AN EXTENSIVE QUASI-OPERATIONAL COMPARISON BETWEEN A SPECTRAL AND A GRID POINT MODEL

M.Jarraud and C.Girard
European Centre for Medium Range Weather Forecasts
Reading, UK

Note: This presentation is mainly a summary of a more thorough investigation reported in Girard and Jarraud (1982).

1. INTRODUCTION

At almost the same time ECMWF selected the finite difference technique for its first operational model, it was decided to develop and test in parallel a spectral model. This led to an uncommon situation; such comparisons usually involve a newly developed model and a more ancient one differing from each other in many important aspects. Here, neither model is more modern than the other and they are as similar as possible except for the technique used for the horizontal discretization. The grid point model (referred to as "N48" in the following) is described in Burridge and Haseler (1977). In brief, it works on a regular latitude-longitude grid ("C-grid" in Arakawa's classification) with 48 lines of latitude betwen pole and equator. The vertical coordinate is terrain following ("sigma") with 15 irregularly spaced levels. The spectral model (described in Baede et al. 1979) is referred to as "T63" since it uses a triangular truncation retaining all zonal wavenumbers up to 63. The vertical coordinate and discretization are identical to those of the grid point model. Both models use a semi-implicit time discretization scheme with a 15 minutes time step for N48 and 18 minutes for T63. Except for horizontal diffusion they also use the same parameterizations of diabatic processes (described in Tiedtke et al. 1979): these include radiation, large scale condensation, moist convection, boundary layer and vertical turbulent fluxes.

For the horizontal diffusion, both models use a 4th order scheme, linear for T63 and nonlinear at the start for N48. During the course of the comparison, this was changed to a linear scheme similar to the one used by T63, however with a much larger coefficient which was reduced later. All other changes to the operational physical package were carried simultaneously into T63.

Since the main purpose of the comparison was to test a potential alternative to the operational N48 model, the horizontal resolution of the spectral model was selected such that the two models required similar computer resources.

The experiments were done regularly once a week between September 1979 and August 1980 covering thus a wide range of synoptic situations. The initial conditions were the operational ECMWF analysis (which means that N48 was used both in analysis and initialization steps).

It was decided to concentrate the evaluation over the extratropical part of the Northern Hemisphere where data and forecasts are more reliable. The size of the sample (53 cases) allowed us to gather significant evidence on forecasts differences between the models and on their nature (Sects. 2 to 5). We have also tried to relate them to theoretical differences in technique (or to other known model differences) (Sect.5.).

2. OBJECTIVE EVALUATION OF RESULTS

All results presented here are based on anomaly correlations (AC) of the height field (z) in the band 20° to 82.5°N, since this score was found to be more sensitive to the forecast differences being investigated. Other scores (standard deviation) for height and anomaly correlation of temperature are also presented in Girard and Jarraud (1982).

Fig.1 gives an overall impression of the results by displaying scatter diagrams of AC(z) vertically integrated from 1000 to 200 mb on D+3, D+4, D+5, D+6, D+8 and D+10. A point above the diagonal corresponds to a better forecast by T63 (and vice versa for N48). At almost all ranges there is a larger variability between cases than between the models (except may be at D+10). The superiority of T63 increases from D+3 to D+5. Note that it is almost systematic on D+4 (less than 20% of the points are in favour of N48). Note also some very large differences on D+5 in cases where at least one of the models keeps some predictive skill. (e.g. AC>60%). Most of them represent improvements of T63 over N48.

Beyond D+6 the number of cases with predictive skill left decreases sharply suggesting that other sources of errors (data, resolution, physics,...) have become predominant and making thus more difficult to investigate technique differences.

Furthermore, the variability is such that the statistical significance of the results is getting doubtful. Therefore in the rest of this presentation we shall concentrate mainly on the range D+3 up to D+5 where the differences are most significant and less affected by the other sources of errors.

As seen from Fig.2, the improvement of T63 over N48 is not uniform in the vertical: it is larger at 1000 mb where very large differences are already seen on D+3. It is also mainly concentrated in the long waves (Fig.3) which are usually the most predictable components of the flow. If we look at the results from a predictability (P) point of view (Fig.4), it is again clear that the improvement brought by the spectral model is larger at 1000 mb than at 500 mb. For a quality level corresponding to anomaly correlations of 60% (generally accepted to be a reasonable limit beyond which forecasts are no longer useful), T63 increases P(60%) at 1000 mb by more than a day in 9% of the cases (while the reverse never ocurred) and by more than 12 h in almost 20% of the cases (the reverse was true in only one case). Note again that for good quality forecasts (AC>60%) almost all the very large improvements are in favour of the spectral model.

