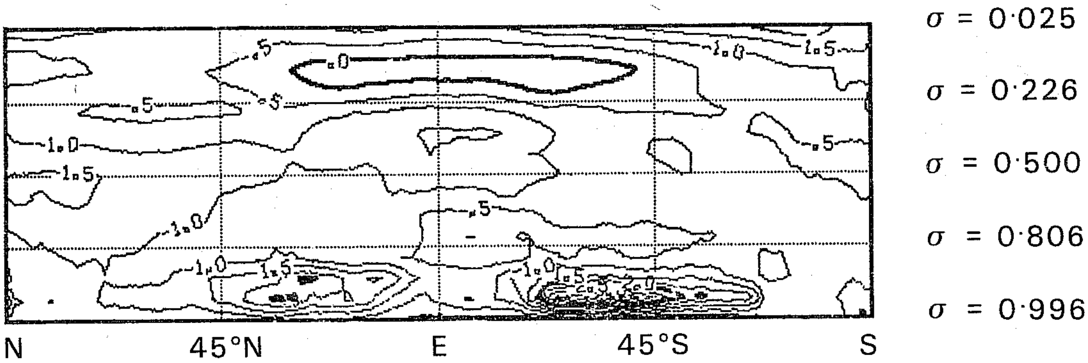


Fig.10.1.4 Vertical distribution of global mean heating by radiation (left part) and by all other physical processes (latent heat release + surface fluxes + energy conversion) for ECM-run and GFD-run.

GFD: $(\Delta T)_{rad}$



ECM: $(\Delta T)_{rad}$



DOPPLICK (1972): $(\Delta T)_{rad}$

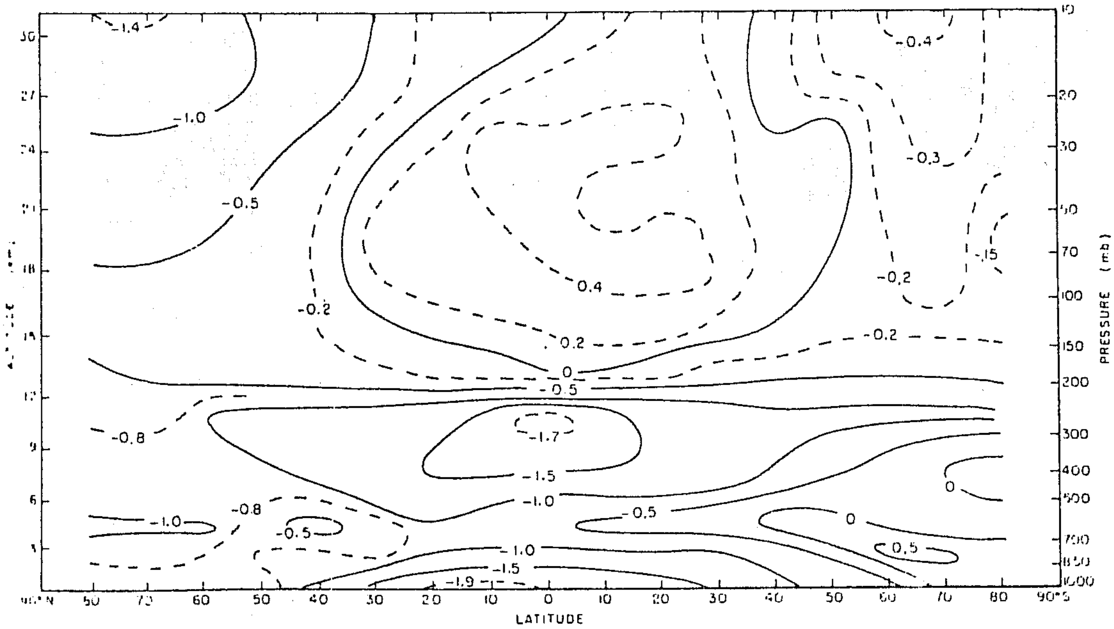


Fig.10.1.5 Latitude-height distribution of zonal means of radiative heating (^{o}C/day) for the 10 day forecast period for GFD-run (top), ECM-run (middle) and mean radiative heating obtained by Dopplick (1972) for the period December - February.

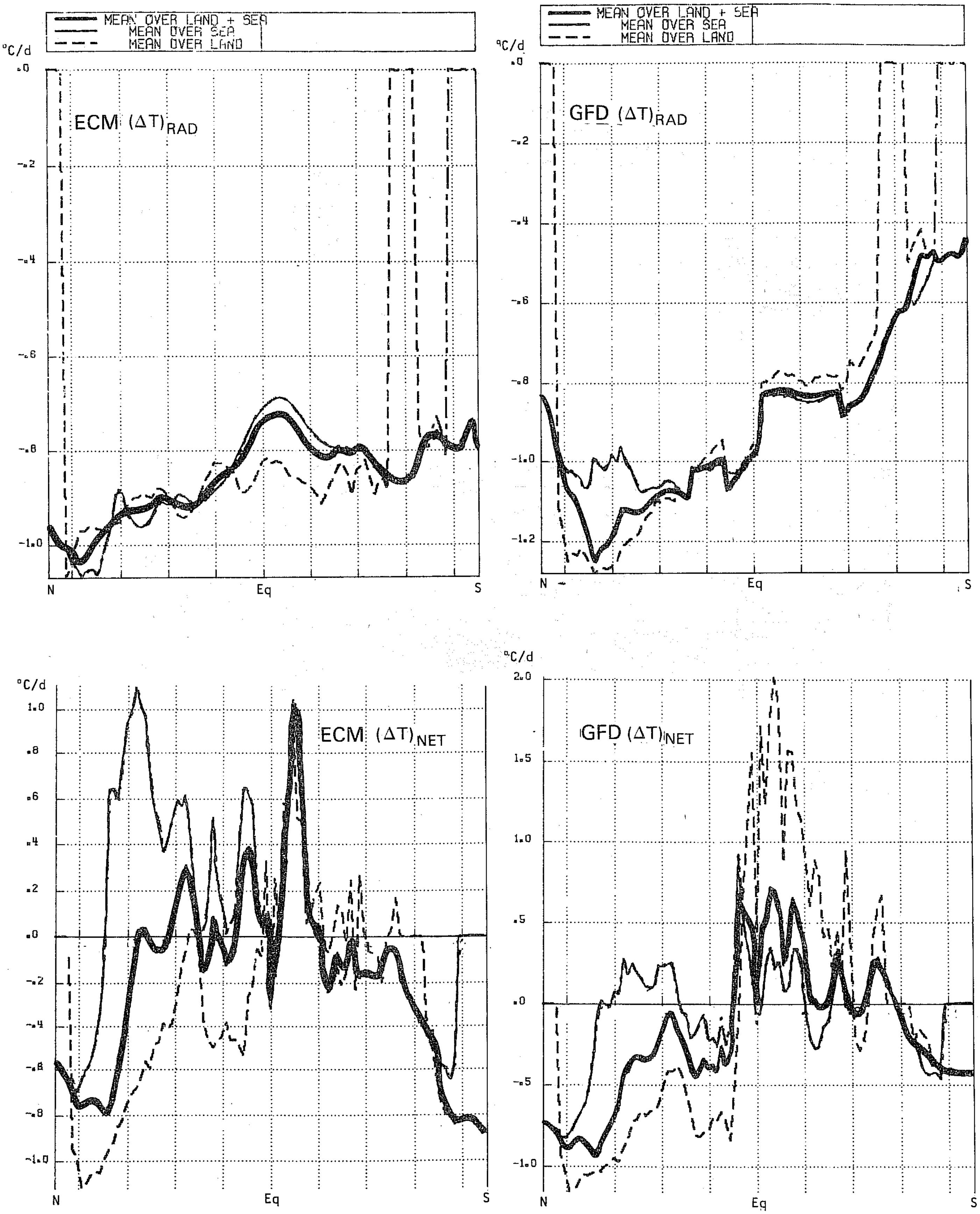


Fig.10.1.6 Meridional distribution of mean heating rates (^{o}C/day) for the 10-day forecast period for ECM-run (left) and for GFD-run (right) a) due to radiation (top) and b) due to all diabatic processes (bottom)

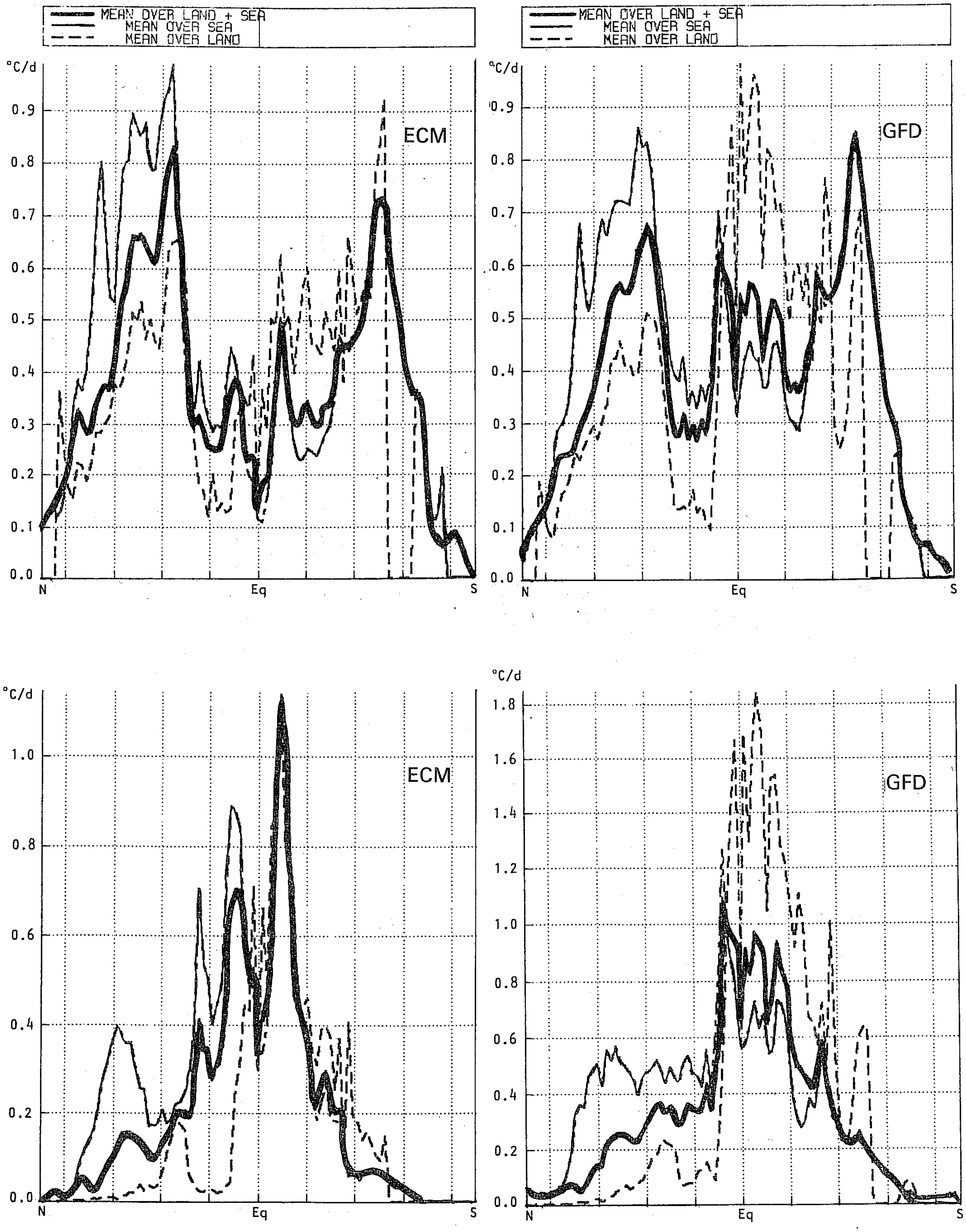


Fig. 10.1.7 As for Fig. 10.1.6 but a) due to large-scale condensation processes (top)
b) due to moist convection (bottom)

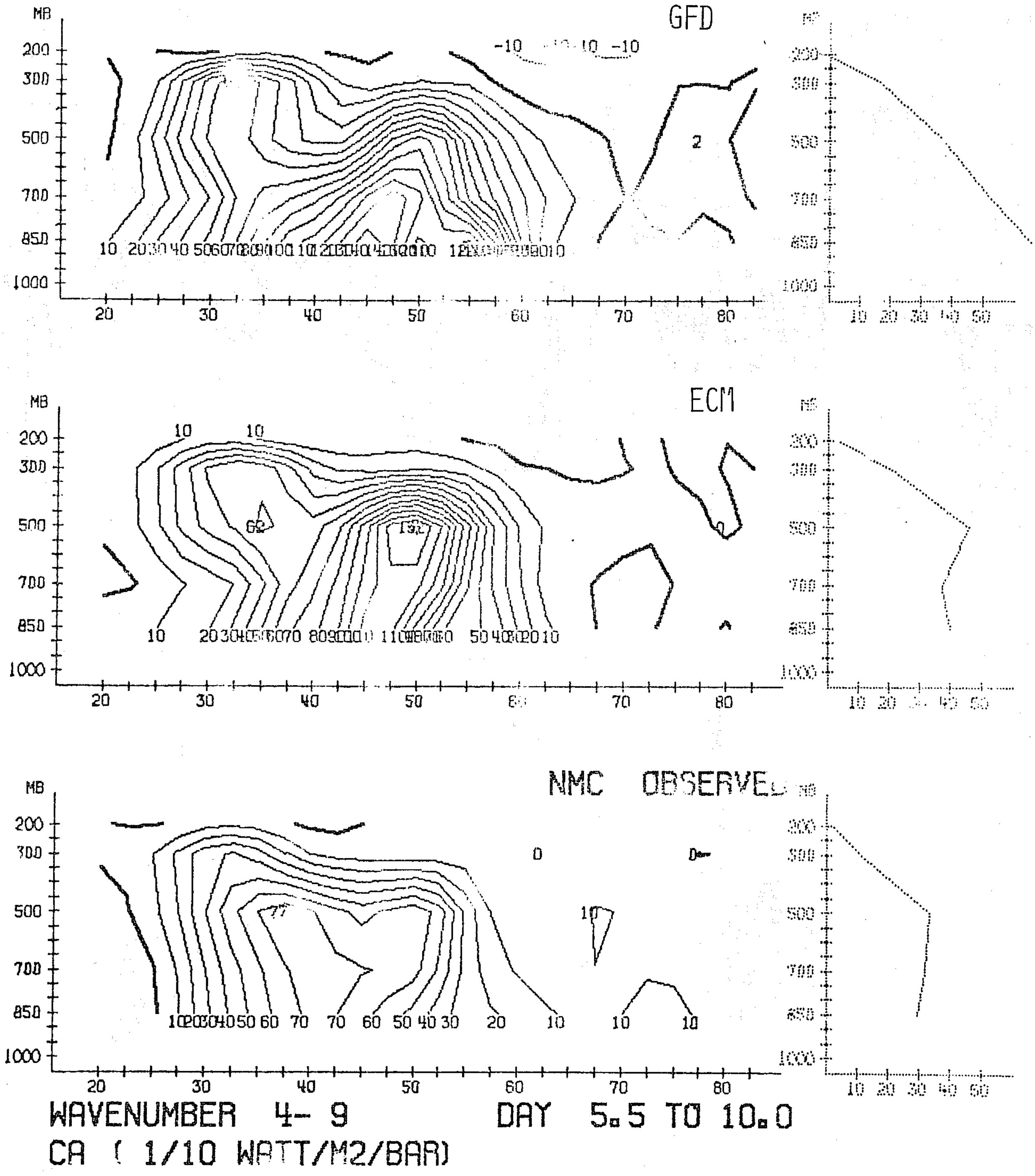


Fig.10.1.8 Latitude-height distribution of generation of available potential energy at wavenumbers 4-9 by conversion from zonal to eddy available potential energy for EC-run, for GFDL-run and for observed state.