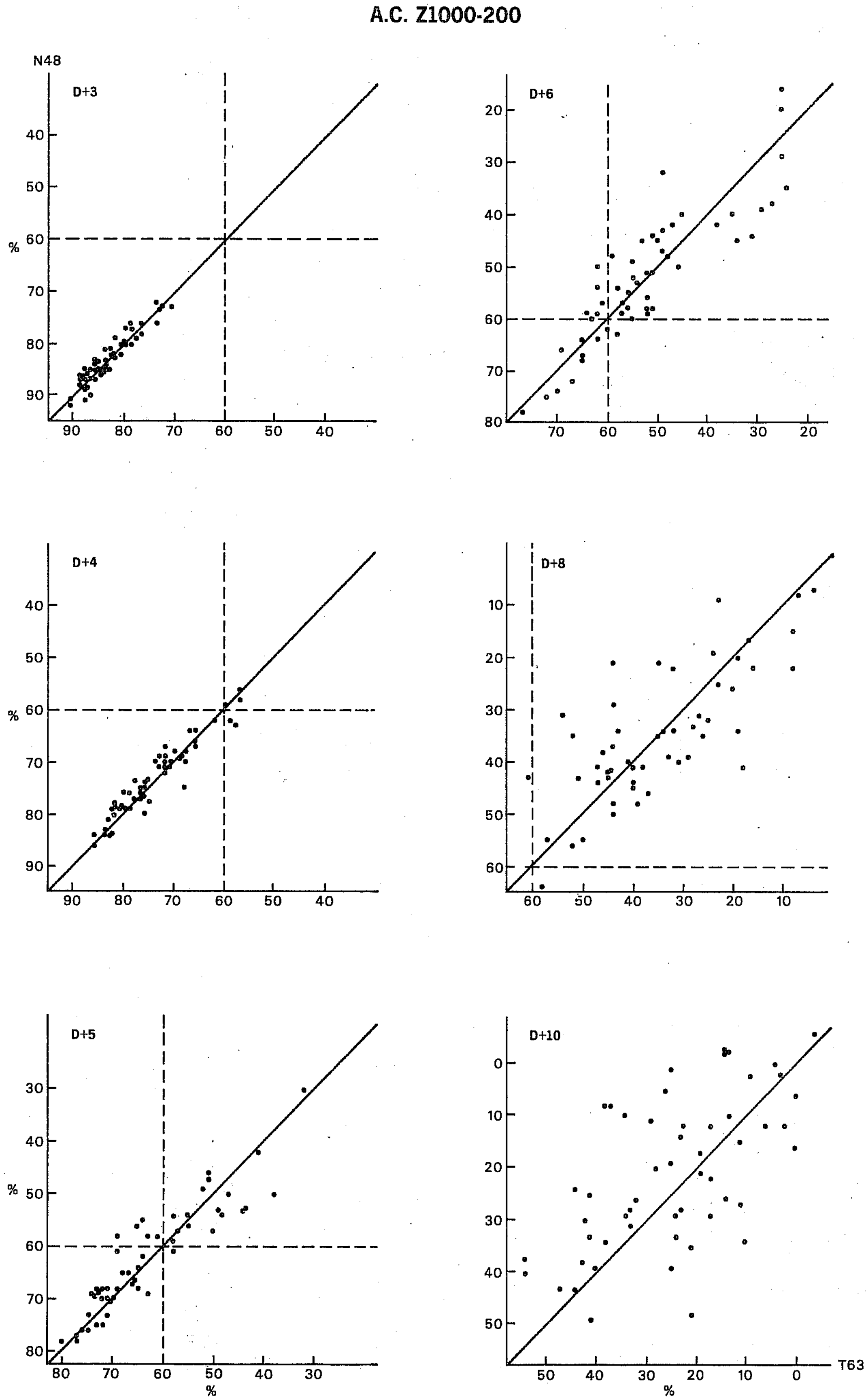


Fig. 1 Scatter diagrams showing AC(Z) on D+3, D+4, D+5, D+6, D+8, D+10 averaged horizontally from 20°N to 82.5°N and vertically from 1000 to 200 mb, comparing 53 forecasts by T63 (horizontal axis) and N48 (vertical axis).