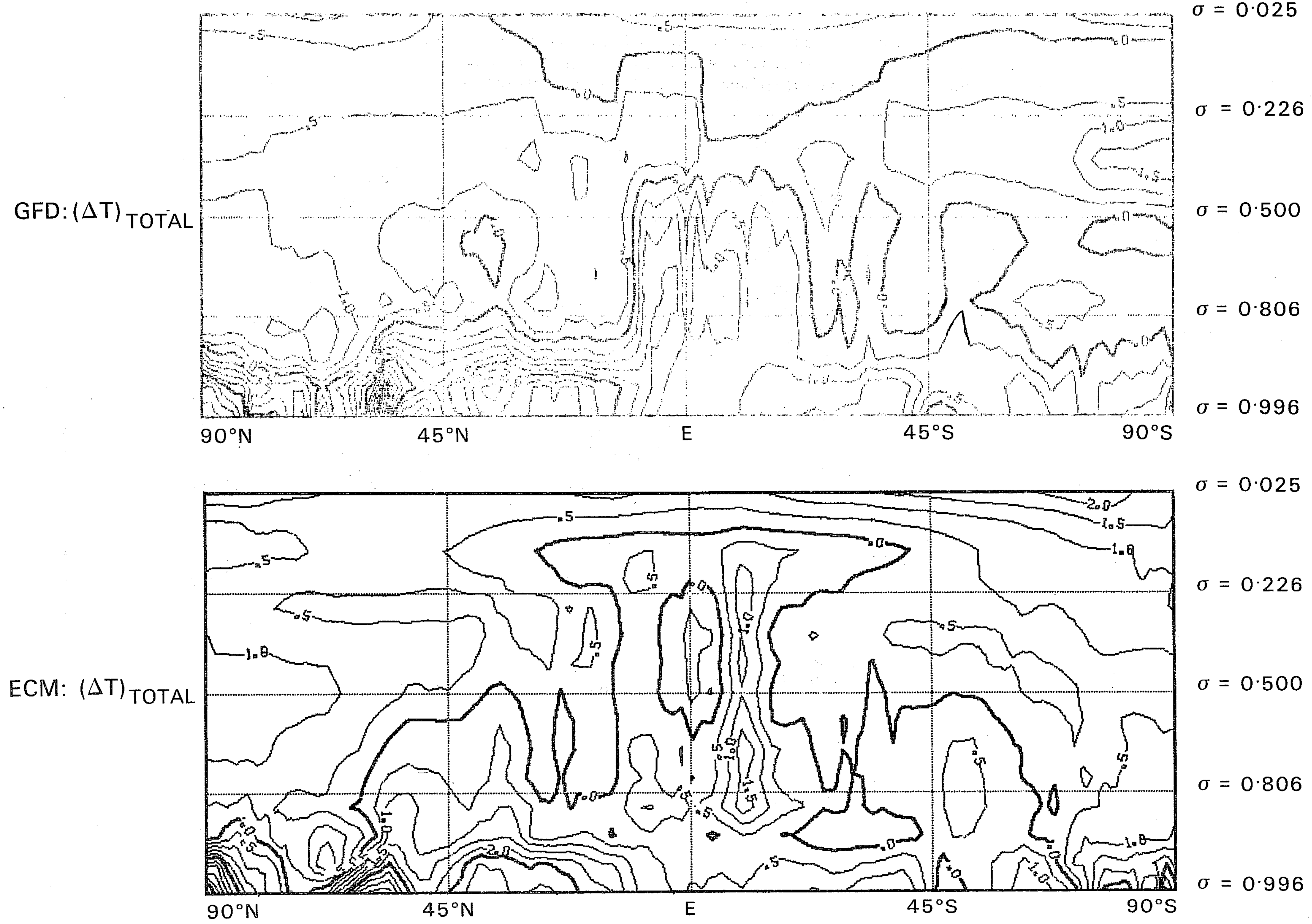


Fig.10.1.9 Net heating rates (^{o}C/day) due to all physical processes for GFD-run (top) and for ECM-run (bottom).

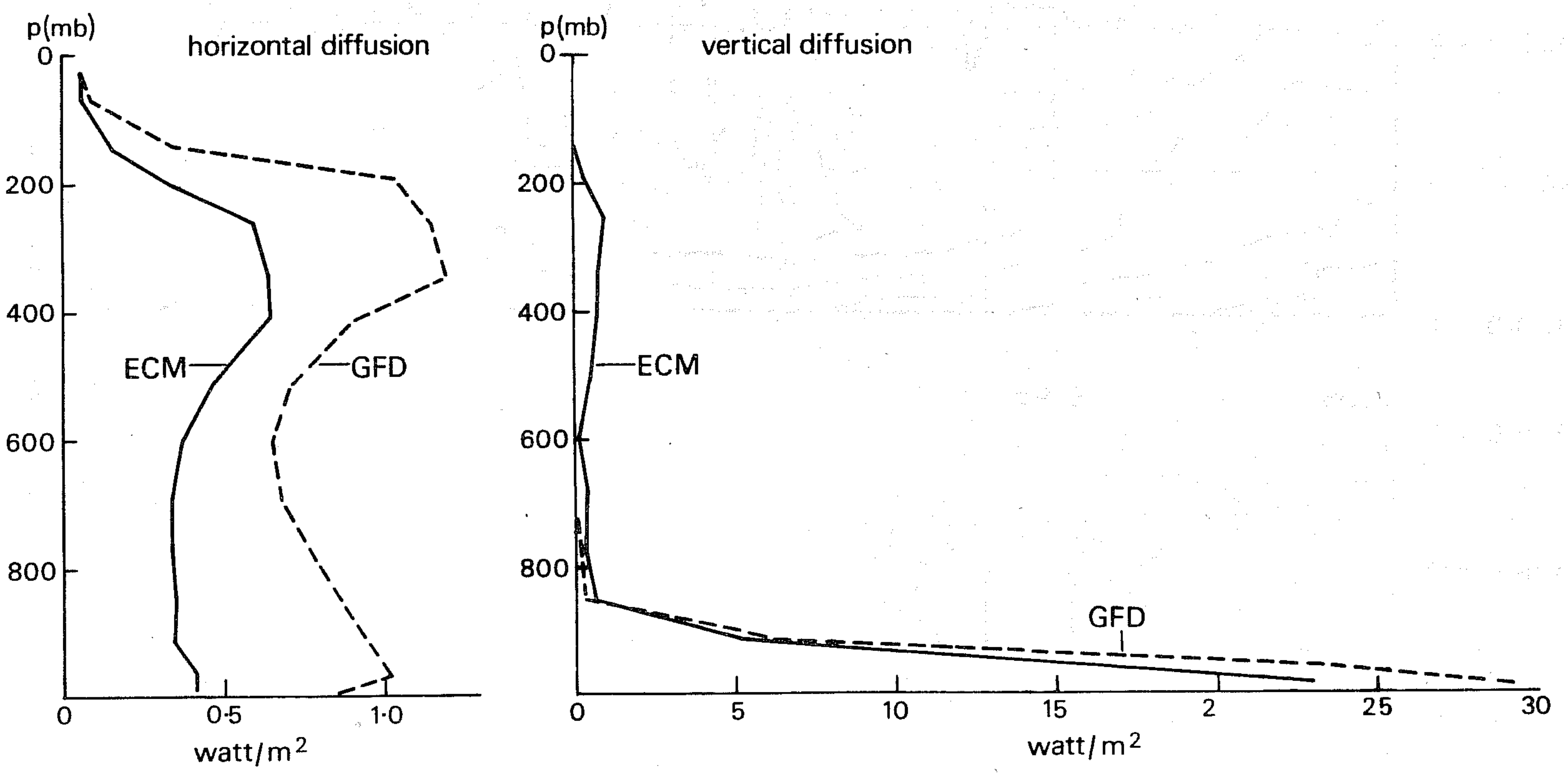


Fig.10.2.1 Vertical distribution of global mean dissipation rate due to vertical momentum diffusion and horizontal momentum diffusion for the 10-day forecast period for ECM-run and GFD-run.

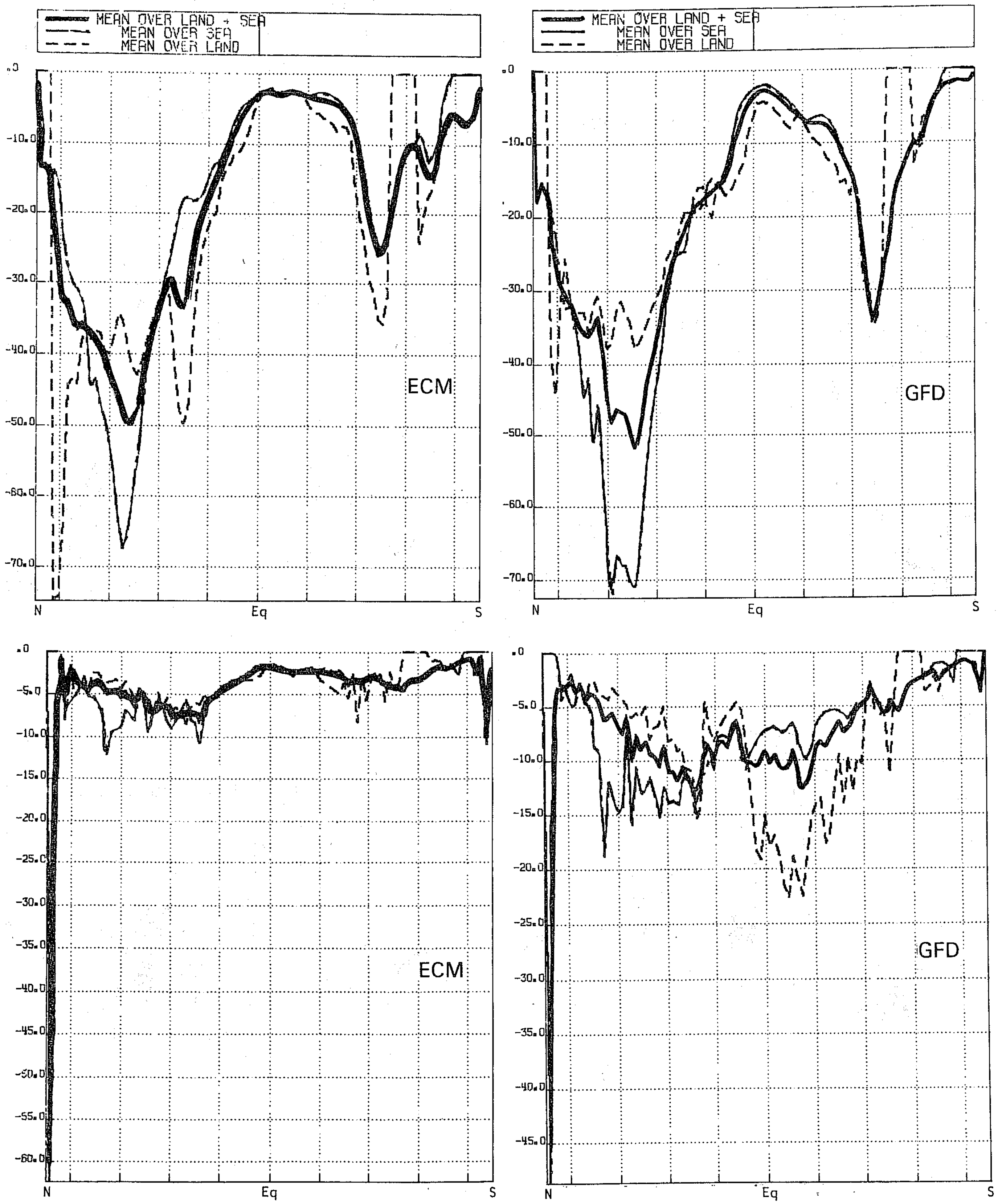


Fig.10.2.2 Meridional distribution of dissipation for the 10-day forecast period for ECM-run (left) and for GFD-run (right) a) due to vertical diffusion (top) and b) due to horizontal diffusion (bottom). Units: $m^2s^{-2}d^{-1}$.

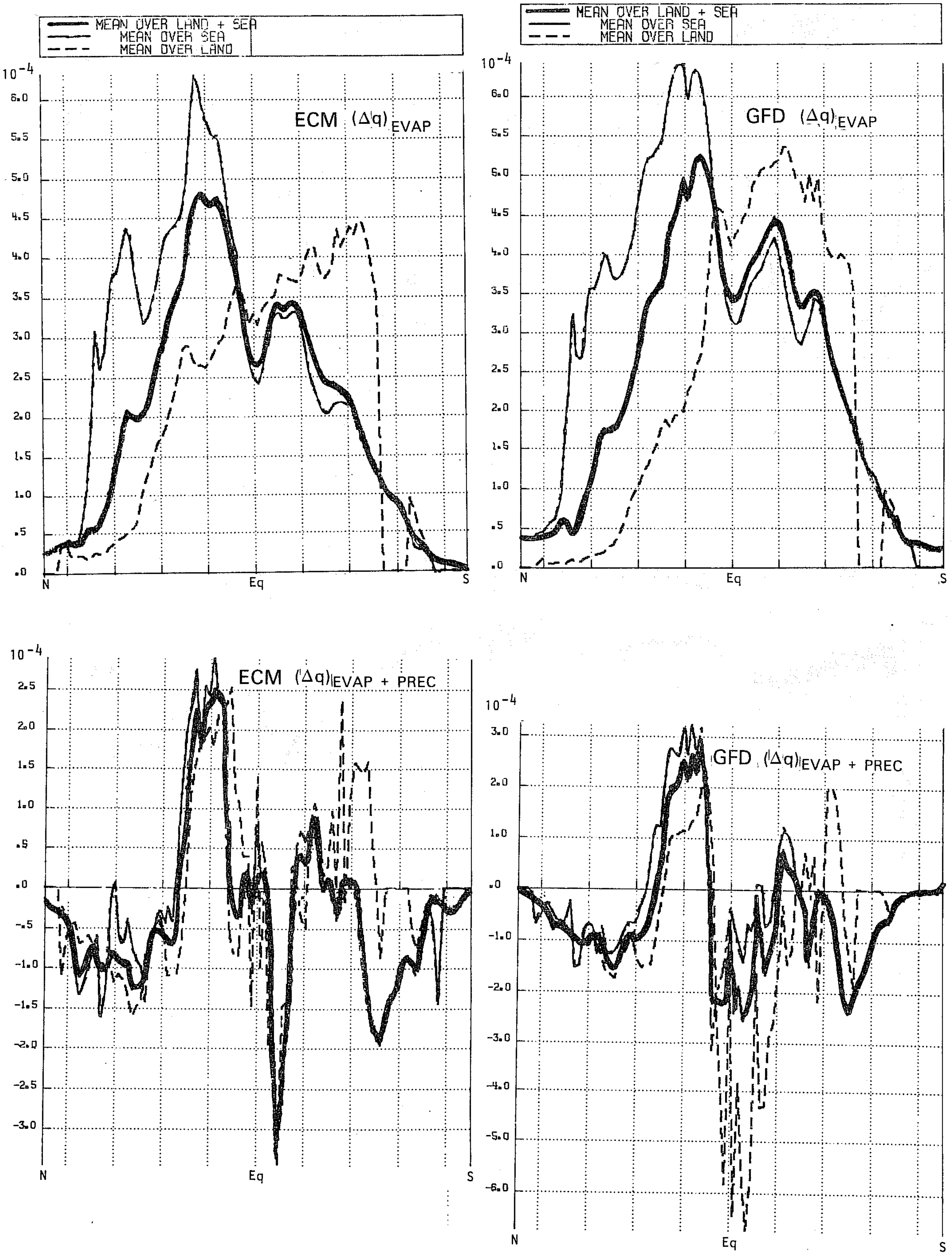


Fig.10.3.1 Meridional distribution of change of moisture for the 10-day period for ECM-run (left) and for GFD-run (right) a) due to surface fluxes (top) and b) due to all physical processes. Units: day^{-1}.

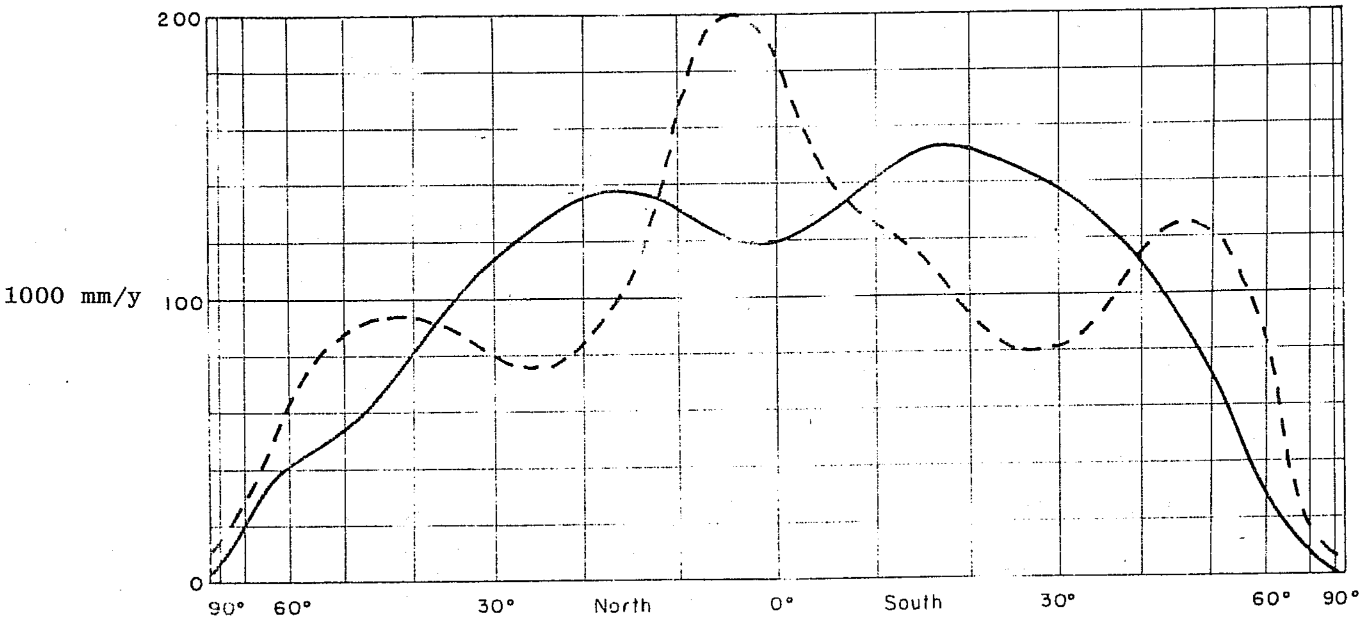


Fig.10.3.2 Average annual evaporation (solid curve) and precipitation (dashed curve) per unit area as given by Sellers (1966). Values are in centimetres of water per year, or g cm^{-2} year^{-1} (scale on left). Taken from Lorenz (1967).