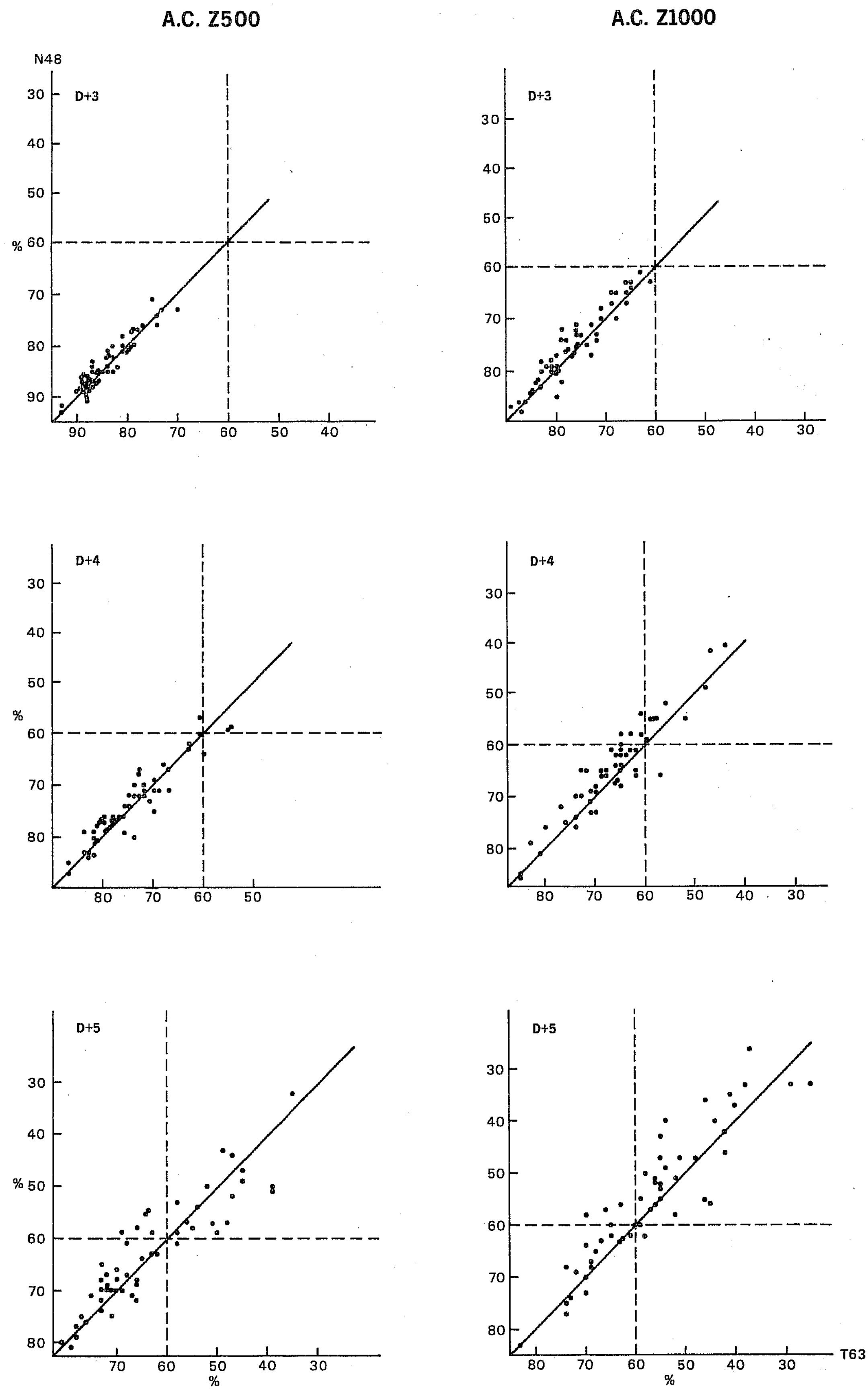


Fig. 2 Scatter diagrams showing AC(Z500) and AC(Z1000) on D+3, D+4, D+5 averaged horizontally from 20°N to 82.5°N, comparing 53 forecasts by T63 (horizontal axis) and N48 (vertical axis).