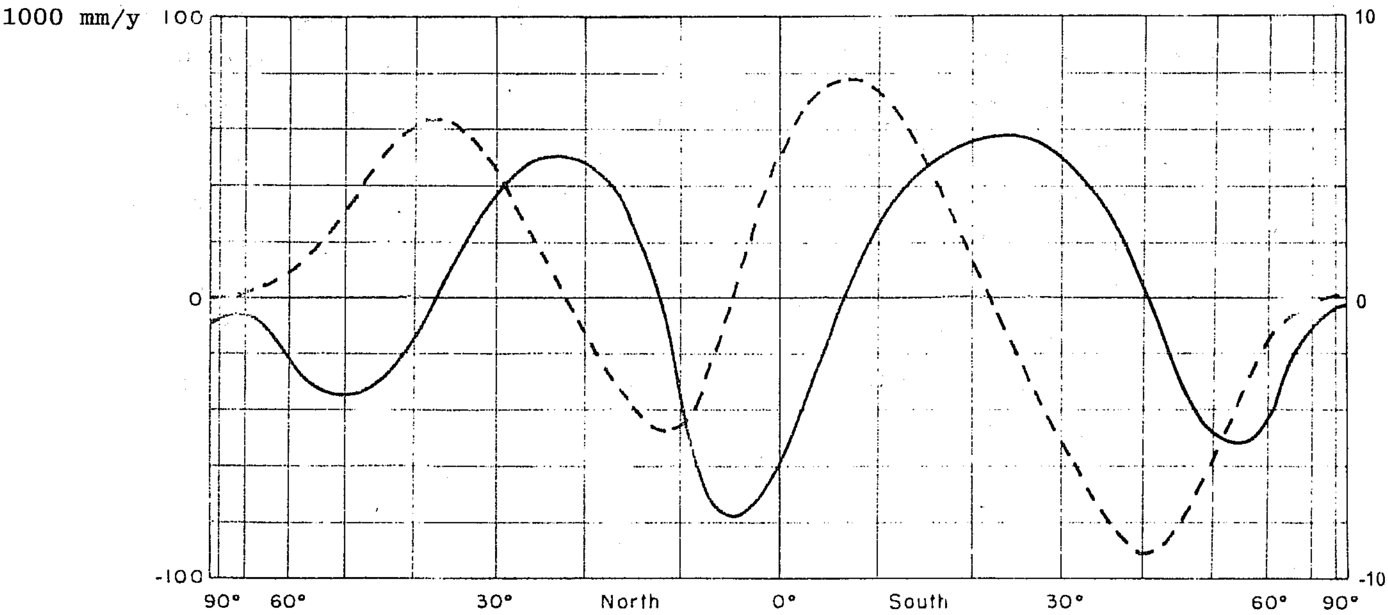


Fig.10.3.3 Excess of evaporation over precipitation (solid curve) as given by Sellers (1966), in g cm^{-2} year $^{-1}$ (scale on left), and northward transport of water in the atmosphere required for balance (dashed curve) in units of 10^{11} g sec^{-1} (scale on right), taken from Lorenz (1967).

11. Wavenumber decomposition of the budgets of energy sensible heat and momentum in the extra-tropics

In this section we consider the energetics of the models in some detail for the region north of 20 N between the surface and 200 mb. We shall try to show that the dominant errors in the long wave energetics are reduced amplitudes in that part of the flow which should be stationary. The wavenumber band four to nine is largely transient and amplitudes here are rather too large.

11.1 Time evolution of the ensemble averaged energy statistics and the evolution of the energy statistics of the ensemble averaged flow

Fig. 11.1.1a shows the evolution in time of the kinetic energy in different wavenumber groups, averaged over the seven forecasts, for the observations and for both of the models for the extra-tropics below 200 mb. If the ensemble of cases were sufficiently large then the curves for the observations would be horizontal straight lines. Several features are noteworthy. The zonal kinetic energy for both models steadily increases through the forecast period with the GFD model showing a 35% increase in zonal kinetic energy by the tenth day while the EC model shows an increase of 15%. Between days four and nine both models show an underestimate of long wave kinetic energy. Between days four and ten both models also show an overestimate of kinetic energy of up to 25% in wavenumbers 4-9, with largest errors in the EC model.

Fig. 11.1.2a shows the corresponding curves for the available potential energy. The zonal available energy (ZAPE) increases pari passu with the zonal kinetic energy (ZKE). Our energetics calculations are incomplete because we cannot at the moment calculate the conversions zonal kinetic energy (ZKE) to zonal available energy (ZAPE) or eddy available energy (EAPE) to eddy kinetic energy (EKE). Moreover we do not know the boundary fluxes of energy across 20 N or across the 200 mb level. Nevertheless the simultaneous increase of ZAPE and ZKE would suggest that the increase in ZAPE occurs, at least in part, through an erroneously large generation of ZAPE by the physical parameterizations.

In wavenumbers 1-3 and 4-9 the level of EAPE for the ECM model is higher than observed, while the GFD model shows closer agreement with observation.

Fig. 11.1.3a shows the evolution of the term CA$(-\overline{v'T'}\,\frac{1}{a}\frac{\partial \overline{T}}{\partial \phi})$. For most of the time this term is underestimated by the models in the long waves and overestimated in the medium waves. For both models we notice that in the medium waves there is a good correlation between EAPE and CA for most of the time with the conversion term leading by about a day. In the ECM model at days 7 - 8 the contribution of the long waves to CA is about half that of the medium waves while in the observations the contributions are almost the same.

Fig. 11.1.4a shows the time evolution of the term CK $(-\frac{\overline{u'v'}\cos\phi}{a}\frac{\partial}{\partial\phi}\frac{\overline{u}}{\cos\phi})$ by wavenumber group. In the long waves the term is predominantly too negative. This is even more true in the medium waves, particularly for the EC model. Despite the fact that the implied driving for the ZKE through the term CK is larger for the ECM model the growth of ZKE is larger in the GFD model, and the dissipation is also larger there (cf Section 10).

To summarize these results then we see that,

(i) The ZKE and ZAPE grow simultaneously in both models indicating the likelihood that both parameterization schemes have too large generation of ZAPE.

(ii) The level of energetic activity in the long waves is too low in both models.

(iii) The level of energetic activity is too high in the medium waves.

We now discuss the sense in which the errors in the long wave energetics are in the "stationary" part of the flow.

In Section 9 we discussed the systematic errors of the forecasts, i.e. the differences between the day-n ensemble mean of the forecasts and the day-n ensemble mean of the observations. We saw there that the dominant components of the systematic errors were in zonal wavenumbers 2 and 3. If we had sufficiently large ensembles of observations then all the day-n ensemble means for the observations would be essentially the same climatological mean. The evolution, with n, of the day-n ensemble means of the forecasts could then be interpreted as the erroneous evolution of that part of the flow which should be stationary.

We now want to apply this concept to the budget calculations. We shall speak of the ensemble average of a particular term (e.g. the ensemble average of the kinetic energy) as the "total" part of the term. We shall speak of the same term calculated for the ensemble mean flow as the "S" part of the term ("S" for "should be stationary") and we shall speak of the difference between the two as the "T" (for "transient") part of the flow. The distinction between the total and the S parts of a term is the distinction between the average of a statistic on the one hand and the same statistic calculated for the average state, on the other.

Fig. 11.1.1b shows the evolution of the kinetic energy in the forecast and observation ensemble means. This is the part of the flow that would be constant with day-n (at least for the observations) if the ensembles were sufficiently large. The bulk of the energy is contained in the zonal component and in wavenumbers 1-3. We see a progressive increase in the forecast zonal kinetic energy and a progressive decline, from the start of the forecast, in the long wave kinetic energy. A comparison of Figs. 11.1.1a and b shows that for the observations the zonal kinetic energy is almost entirely S, the kinetic energy in wavebands 4-9 and 10-20 is almost entirely T and that at least half of the kinetic energy in waves 1-3 is S. The model errors in the zonal kinetic energy are due almost entirely to the S part while the model errors in the kinetic energy of waves 4-9 is almost entirely due to the T part. The forecast underestimate of total long wave kinetic energy is smaller than the underestimate in the S part. If we take day 8 as representative of the last part of the forecast then the observations show 29 units of energy in long wave S energy and 22 units in T. The EC model shows 17.5 units in S and 27.5 units in T. Thus the more serious error is in the S part of the flow, for the long waves. Fig.11.1.2b shows the corresponding results for available potential energy. Significant amounts of S available potential energy occur only in the zonal component and the long waves. We can see that most of the growth of the zonal available potential energy can be assigned to the growth of the S part of the zonal available potential energy. In the long waves the S part of the eddy available potential energy is consistently underestimated by the models except in the last day and a half.

Fig. 11.1.3b shows the corresponding results for the term CA. We see that the models' underestimate of the CA term in the long waves may be largely attributed to the underestimate of this term by the S part of the flow.

Fig. 11.1.4b shows the corresponding picture for CK. There we see that in the models the contribution of the S part in the long waves is too negative just as we saw for the total flow.

These results strongly suggest that the errors in the forecasts of the long waves are largely due to errors in the forecast of the S part of the flow. So far we have only looked at volume integrals of the energetic quantities. In Section 11.3 we shall examine this question in more detail but as a preliminary we examine the spectra of kinetic energy at different levels.

11.2 Spectra of kinetic energy

Fig. 11.2.1 shows the spectra of observed and forecast kinetic energy averaged over the seven cases and days 7 to 10 at 1000, 850, 500, and 300 mb (between 40 N and 60 N). There is a change of scale between the first two and the second two frames. In the lower levels we see that the models tend to overestimate the kinetic energy, particularly in wavenumbers 2-10. At 300 mb there is a tendency for the models to underestimate both the long wave energy and the short wave energy with an overestimate at some intermediate wave lengths. The most successful forecasts from the point of view of spectra, are, perhaps, the 500 mb forecasts. The models appear to follow a k^{-3} spectrum reasonably well at 500 mb but appear to have a somewhat steeper spectrum at 300 mb and a less steep spectrum below 500 mb.

The models' overestimates of kinetic energy at low levels is of course associated with the excessively strong surface winds in both models which have already been referred to. In order to study this deficiency and the underestimates of long wave kinetic energy at high levels we present latitudes height cross-sections of the kinetic energy by wavenumber groups.

11.3 Latitude height cross-sections of budget calculations for the zonal flow for days 7-10

Fig. 11.3.1 shows the forecast and observed ensemble means of zonal wind together with the meridional integral of this quantity for days 7 to 10. We have already seen (Fig. 11.1.1a) that the total kinetic energy is largest in the GFD model.

Fig. 11.3.2 shows the differences in zonal wind (forecast-observed) for both models averaged over the ensemble for days 7 to 10. These differences show a consistent structure in space and time in both models. The two models show similar behaviour in mid latitudes with the mean zonal flow of the models being four or five ms^{-1} faster than the observed flow in the later period. The largest differences between the models occur in the subtropics. There, in the GFD model, for days 7 to 10, the zonal flow has accelerated at all levels with largest increases of 4 ms^{-1}, while in the ECM model there has been a marked deceleration of the flow above 500 mb with largest decreases of 6 ms^{-1}. This feature is the most striking difference between the behaviour of the models.

Figs. 11.3.3, 11.3.4 shows the corresponding averages, for models and observations, of the two main terms contributing to the momentum budget, i.e. the mean meridional circulation and the polward momentum flux. No observations are available for the mean meridional circulation, which is set to zero in the DST data. The values shown at low levels for these quantities may be slightly suspect because of the problem of extrapolating below topography. There are certainly important differences between both models at high levels. The Hadley circulation seems to be stronger in the GFD model by ~ 20 cm/sec at 20 N. The Coriolis torque operating on a wind of 1 cm/sec can produce an acceleration of 1 m/sec in ten days. On the other hand the Ferrel cell at high levels in the ECM model seems to be more intense than in the GFD model. In higher latitudes the mean meridional circulation appears to be weak in both models.

There are important differences between the high level momentum fluxes in both models. For the momentum fluxes we do have observations. At 20 N we see that the momentum fluxes out of the tropics are much stronger in the ECM model than in the GFD model, which is close to the observed values. The differences between the models in this term and in the $\overline{v}$ term is at least in the correct sense to explain the large differences in zonal wind at high levels at 20 N, provided we assume that the momentum fluxes are small at the equator. Between 30 N and 40 N, the region of the maximum of the momentum flux, the ECM model overestimates the maximum while the GFD model slightly underestimates it. Between 40 N and 50 N both models show a far too intense convergence of the momentum flux which is presumably related to the large accelerations that both models show in these latitudes. Betwen 50 N and 60 N both models show equatorward momentum fluxes where the observations show substantial poleward fluxes. This is presumably associated with the substantial decelerations which both models show between 60 N and 80 N. We have no indication from these results as to the reasons for the differences between the models in the region 60 N to 80 N.

It seems clear then that differences between the models both in the tropical circulation and in the structure of extratropical disturbances are contributing to the errors in the zonal wind. We cannot say much more about the tropical circulation but the extratropical aspects of the flow will be discussed further below.

We turn now to discuss the zonally averaged temperature field.

Fig.11.3.5 shows the zonally averaged temperature differences (forecast minus observed) for the same region and time as Fig. 11.3.2. There are several noteworthy differences in the behaviour of the models. In the ECM model the extra-tropical troposphere has warmed on the whole, while the reverse is the case for the GFD model. In the GFD model there is a tendency for stabilization, relative to observation in the lower troposphere while the reverse is the case for the ECM model. Both models show too much cooling near 200 mb. The EC model shows a distinct maximum of temperature increase (2 K) near 700 mb, between 20 N and 30 N. The EC model also shows a maximum of heating (3 K) over the polar ice.

Fig. 11.3.6 shows the poleward flux of sensible heat $\overline{v'T'}$ for the wavenumber group 1-20. In both models we see that the poleward flux is too strong almost everywhere and that this feature is particularly marked near the surface and again near the tropopause where the flux increases with height in marked disagreement with observations. Both these points are discussed further below.

11.4 Latitude height cross-sections of budget calculations for the long waves (zonal wavenumbers 1-3) for days 7-10

Fig. 11.4.1a shows the observed and forecast latitude-height distribution of long wave kinetic energy for the last three days of the forecast period, averaged over the seven cases. The overall features we noted from the spectra are clearly seen, viz, that there is a forecast overestimate of long wave energy at low levels and a forecast underestimate at high levels. These latter are the levels that predominate in the volume integral. We note too that both models are very similar to each other in this regard. The largest decreases in long wave kinetic energy in the models occur in the upper troposphere at 30 N and at 47 N.

Fig. 11.4.1b shows the latitude height distribution of the long wave kinetic energy in the observed and forecast ensemble average fields, i.e. the S fields. Figs. 11.4.1a and b bear very marked similarities to each other both for the observations and for each of the models. Both observed fields show upper tropospheric maxima near 30 N and near 45 N. Moreover the differences between forecast and observed fields are very similar for both the S field and the total field.

A comparison of Figs. 11.4.1a and 11.4.1b strongly suggests that the major error in the models long wave kinetic energy is the error in the S part of the kinetic energy, particularly at high levels. For example at 300 mb, using the units of Fig. 11.4.1, the observed meridionally-averaged long wave kinetic energy is composed of 63 units of S energy and 53 units of T energy. For the EC model the corresponding numbers are 39 units of S energy and 55 units of T energy. The results for the GFD model are very similar.

Fig. 11.4.2 a,b shows the long wave momentum flux $\overline{u'v'}$ for observations and models for both the total field and the S part of the flow. The observations show that the long wave momentum flux is about half the total momentum flux (cf Fig. 11.3.4) and that the long wave flux is dominated by the stationary part. The models fail to capture the structure of the field; the forecast maximum occurs between 20 N and 30 N instead of between 40 N and 50 N as observed, and this major failure occurs because both models fail to predict the stationary part. If we estimate the T ("transient") part of the term from Fig. 11.4.2a and b we see that in the observations and in the GFD model it is poleward between 60 N and 80 N while in the ECM model it is equatorward.

The effect of differences in the zonal flow and differences in the momentum fluxes on the energy conversion term CK in wavenumbers 1-3 is shown in Fig. 11.4.3. This presents the latitude-height distribution of the total term and of the S component. The differences are largely confined to the south side of the jet maximum and result in the term being much too negative there.

Fig. 11.4.4a show results corresponding to Fig. 11.4.1 for the eddy available energy in wavenumbers 1-3. The models produce a reasonable structure for this field, particularly the EC model. Fig. 11.4.4b shows that a substantial part of this term is in the S field as we have already seen for the kinetic energy.

Fig. 11.4.5 shows the contribution of the long waves to the heat flux $\overline{v'T'}$ for both the total term and the S part of the term for the same area and time period. This term is a substantial contributor to the total heat flux (cf Fig. 11.3.6). For the S part of the term the models are not successful in reproducing the observed structure in that they show a relative minimum in mid levels instead of showing the minimum at 200 mb. If we compare the S field with the total field it is clear that the large values at 200 mb and 850 mb between 50 N and 70 N must be due to the "transient" waves.