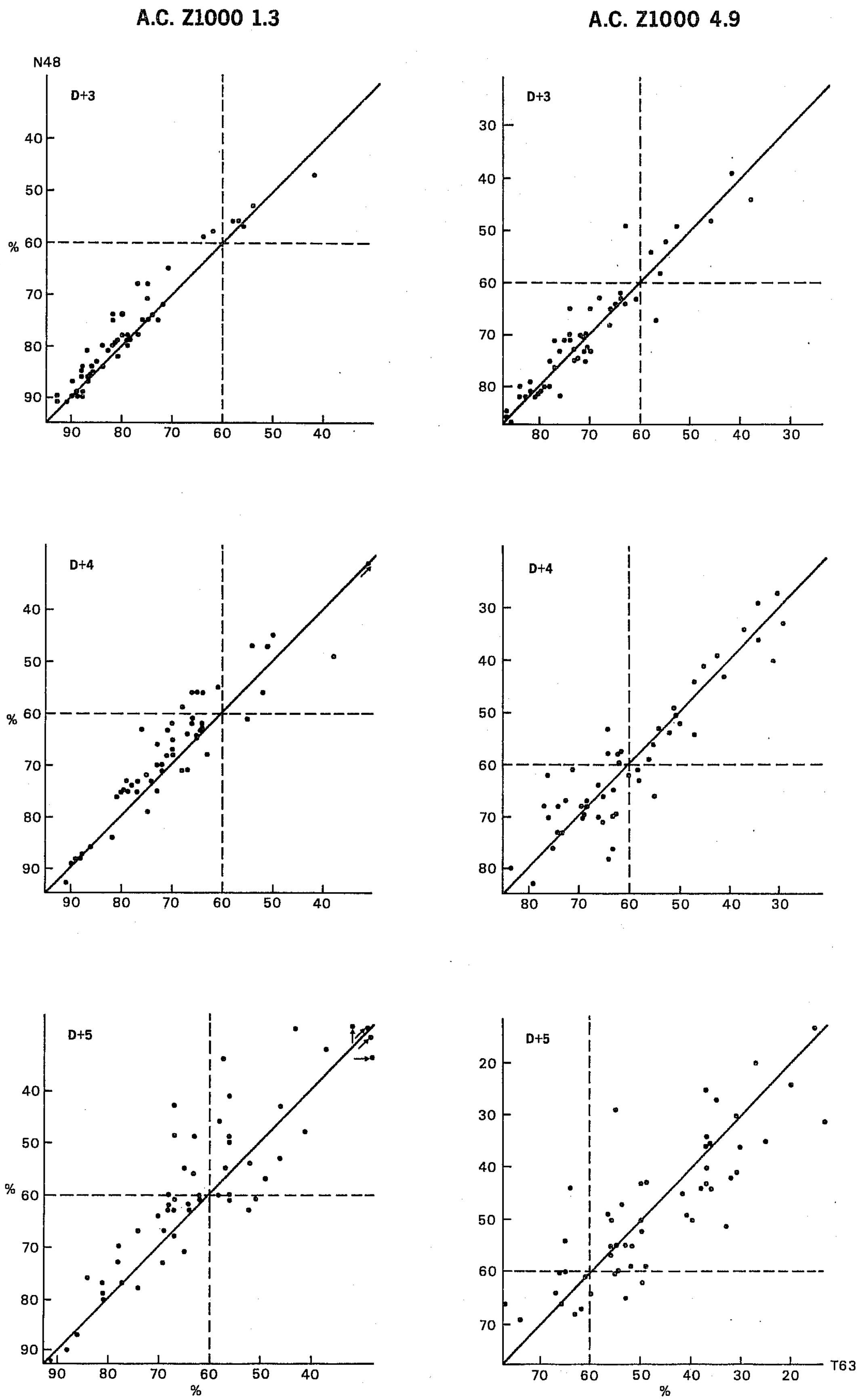


Fig. 3 Same as Fig.2 showing AC(Z1000) for the long (1-3) and medium (4-9) wave components.