In summary then, the main points to be made about the contribution of the long waves to the budgets are that i) the "stationary" long waves are too weak and do not transport enough heat or momentum polewards, especially north of 45 N. ii) the "transient" long waves transport too much heat polewards at 200 mb between 60 N and 80 N.

iii) the "transient" long waves transport momentum equatorwards instead of polewards in the ECM model between 60 N and 80 N. The errors in the "Stationary" long waves are serious because they dominate the long wave structure of the atmosphere and this is the part of the flow we would most like to forecast.

11.5 Latitude-height cross sections of budget calculations for the medium waves (zonal wavenumbers 4-9)

Fig. 11.5.1 shows the observed and forecast latitude height distribution of medium wave kinetic energy averaged over the seven cases for the last three days of the forecasts. The forecasts are quite successful for this quantity in the upper troposphere. In the lower troposphere both models overpredict it by approximately 65%. The contribution of the S fields to this quantity is small in both the forecasts and in the observations.

Fig. 11.5.2 shows the momentum fluxes in this wave band. In the region of the main maximum the GFD model produces a better forecast than the ECM model which produces too large and too extensive fluxes. One may speculate that this difference is due to the difference in the interaction of the dynamics with the convection schemes since the Kuo scheme tends to affect higher levels than the Manabe scheme. In both models the momentum convergence is larger than observed between 40 and 50 N, a feature which was also noted for the long waves. This is particularly true for the EC model where the momentum divergence at high levels in the subtropics is also much stronger than either observations or the GFD model.

Just as for the long waves, the contribution of the medium waves to the term CK is too negative, as may be seen in Fig. 11.5.3 where the largest errors are again in the ECM model. This is due not only to the over prediction of the momentum fluxes but also, for the ECM model, to the error in the position of the jet.

Turning now to the temperature field we see in Fig. 11.5.4. that the contribution of the medium waves to the EAPE is reasonably well forecast although in both models the contribution is too large. The horizontal heat flux (Fig. 11.5.5) is also too large in both models in this wave band. This is true at all levels but is particularly marked in the GFD model in the lower troposphere. Similar comments may be made about the medium wave contribution to the term CA (Fig. 11.5.6).

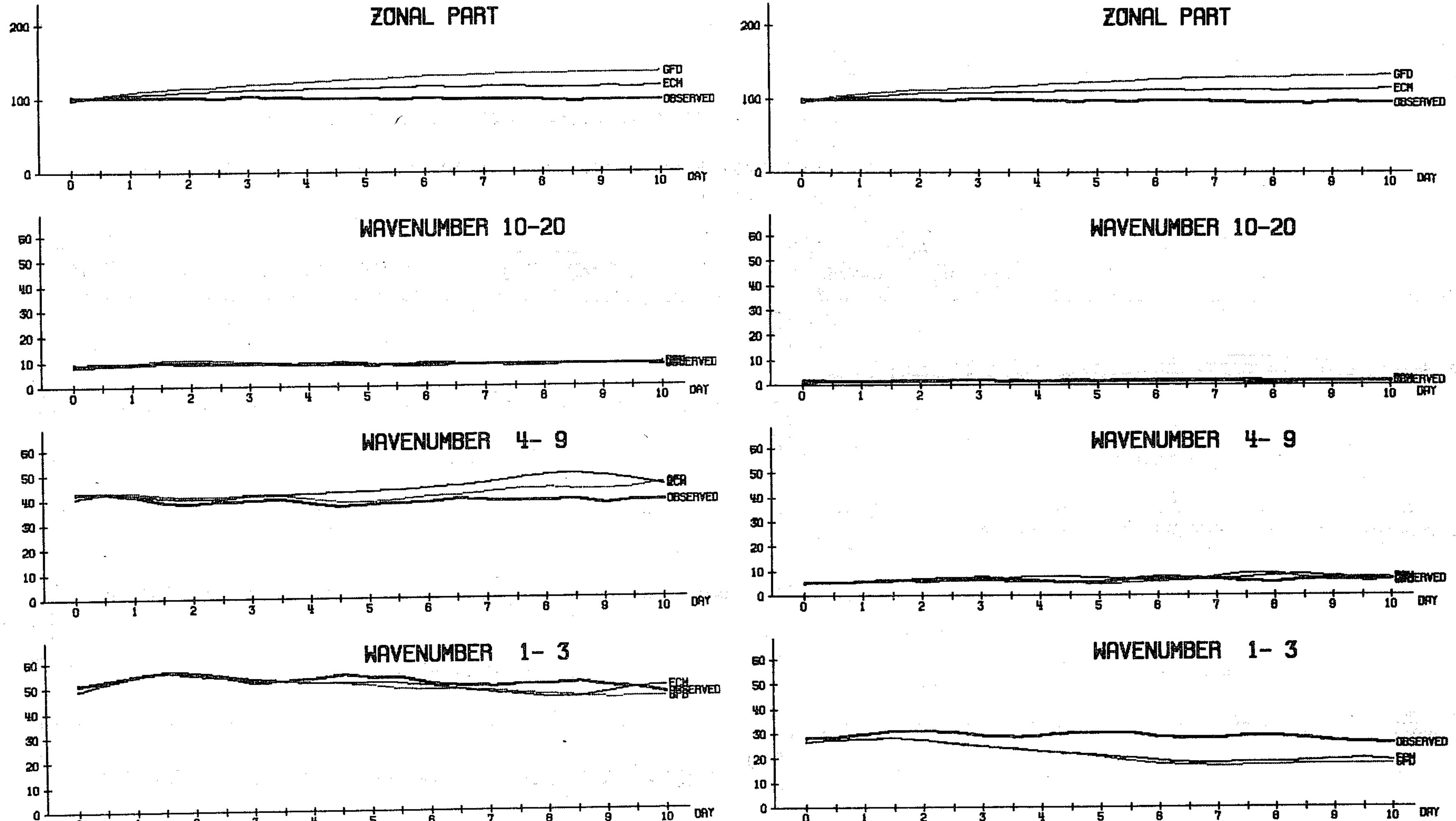
11.6 Effects of the parameterization schemes

The fields which we have studied in this chapter seem to show more sensitivity to differences in the parameterization schemes than those studied in earlier chapters. Broadly speaking the structure of the major errors are similar for both models, in particular the acceleration of the zonal flow in mid latitudes and deceleration at high latitudes. Nevertheless there appear to be important differences which may be attributed to the parameterization schemes such as the structure of the mean meridional circulation, the transient long wave momentum flux at high levels north of 60 N, the subtropical momentum fluxes due to the medium waves and the lower tropospheric sensible heat fluxes in the same medium wave band.

We have also compared the model's energetics with the analyses from the Deutscher Wetterdienst. Although there are some marked differences between the energetics of the two analyses (Arpe 1979) the only conclusions presented above that need to be modified are that the forecast minus observed differences for APE and CA are no larger than the differences between the two analyses.

As was pointed out in the discussion of the systematic errors, further work is required to understand the reasons for the weakness of the "stationary" long waves in the forecasts. The material presented here has documented further this aspect of the models behaviour and also shown clearly the excessive vigour of the medium range waves especially in their associated heat transports.

Continued on page 207

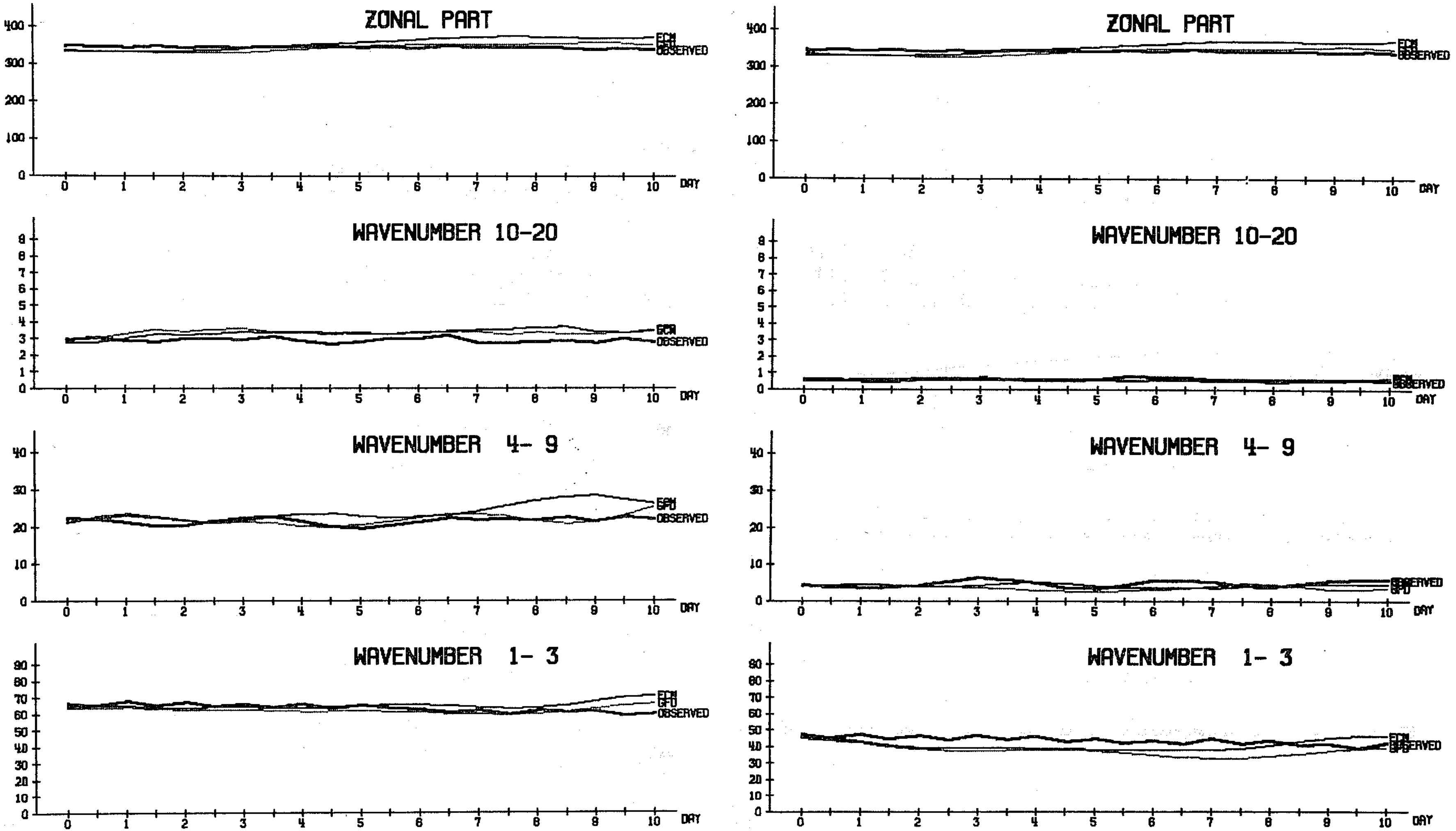


a) INTEGRAL 1000- 200 MB AREA MEAN 20.0- 82.5 N
KE (10 KJ/M2) 7 CASES

b) INTEGRAL 1000- 200 MB AREA MEAN 20.0- 82.5 N
KE (10 KJ/M2)

Fig.11.1.1 a) Time evolution of the kinetic energy in the extra tropical troposphere averaged over the ensemble for the indicated wavenumber groups. The heavy line is for the observations, the thin line for the ECM model and the dotted line is for the GFD model.

b) As a) for the kinetic energy in the ensemble means of the observations and of the forecasts.



a) INTEGRAL 850- 200 MB AREA MEAN 20.0- 82.5 N
AE (10 KJ/M2) 7 CASES

b) INTEGRAL 850- 200 MB AREA MEAN 20.0- 82.5 N
AE (10 KJ/M2)

Fig.11.1.2 a) As Fig. 11.1.1 a for the available potential energy

b) As Fig. 11.1.1 b for the available potential energy

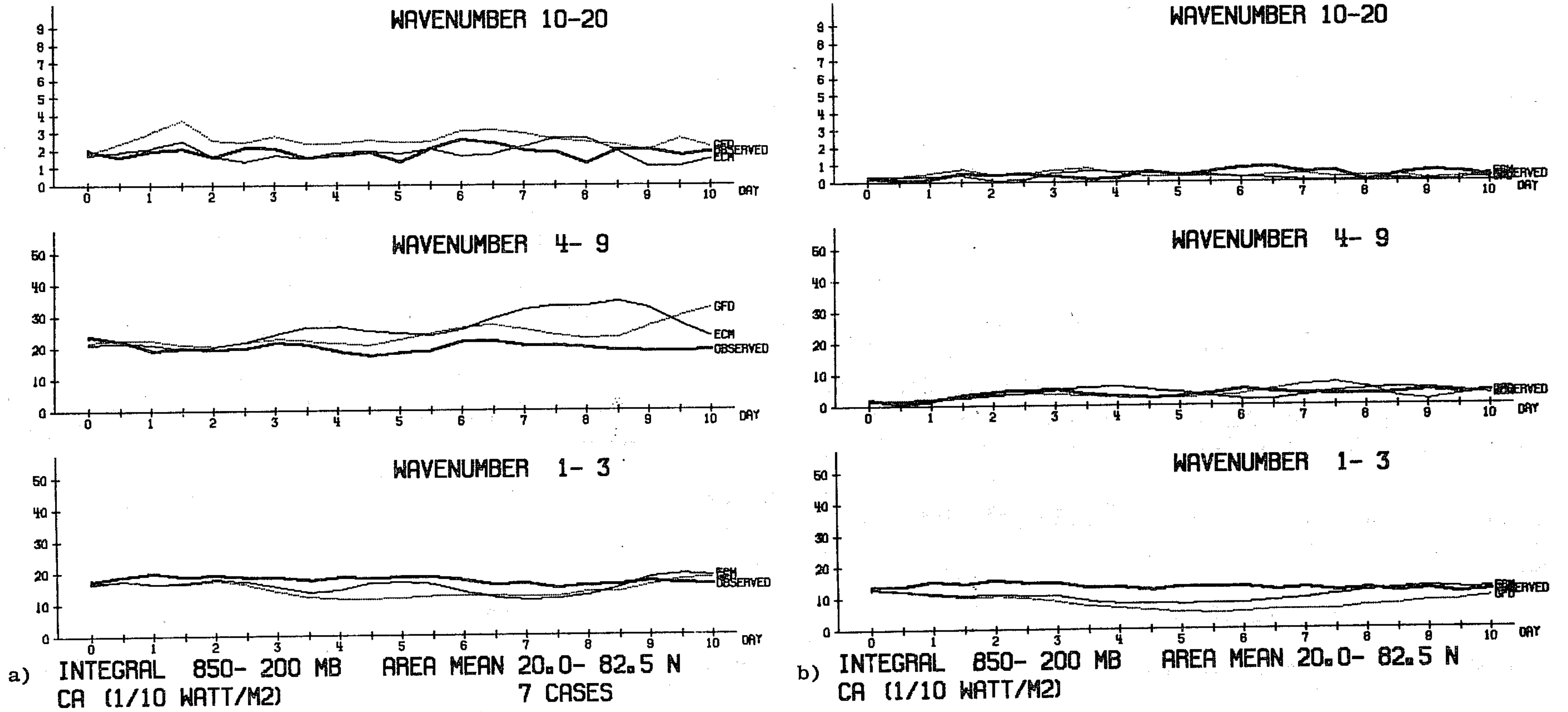


Fig. 11.1.3 a) As Fig. 11.1.1 a for the rate of conversion (CA) of zonal available potential energy to eddy available potential energy.

b) As Fig. 11.1.1 b for CA

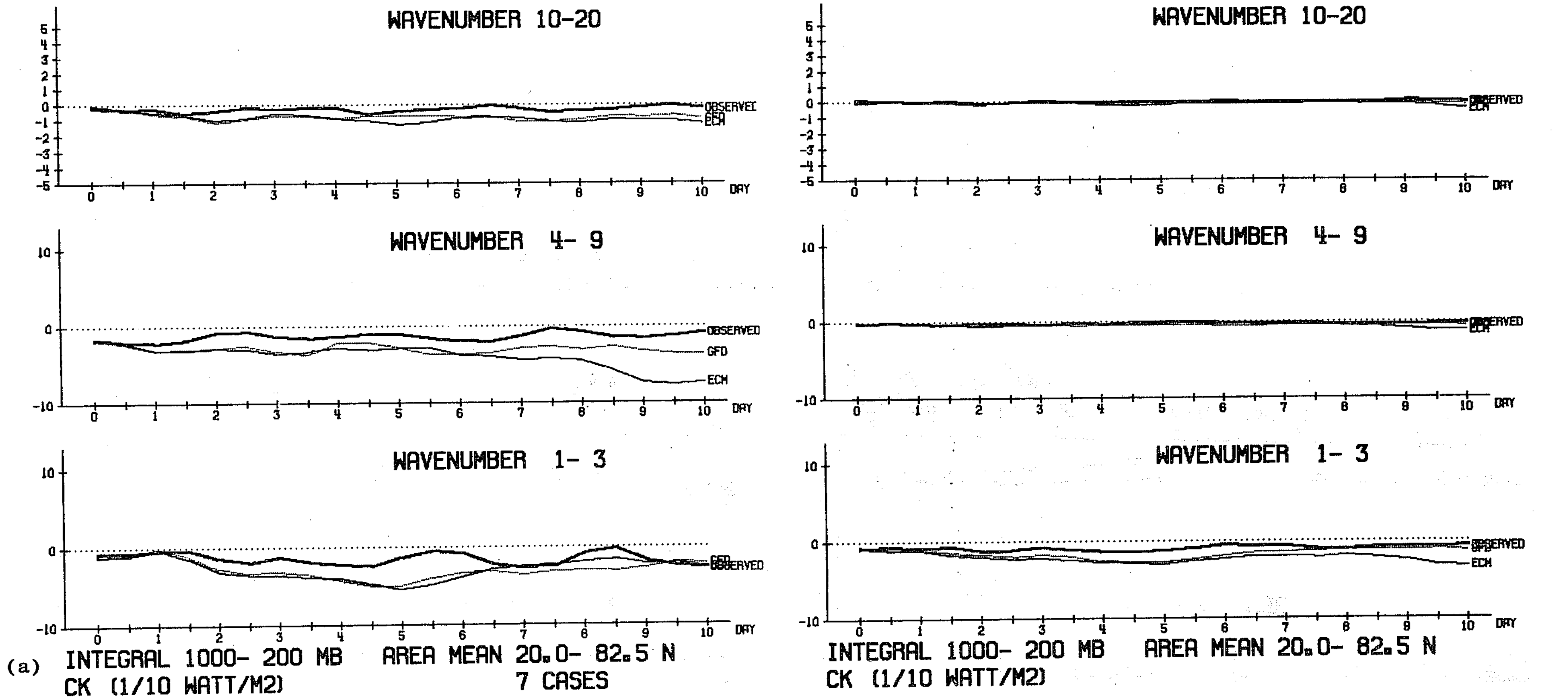
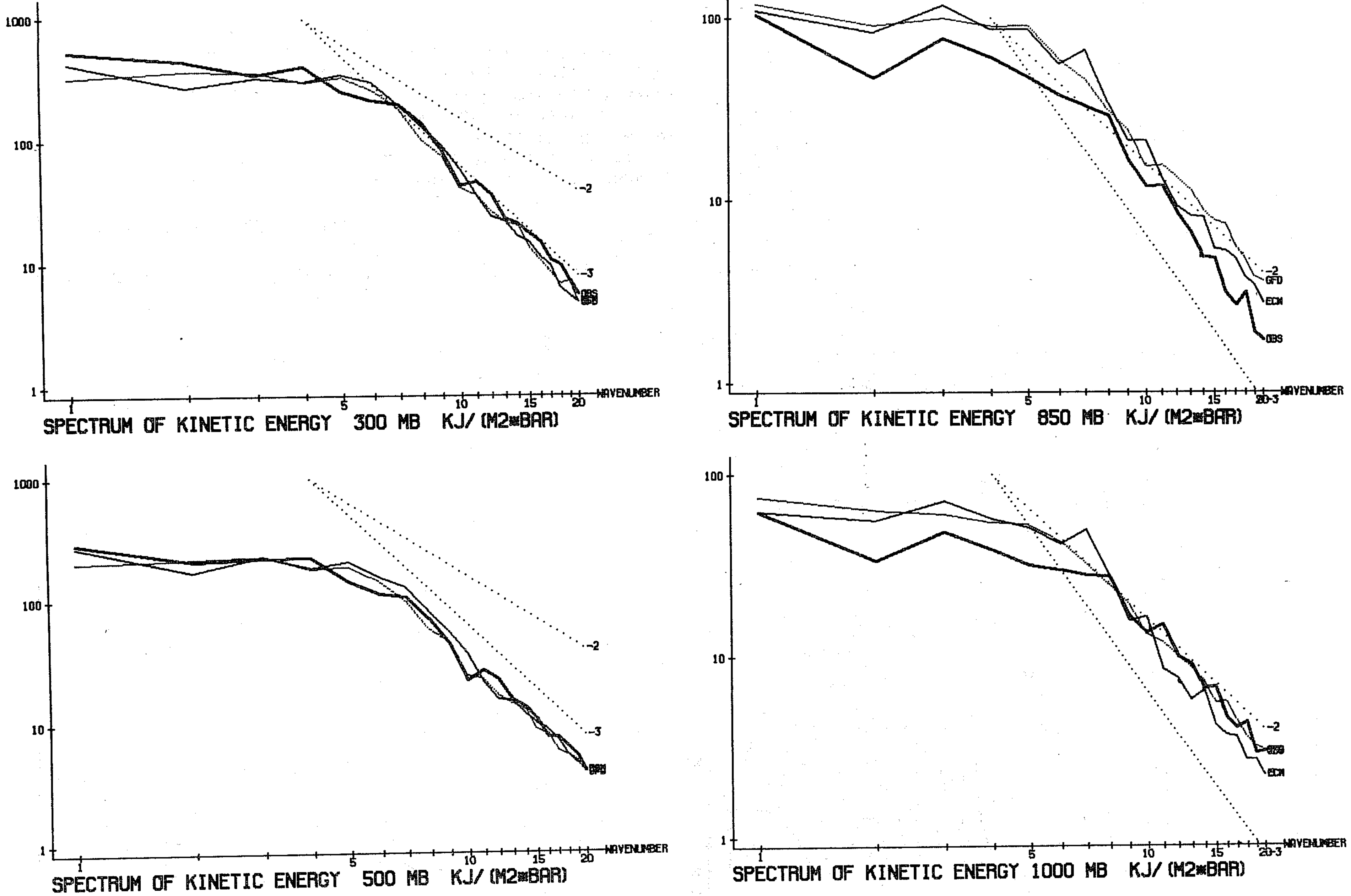


Fig. 11.1.4 a) As Fig. 11.1.1 a for the rate of conversion (CK) of zonal kinetic energy to eddy kinetic energy.

b) As Fig. 11.1.1 b for CK.



SPECTRUM OF KINETIC ENERGY 300 MB KJ/(M2*BAR)

SPECTRUM OF KINETIC ENERGY 850 MB KJ/(M2*BAR)

SPECTRUM OF KINETIC ENERGY 500 MB KJ/(M2*BAR)

SPECTRUM OF KINETIC ENERGY 1000 MB KJ/(M2*BAR)

Fig. 11.2.1 Spectra of kinetic energy at 1000, 850, 500, 300 mb averaged between 40 N and 60 N, averaged over the ensembles and averaged over days 7 to 10, for the observations (thick line) for the ECM model (thin line) and for the GFD model. Note the change of scale between the upper tropospheric spectra (left) and lower tropospheric spectra (right). Also shown are the slopes corresponding to k^{-2} and k^{-3} spectra.

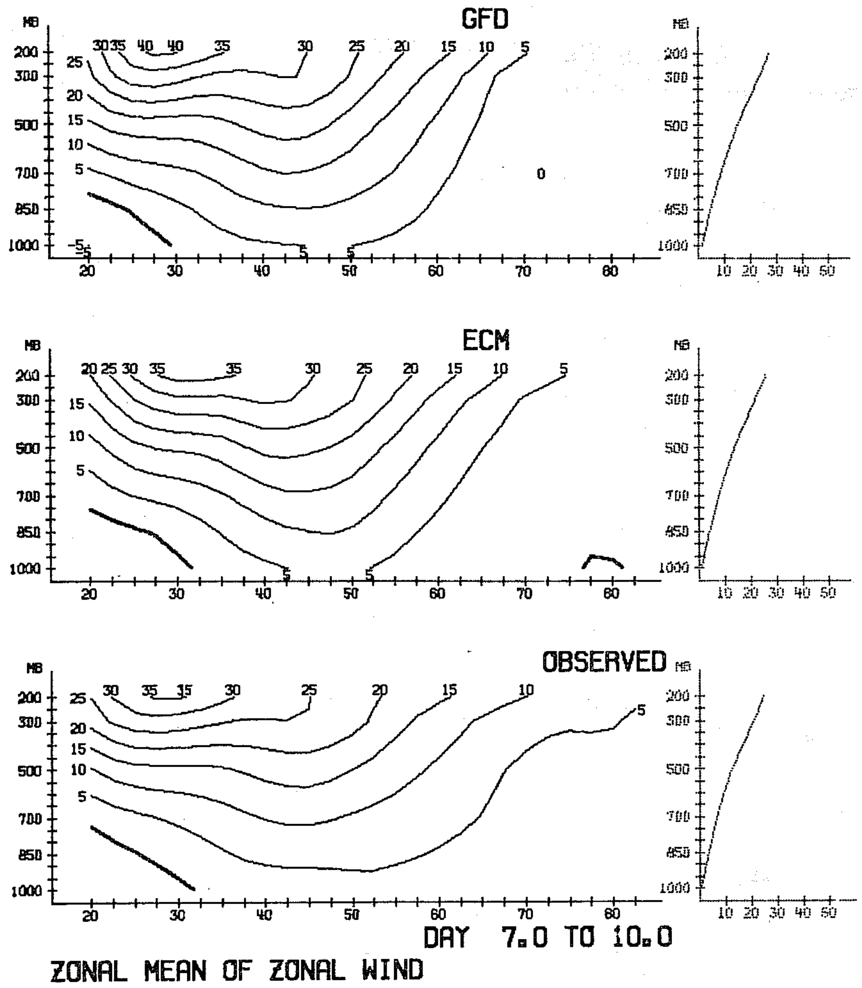


Fig.11.3.1 Latitude height distribution of the average over the ensemble and over days 7 to 10 of the zonal mean of zonal wind for the GFD model (top) ECM model (centre) and observations (bottom). The units are ms^{-1}. The small panels to the right show the vertical profile of the area-weighted meridional integral of the quantity in the main panel.

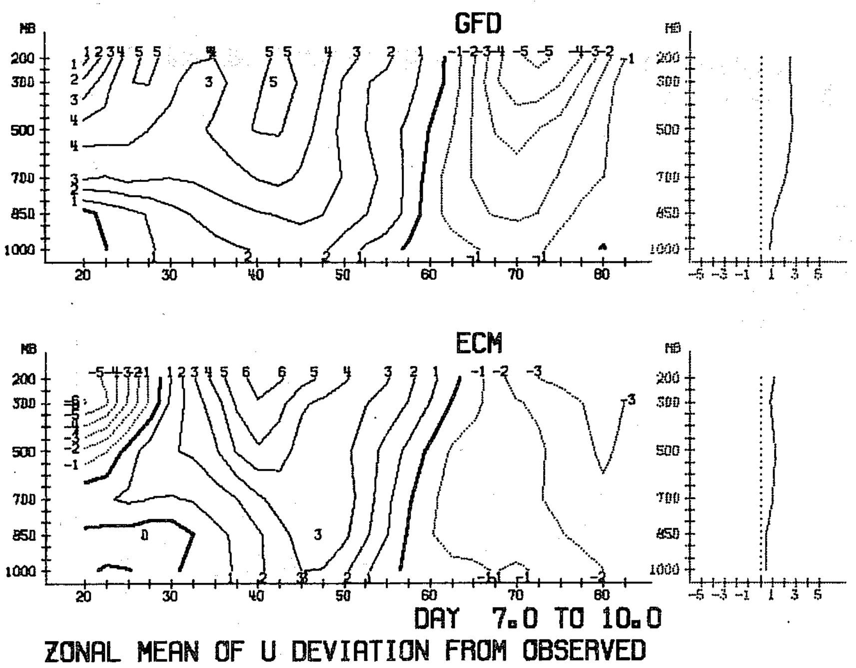


Fig. 11.3.2 Departure of forecast from observation of the zonal wind data shown in Fig. 11.3.1. The vertical profiles of the area weighted mean are shown on the left.

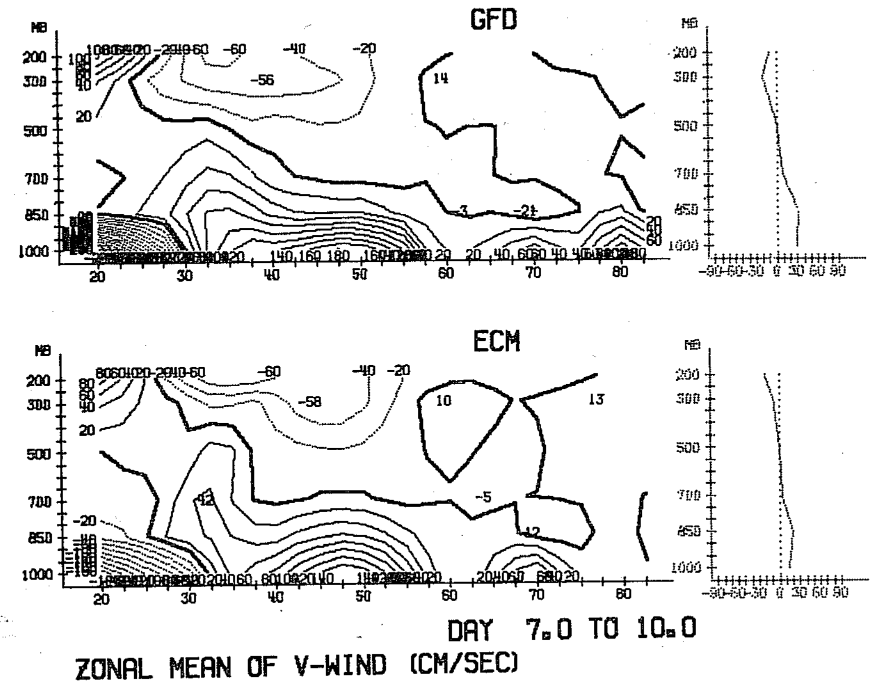


Fig. 11.3.3 As Fig.11.3.1 for the zonal mean of the meridional velocity. No observational data are available. The units are cm s^{-1}

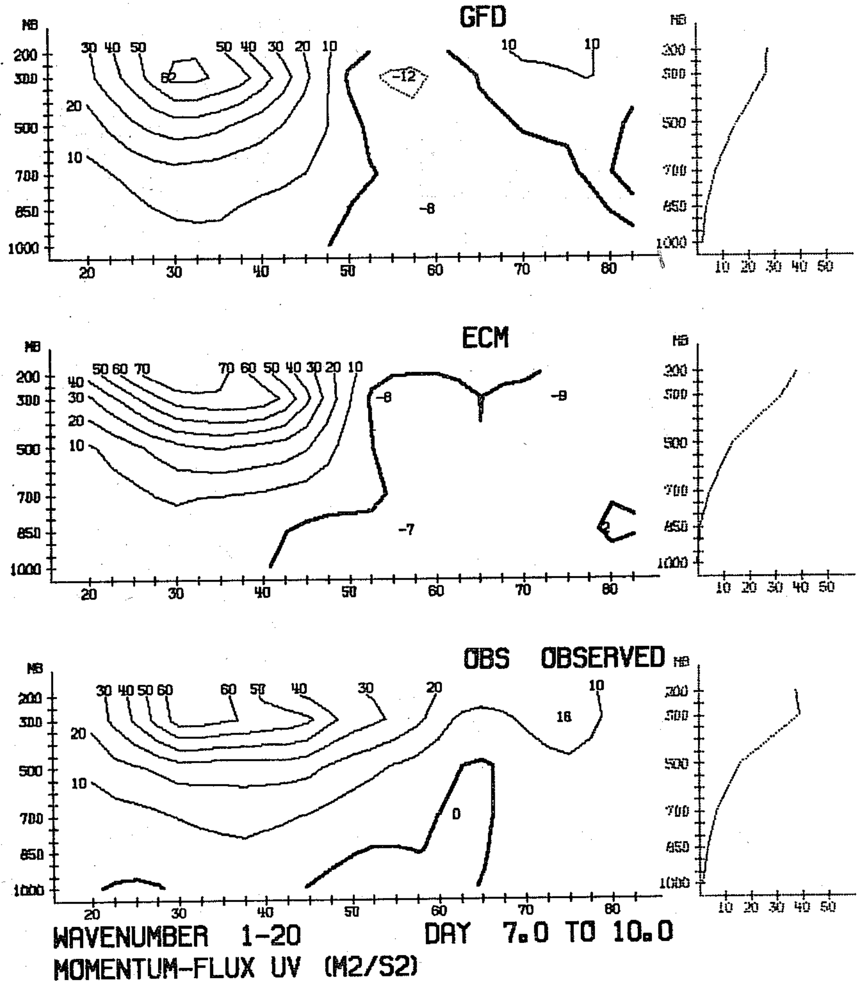


Fig.11.3.4 As Fig.11.3.1 for the momentum flux $\overline{u'v'}$ in the zonal wavenumber band 1-20. The units are m^2 s^{-2}.