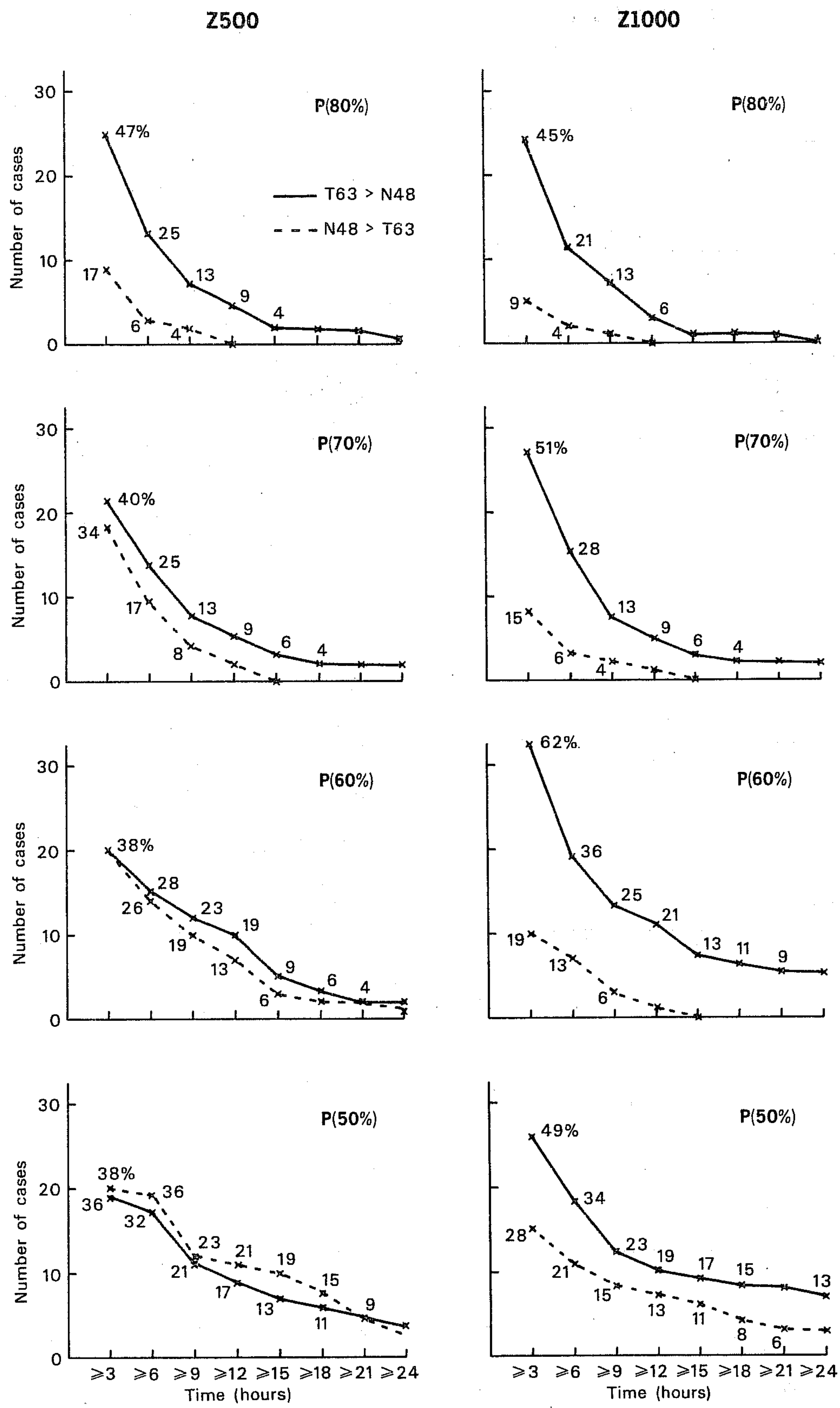


Fig. 4 Predictability improvements (in terms of AC) of one model over the other for Z500 and Z1000 at various fixed levels of quality: P(80%), P(70%) P(60%), P(50%). Crosses on the full (dashed) lines correspond to the number of forecasts where relative predictability improvements greater than 3,6,9,....24 hours were observed in favour of T63(N48). Corresponding percentages (out of the 53 cases) are also given.