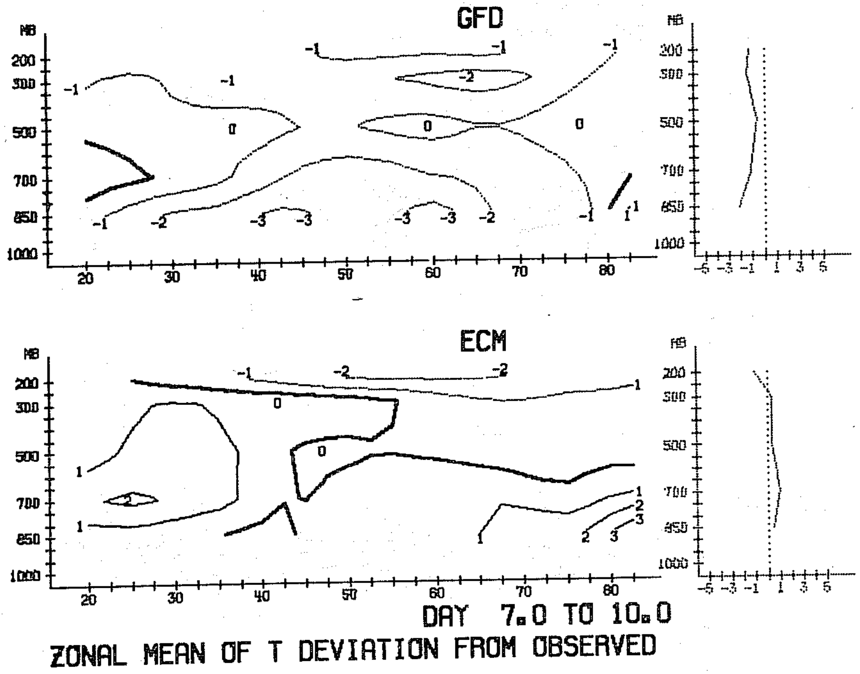


Fig. 11.3.5 As Fig. 11.3.2 for the difference between the forecast and observed zonal mean of temperature.

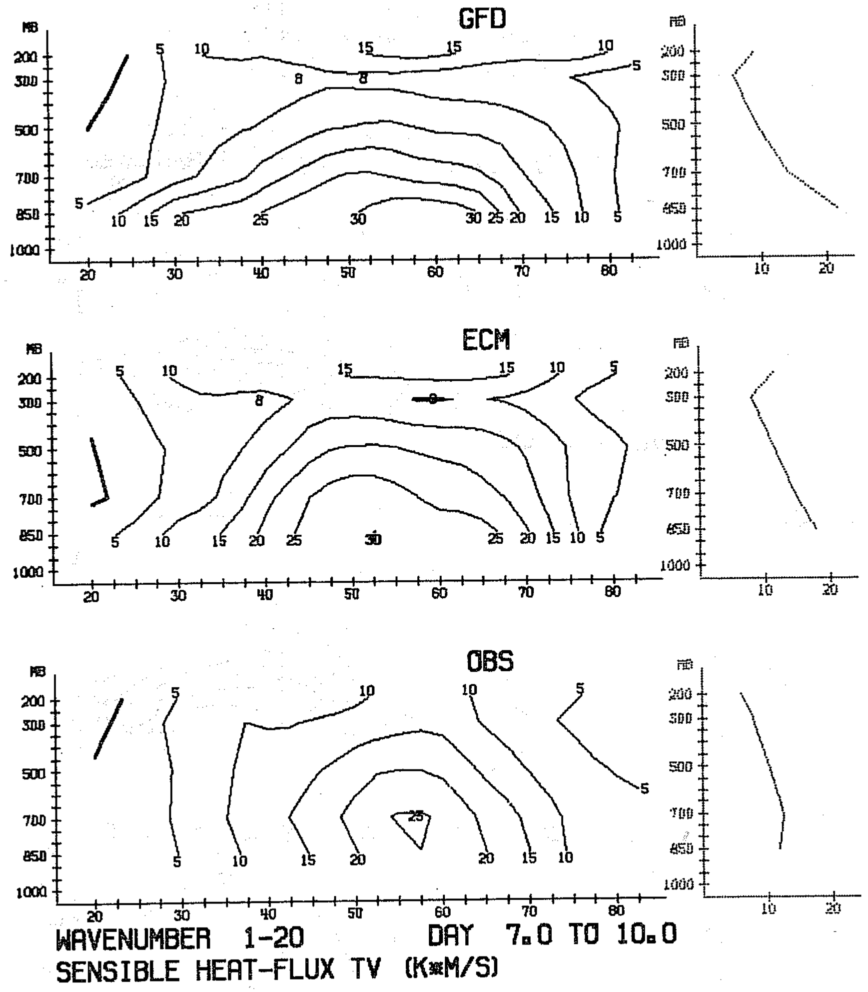


Fig. 11.3.6 As Fig.11.3.1 for the covariance of V and T (which is proportional to the sensible heat flux). The units are Km s^{-1}.

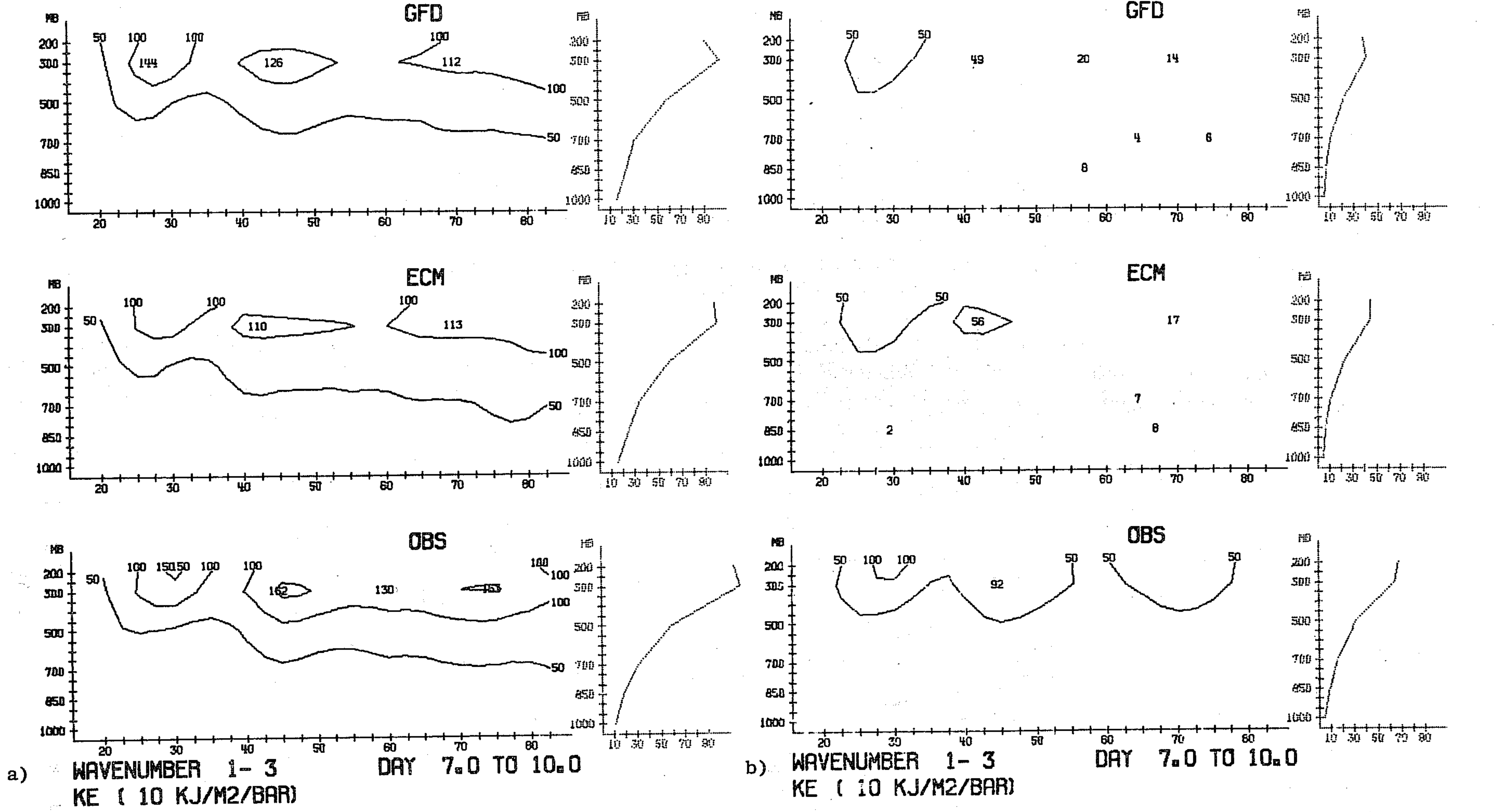
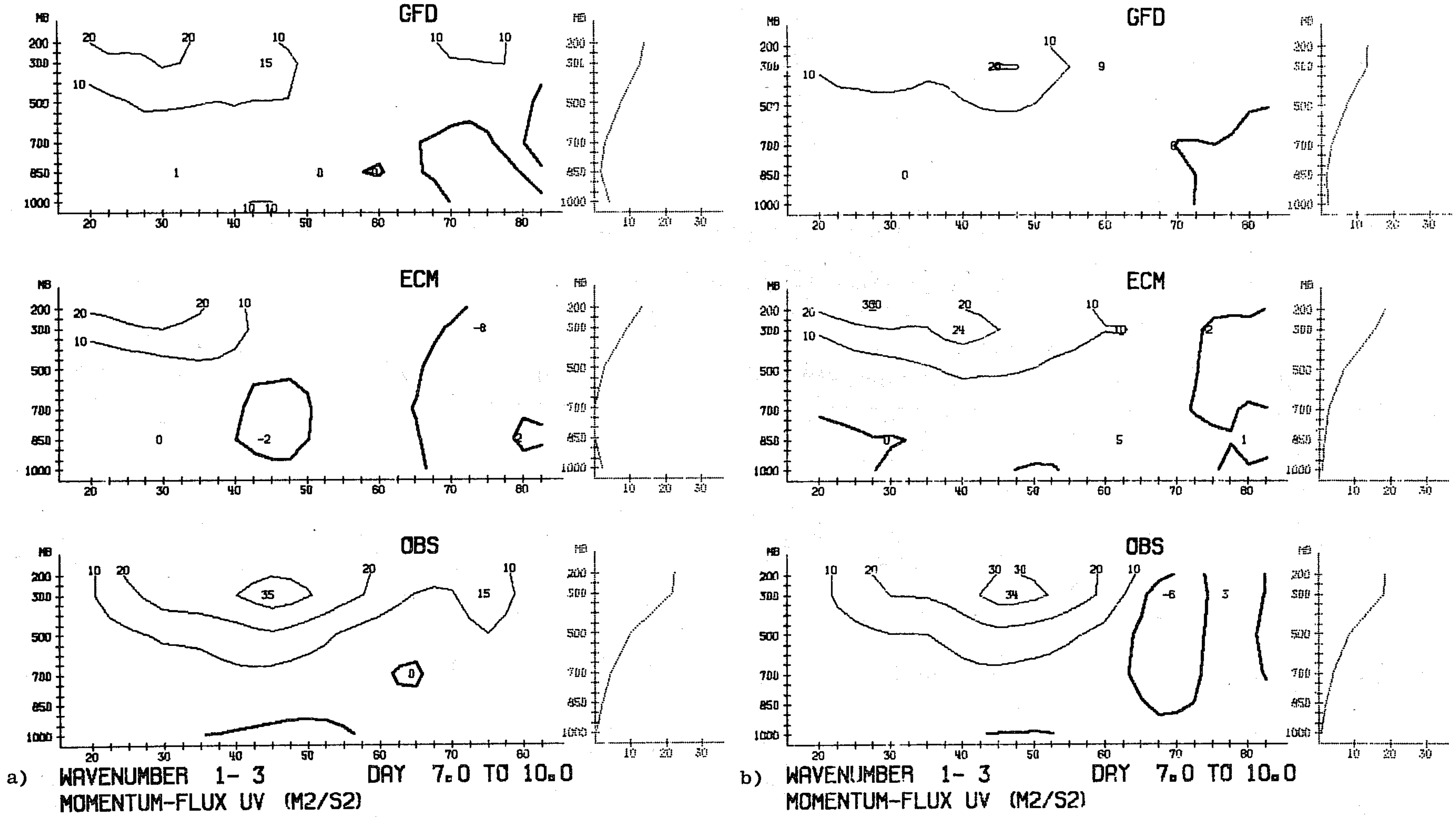


Fig. 11.4.1 a) As Fig. 11.3.1 for the kinetic energy in the zonal wavenumber group 1-3, the long waves.

b) As a) for the kinetic energy in the ensemble averaged fields



a) WAVENUMBER 1- 3 DAY 7.0 TO 10.0
MOMENTUM-FLUX UV (M2/S2)

b) WAVENUMBER 1- 3 DAY 7.0 TO 10.0
MOMENTUM-FLUX UV (M2/S2)

Fig. 11.4.2 a) As Fig. 11.3.1 for the momenum flux in the zonal wavenumber group 1-3, the long waves.

b) As a) for the long wave momentum flux in the ensemble averaged fields.

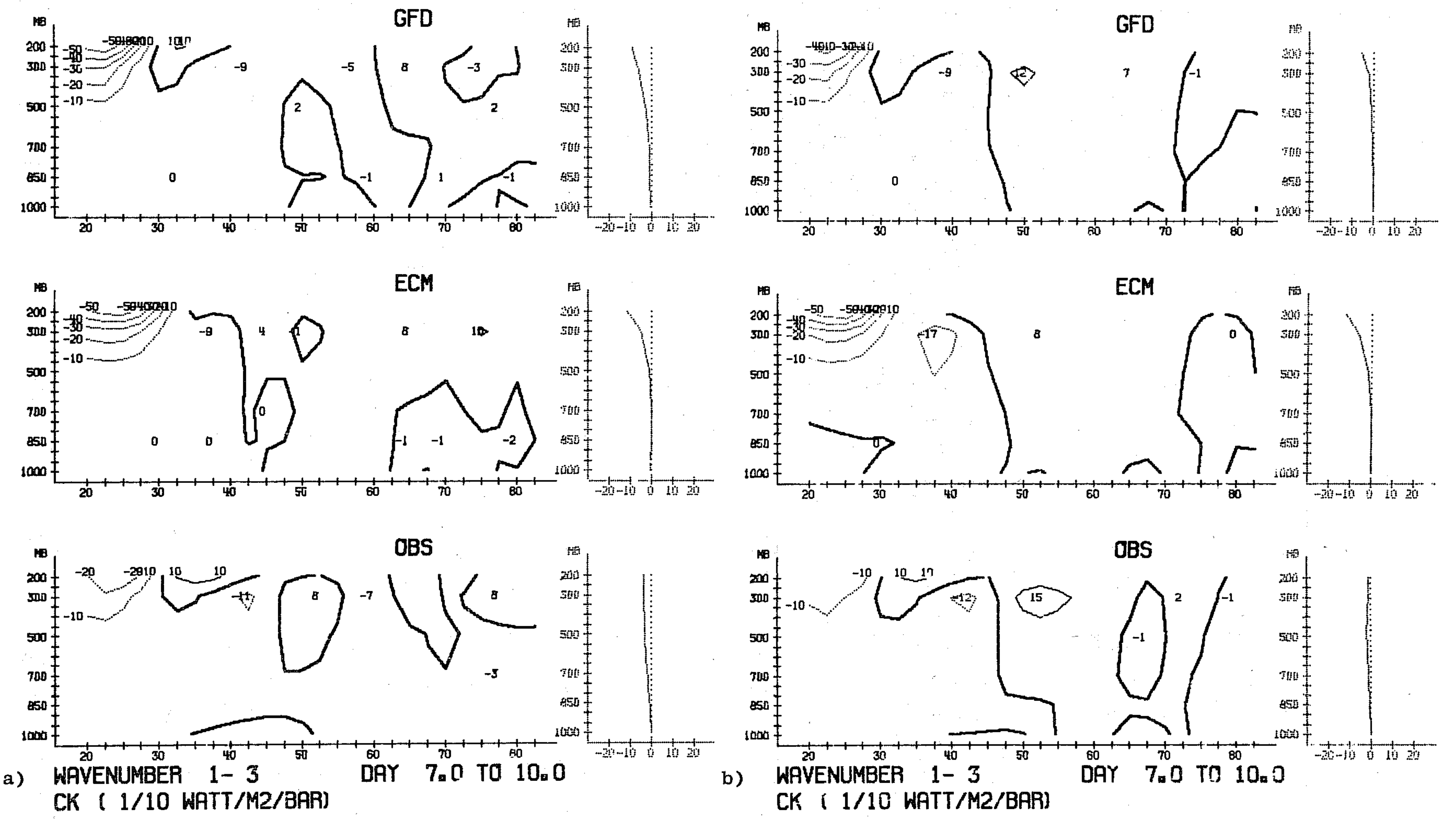
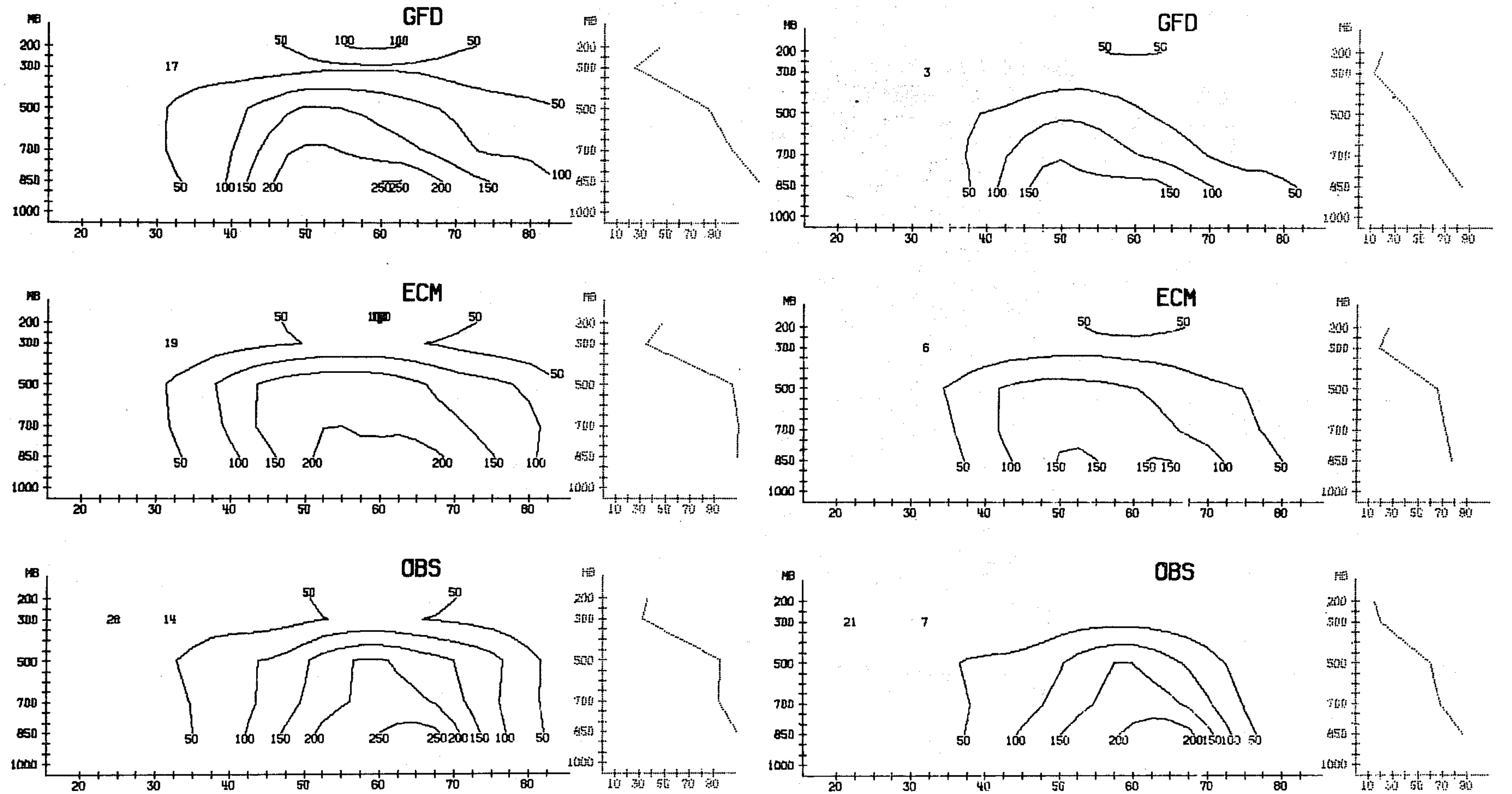


Fig. 11.4.3 a) As Fig. 11.3.1 for the conversion rate from zonal to eddy kinetic energy (CK) in the long waves.

b) As a) for CK in the ensemble averaged fields.

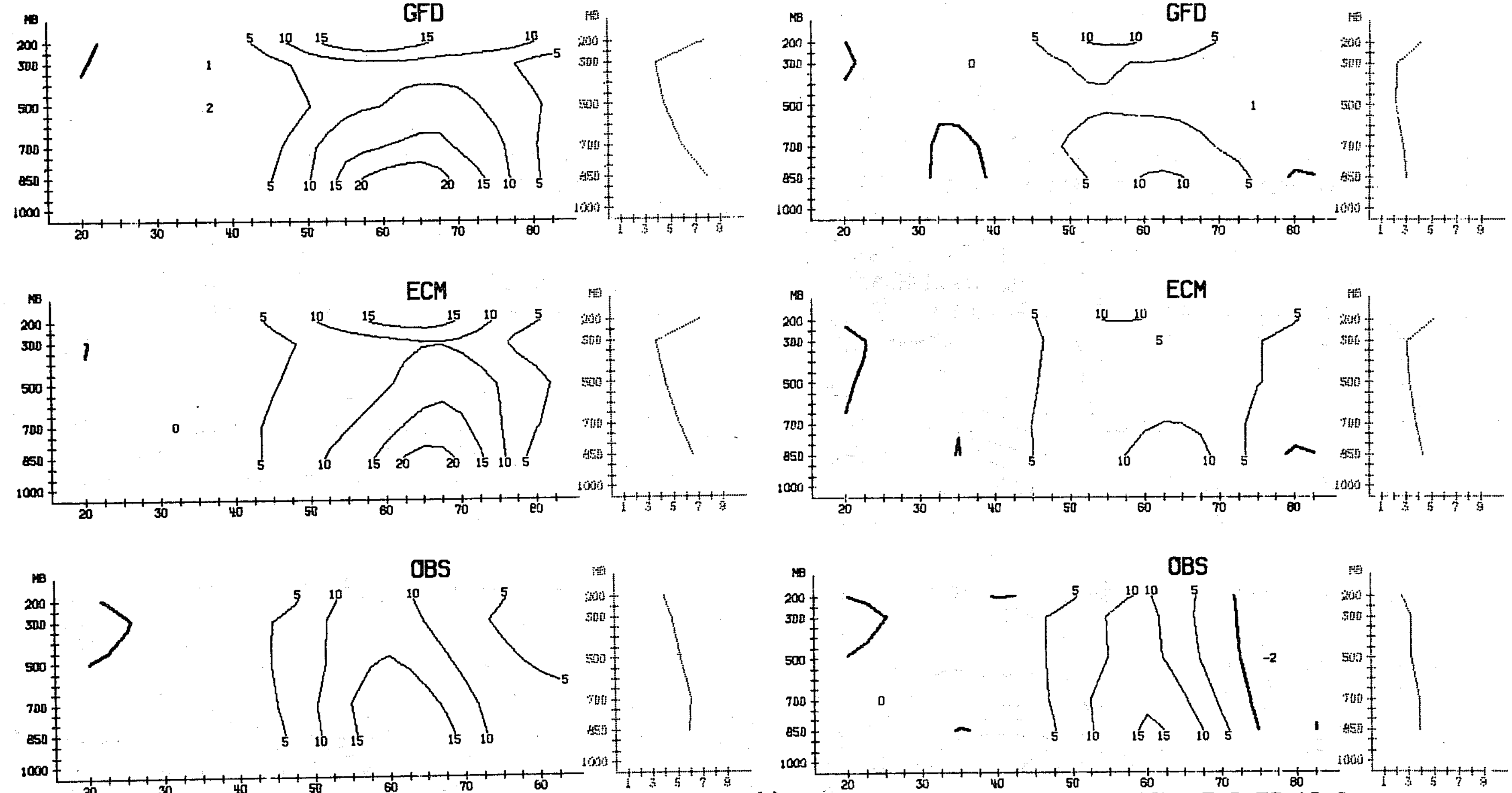


a) WAVENUMBER 1- 3 DAY 7.0 TO 10.0
AE (10 KJ/M2/BAR)

b) WAVENUMBER 1- 3 DAY 7.0 TO 10.0
AE (10 KJ/M2/BAR)

Fig. 11.4.4 a) As Fig. 11.3.1 for the eddy available energy (AE) in the long waves (zonal wavenumbers 1-3).

b) As a) for AE in the ensemble averaged fields.



a) WAVENUMBER 1- 3 DAY 7.0 TO 10.0
SENSIBLE HEAT-FLUX TV (K*M/S)

b) WAVENUMBER 1- 3 DAY 7.0 TO 10.0
SENSIBLE HEAT-FLUX TV (K*M/S)

Fig. 11.4.5 a) As Fig. 11.3.1 for the covariance of V and T (which is proportional to the sensible heat flux in the long waves.

b) As a) for the covariance in the ensemble averaged fields.

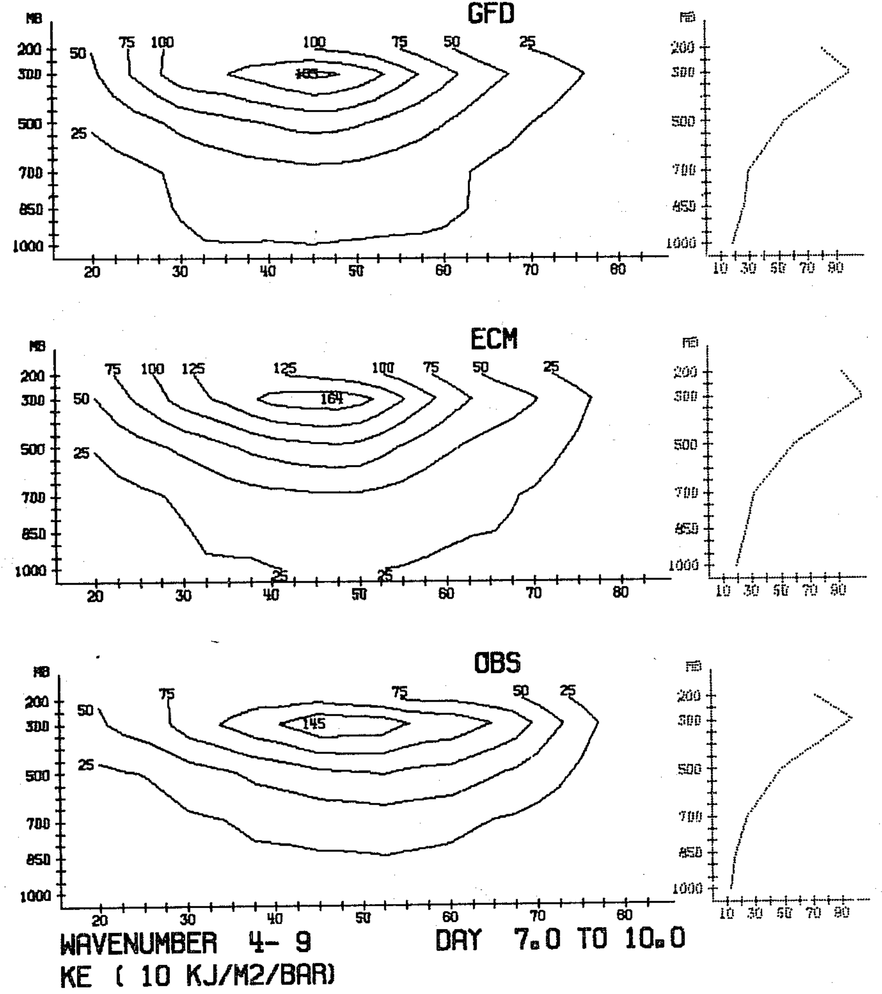


Fig. 11.5.1 As Fig. 11.3.1 for the kinetic energy in the wavenumber group 4-9.

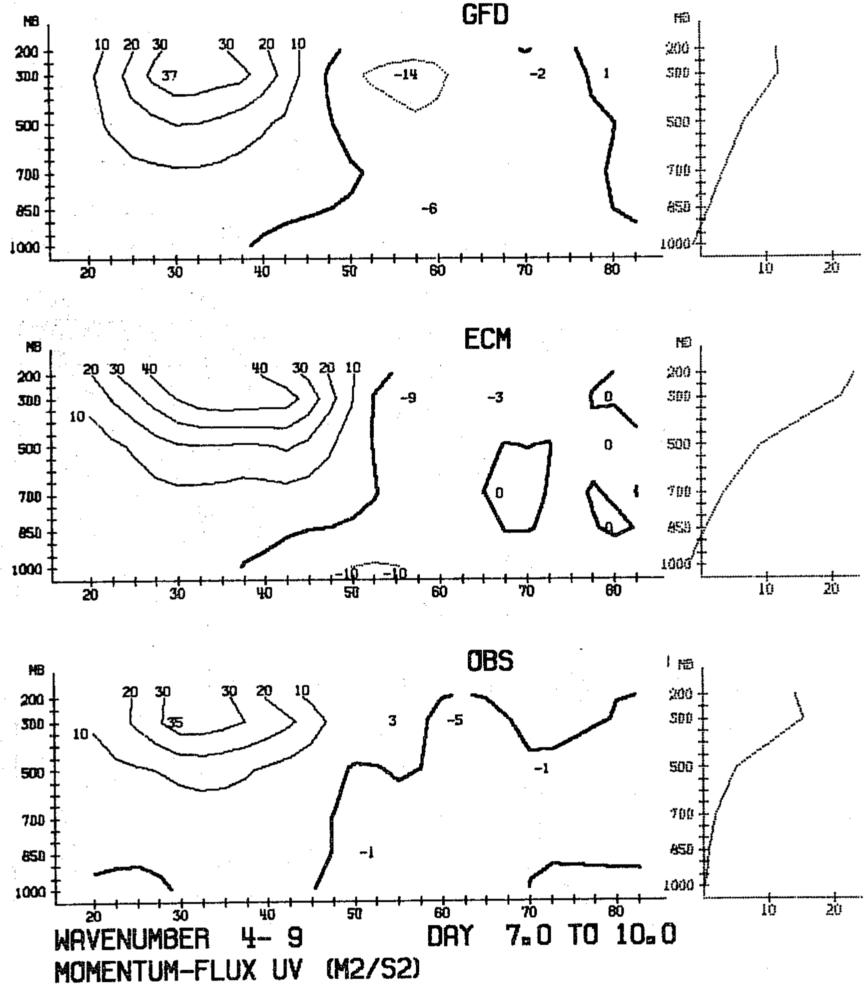


Fig. 11.5.2 As Fig. 11.3.1 for the momentum flux in the medium waves (zonal wavenumber 4-9).

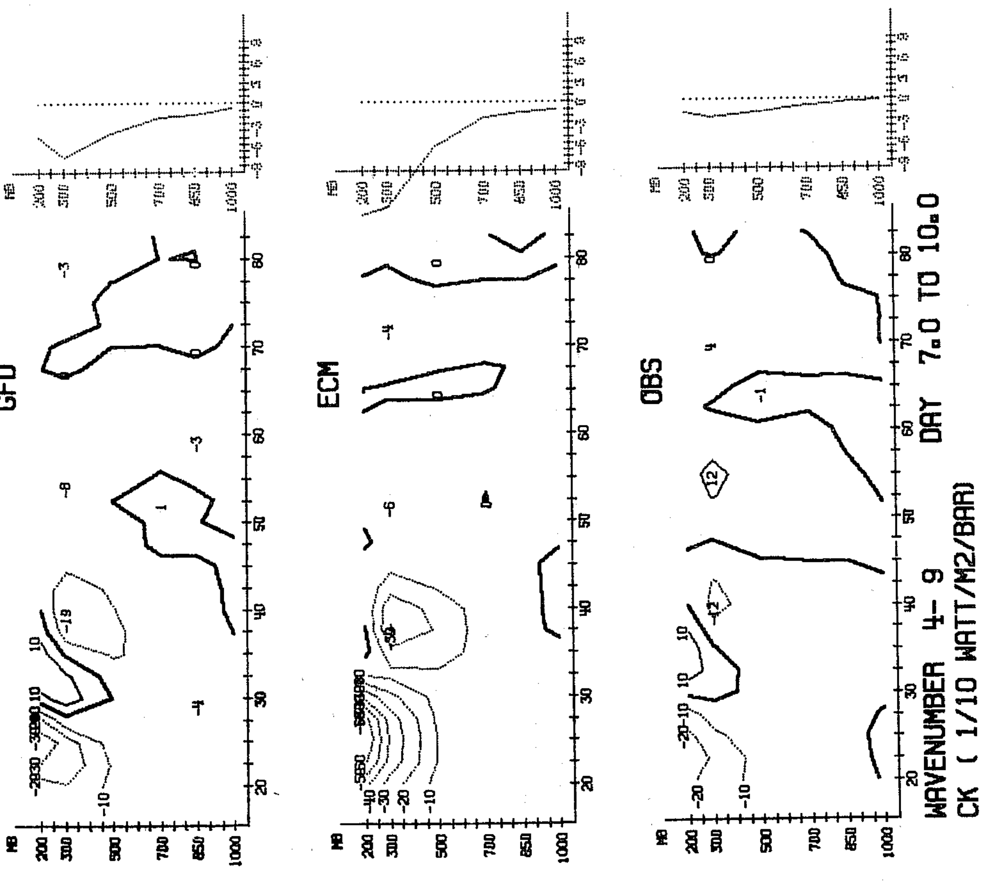


Fig. 11.5.3 As Fig. 11.3.1 for the contribution of the medium waves to the term CK.