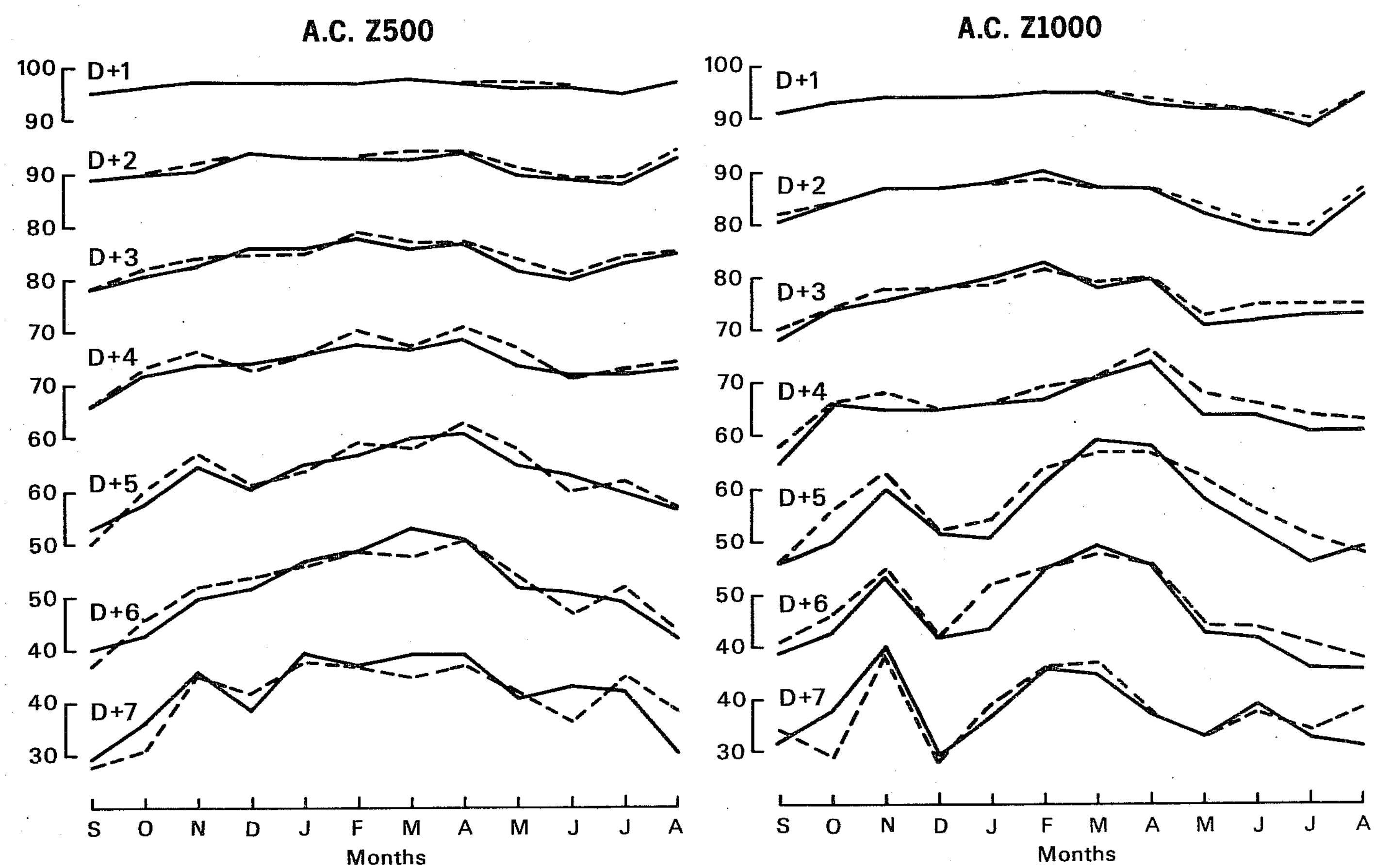


Fig. 5 Monthly mean values of AC(Z500) and AC(Z1000) averaged horizontally from 20°N to 82.5°N on D+1 to D+7. (T63= ---; N48= ——)

In other words, most of the time the two models are very close, closer to each other than each of them to reality (see Fig.2) with only a small mean advantage to T63. However in a significant number of cases, large or very large differences are found, mostly in favour of T63.

The analysis of the mean monthly scores (Fig.5) shows no significant seasonal trend: the difference between the two models in this set of experiments is smaller in winter but a careful investigation of the corresponding forecasts led to the conclusion that it was due to causes other than technique differences (as will be shown in the following). Furthermore extra winter experiments described in Jarraud et al. (1981) and in Sect.6 below showed a large positive impact for the spectral model.

Finally, it is worth mentioning that no significant signal was found in the geographical distribution of the differences between the two models.

3. SYNOPTIC EVALUATION

As mentioned in the introduction, the models are very similar and thus forecast differences should be expected to be small. In practice they are much smaller than the forecast errors, at least up to D+5.