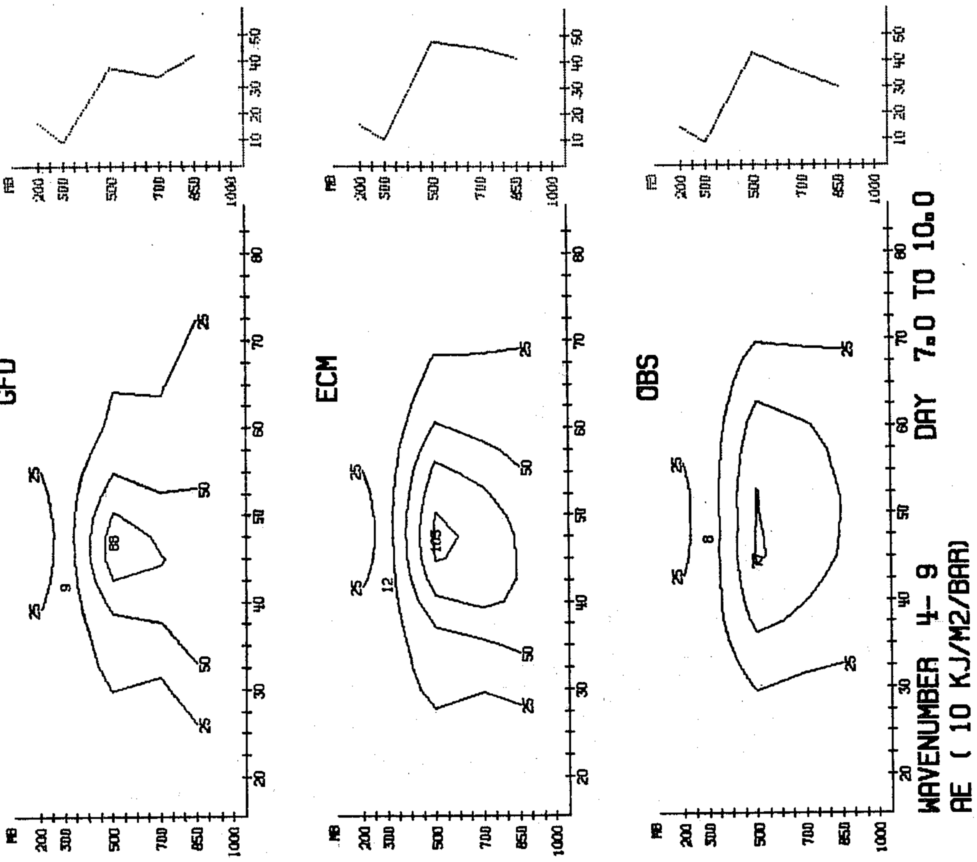


Fig. 11.5.4 As Fig. 11.3.1 for the eddy available energy in the medium waves.

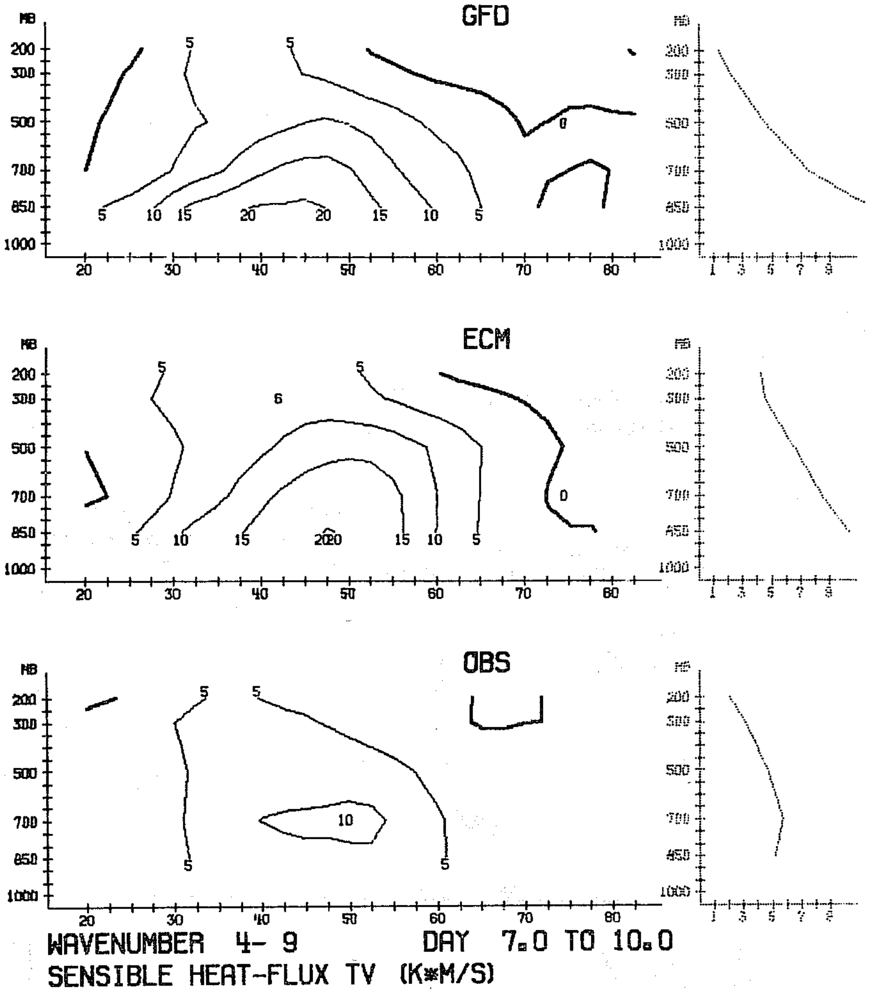


Fig. 11.5.5 As Fig. 11.3.1 for the covariance of V and T in the medium waves.

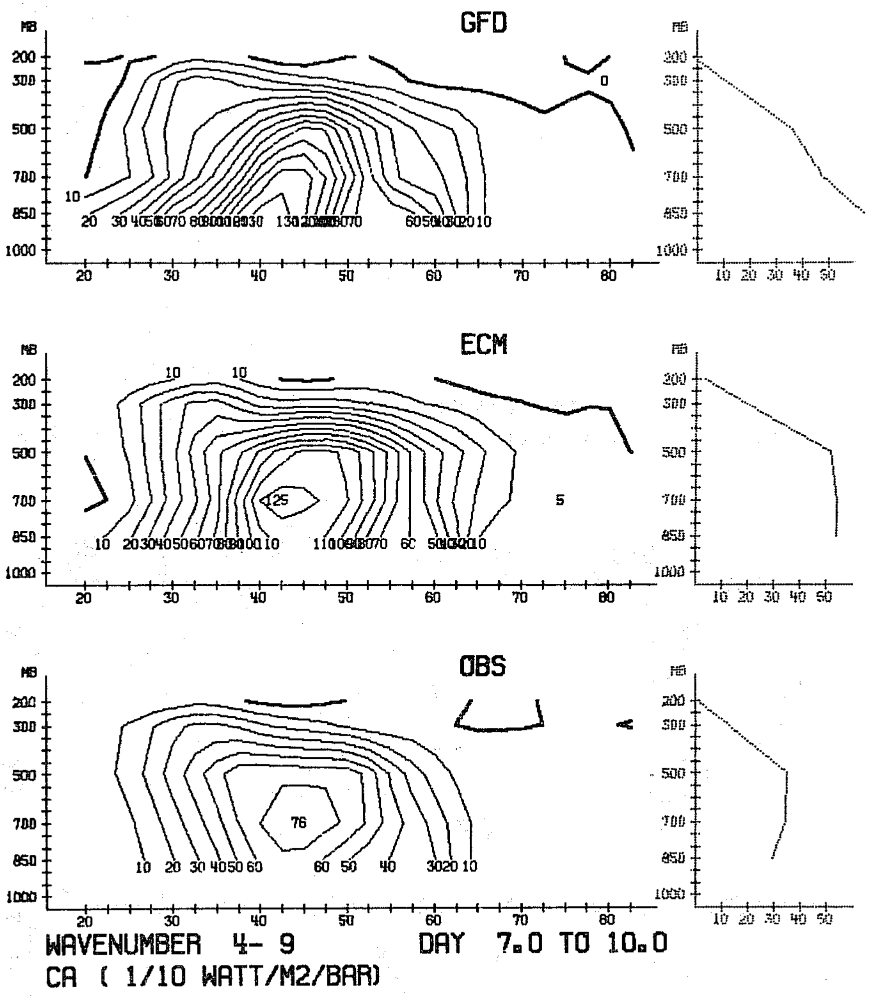


Fig. 11.5.6 As Fig. 11.3.1 for the contribution of the medium waves to the conversion CA.

12. Discussion and outlook

We have presented results of two sets of winter forecasts made with two models differing only in their physical parameterizations. One set of physical parameterizations has been in use in general circulation and forecasting experiments for 15 years while the second set relaxed some of the restrictions of the first and incorporated several new developments.

The forecasts made with the models indicate that we have predictability for five to six days in winter. Up to the point where the models lose predictive value there is little difference between them in forecast quality. Moreover the systematic errors in each model are very similar.

The stationary systematic errors are a substantial element of the total error. Three likely sources for this error have been identified, over-intense baroclinic storms, errors in the direct forcing of long waves and errors in the treatment of some features of the tropical circulation. There is every reason to expect that these errors can be reduced as a result of further work. Crude arguments indicate that the elimination of these errors would bring about a 20% gain in predictability. Such a gain would be substantial progress in extending the range of practical predictability.

A study similar to this for August 1975, which will be the subject of a later report, shows similar results as regards forecast quality. We have, then, grounds for expecting that the present results will not be unrepresentative of the model's performance in operations. This depends of course on the operationally available data being of comparable quality to the DST data. This should be the case during the Global Weather Experiment which ends in December 1979. The availability of the space-based observing systems thereafter should help ensure the maintenance of data quality.

Acknowledgements

We are grateful to L. Bengtsson, D.M. Burridge, J.F.-Louis and J.F.-Geleyn for many helpful discussions in the course of this work. The practical aspects were capably handled by J.R. Gibson, U.Pilz, C.Clarke and W.Loechle. D.Dent provided much helpful advice on the operating systems. The manuscript was prepared by Mrs I. Rhodes and Mrs. M. Foster-Moore.

References

Alexander, R.C. and Mobley, R.L. 1974 "Monthly Average Sea-Surface Temperatures and Ice-pack Limits on a 1^{o} Global Grid" RAND Report R-1310-ARPA.

Arakawa, A. and Lamb, V. 1977 "Computational design of the basic dynamical processes of the UCLA general circulation Model" in "Methods of Computational Physics", Vol.17, Ed. J. Charney, Academic Press.

Arpe, K. 1979 "Confidence limits for verification and energetical studies", to be published.

Asselin, R. 1972 "Frequency filter for time integrations", Mon. Wea. Rev., 100, 487-490.

Baumhefner, D. and Downey, P. 1978 "Forecast intercomparisons from three Numerical Weather Prediction Models", Mon. Wea. Rev., 106, 1245-1279.

Berkofsky, L. and Bertoni, E.A. 1955 "Mean topographic charts for the Entire Earth", Bull. Amer. Meteor. Soc., 36, 350-353.

Burridge, D.M. and Haseler, J. 1977 "A model for medium range weather forecasting", Technical Report No.4, ECMWF, Bracknell, England.

Burridge, D.M. and Haseler, J. 1979 "A semi-implicit global gridpoint model", Technical Report, ECMWF, Shinfield Park, Reading, England, to be published.

Charney, J.G. and Eliassen, A. 1949 "A numerical method for predicting the perturbations of the middle latitude westerlies", Tellus, I, 38-54.

References (contd)

Crutcher, H.L. and Jenne, R.J.	1970	"An interim note on northern hemisphere climatological grid data tape", NOAA Environmental Data Service, NWRC, Asheville.
Desmarais, J.G. and Derome, J.F.	1978	"Some effects of vertical resolution on modelling forced planetary waves with a time-dependent model", Atmosphere-Ocean, 16, 212-225.
Dopplick, T.G.	1972	"Radiative heating of the global atmosphere", J. Atm. Sci., 29, 1278-1294.
Druyan, L.M., Somerville, R.C.J. and Quirk, W.J.	1975	"Extended-range forecasts with the GISS model of the global atmosphere", Mon. Wea. Rev., 103, 779-795.
Fawcett, E.B.	1969	"Systematic errors in operational baroclinic prognoses at the National Meteorological Center", Mon. Wea. Rev., 97, 670-682.
Gandin, L.S.	1963	"Objective Analysis of Meteorological Fields", in Russian, Gidrometeorologicheskoe Izdatelstvo (GIMIZ), Leningrad. (English translation, Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.
Gates, W.L. and Nelson, A.B.	1975	"A new (revised) tabulation of the Scripps Topography"
Hollingsworth, A. and Geleyn, J-F.	1979	"A comparison of some diffusion operators in one and three dimensions", Internal Memorandum ECMWF, Shinfield Park, Reading, England. To be published.

References (contd.)

Hollingsworth, A. and Kållberg, P. 1979 "Spurious energy conversions in an energy/enstrophy conserving finite difference scheme", Internal Report No. 22, ECMWF, Shinfield Park, Reading, England.

Hovmöller, E. 1949 "The trough and ridge diagram", Tellus, I, 62-66.

Karhila, V. and Rinne, J. 1977 "Determination of EOF's from a large sample", ECMWF Workshop on the Use of Empirical Orthogonal functions in Meteorology, 84-96.

Kuo, H.L, 1974 "Further studies of the parameterization of the influence of cumulus convection on large-scale flow", J. Atmos. Sci., 31, 1232-1240.

Lambert, S.J. and Merilees, P.E. 1978 "A study of Planetary Wave errors in a spectral Numerical Weather Prediction Model", Atmosphere-Ocean, 16, 197-211.

Lau, N.C. 1978 "On the three dimensional structure of the observed transient eddy statistics of the Northern Hemisphere", J. Atmos. Sci., 35, 1900-1923.

Lorenc, A., Rutherford, I. and Larsen, G. 1977 "The ECMWF analysis and data-assimilation scheme: Analysis of mass and wind fields", Technical Report No. 6, ECMWF, Bracknell, England.

Lorenz, E.N. 1967 "The nature and theory of the general circulation of the atmosphere", WMO, 1-161.

References (contd.)

Manabe, S., Smagorinsky, J. and Strickler, R.F. 1965 "Simulated climatology of a general circulation model with a hydrologic cycle", Mon. Wea. Rev., 93, 769.798.

Miyakoda, K., Smagorinsky, J. Strickler, R.F. and Hembree, G.D. 1969 "Experimental extended predictions with a nine-level hemispheric model", Mon. Wea. Rev., 97, 1.76.

Miyakoda, K., Hembree, G.D., Strickler, R.F. and Shulman, I. 1972 "Cumulative results for extended forecast experiments. I Model performance for Winter cases", Mon. Wea. Rev., 100, 836-855.

Murray, F.W. 1967 "On the computation of saturation vapour pressure", J. Appl. Meteor., 6, 203-204.

Phillips, N.A. 1957 "A coordinate system having some special advantages for numerical forecasting", J. Meteor. 14, 184-185.

Robert, A.J., Henderson, J. and Turnbull, C. 1972 "An implicit integration scheme for baroclinic modes of the atmosphere", Mon. Wea. Rev., 100, 329-335.

Rossby, C.G. 1945 "On the propagation of frequencies and energy in certain types of oceanic and atmospheric waves", J. Meteor., 2, 187-204.

Rutherford, I.D. 1976 "An operational three-dimensional multivariate statistical objective analysis scheme", GARP Report II. Proceedings of the JOC Study Group Conference on Four-Dimensional Data Assimilation, Paris, 17-21 November 1975, WMO, Geneva, pp.98-121.

References (contd.)

Sadourny, R. 1975 "The dynamics of finite difference models of the shallow water equations", J. Atmos. Sci., 32, 680-689.

Schlatter, T.W. 1975 "Some experiments with a multivariate statistical objective analysis scheme", Mon. Wea. Rev., 103, 246-257.

Shuman, F. and Newell, J. 1970 "Revised Method of 1000mb height computation in the PE model", Technical Procedure Bulletin No. 57, National Weather Service, NOAA, Washington.

Simmons, A.J. and Hoskins, B.J. 1979 "The downstream and upstream development of unstable baroclinic waves", J. Atmos. Sci. (in press).

Smagorinsky, J. 1953 "The dynamical influences of large scale heat sources and sinks on the quasi-stationary mean motions of the atmosphere", Quart. J. Roy. Meteor. Soc., 79, 342-366.

Smagorinsky, J. Manabe, S. and Holloway, J.L. 1965 "Numerical results from a nine-level general circulation model of the atmosphere", Mon. Wea. Rev., 93, 727-768.

Temperton, C. and Williamson, D.L. 1979 "Normal Mode initialization for a multi-level grid point model", Technical Report No. 11, ECMWF, Shinfield Park, Reading, England.

Teweles, S. and Wobus, H.B. 1954 "Verification of Prognostic Charts", Bull. Amer. Meteor. Soc., 35, 455-463.

References (contd.)

Tiedtke, M., Geleyn, J-F., Hollingsworth, A. and Louis, J-F. 1979 "Parameterization of sub-grid Scale Processes", Technical Report No. 10, ECMWF, Shinfield Park, Reading, England.

van Loon, H. 1965 "A climatological study of the atmospheric circulation in the Southern Hemisphere during the IGY, Part I: 1 July 1957-31 March 1958", J. Appl. Meteor., 4, 479-491.

Yeh, T-C. 1949 "On energy dispersion in the atmosphere", J. Meteor., 6, 1-16.