However, in the following we shall try to show their meteorological significance. The differences are found to occur preferentially in regions of tight gradients in the height field, in relation with large scale flow features such as troughs and ridges. Largest differences are observed in association with fast moving or developing disturbances and preferred locations for their occurrence depend on the type of atmosphere circulation prevailing. Their amplitude is of course growing with the forecast range. We can classify these differences with respect to the amplitude or the position of the observed meteorological features.

First, it is worth mentioning that both models share a number of deficiencies, in particular in simulating the evolution of baroclinic disturbances: both models tend to overdevelop large mature cyclones, T63 even slightly more than N48 (an example is given in Fig.6). These overdevelopments occur mostly over the North east Pacific and European areas. Both models also underestimate the speed and growth of new developing cyclones, mainly near the east coast of the continents.

This is however less pronounced for T63 and it represents the most common and most characteristic type of forecast differences in the range D+2 to D+5 between the two models. The systems in the N48 forecasts are almost always found to lag behind those of T63 which are also usually better in phase with the observed ones. Many examples in all seasons can be found, mainly over the Pacific and Atlantic oceans, following the cyclone tracks. Since these tracks have often a SW-NE orientation, so have the phase difference patterns. They are typified by a plus-minus dipole on the maps of the difference N48-T63. Examples are shown in Figs. 7 to 9 for various dates (10 September, 7 February and 17 July). Note in all of them the typical dipole signature.

Track differences are also obvious in many cases: on average the cyclones in T63 move not only faster, but also on more northern trajectories, this being again, in general, nearer reality. This can be seen in Fig.7 for the low near the east coast of Canada and in Fig.8 for both lows. Several other examples are shown in Girard and Jarraud (1982).

Another difference between the two models is that T63 maintains throughout the forecast a higher level of variance than N48 as will be shown in the next subsection. At 1000 mb this is connected with the overdevelopment already

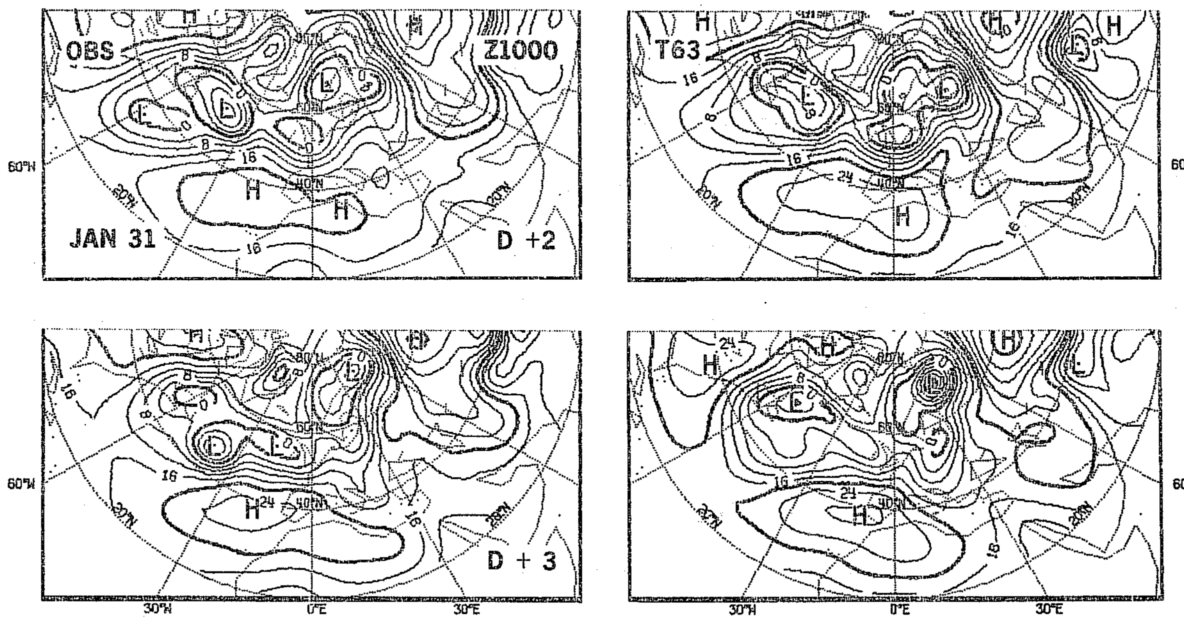


Fig. 6 JAN. 31, Z1000: D+2, D+3

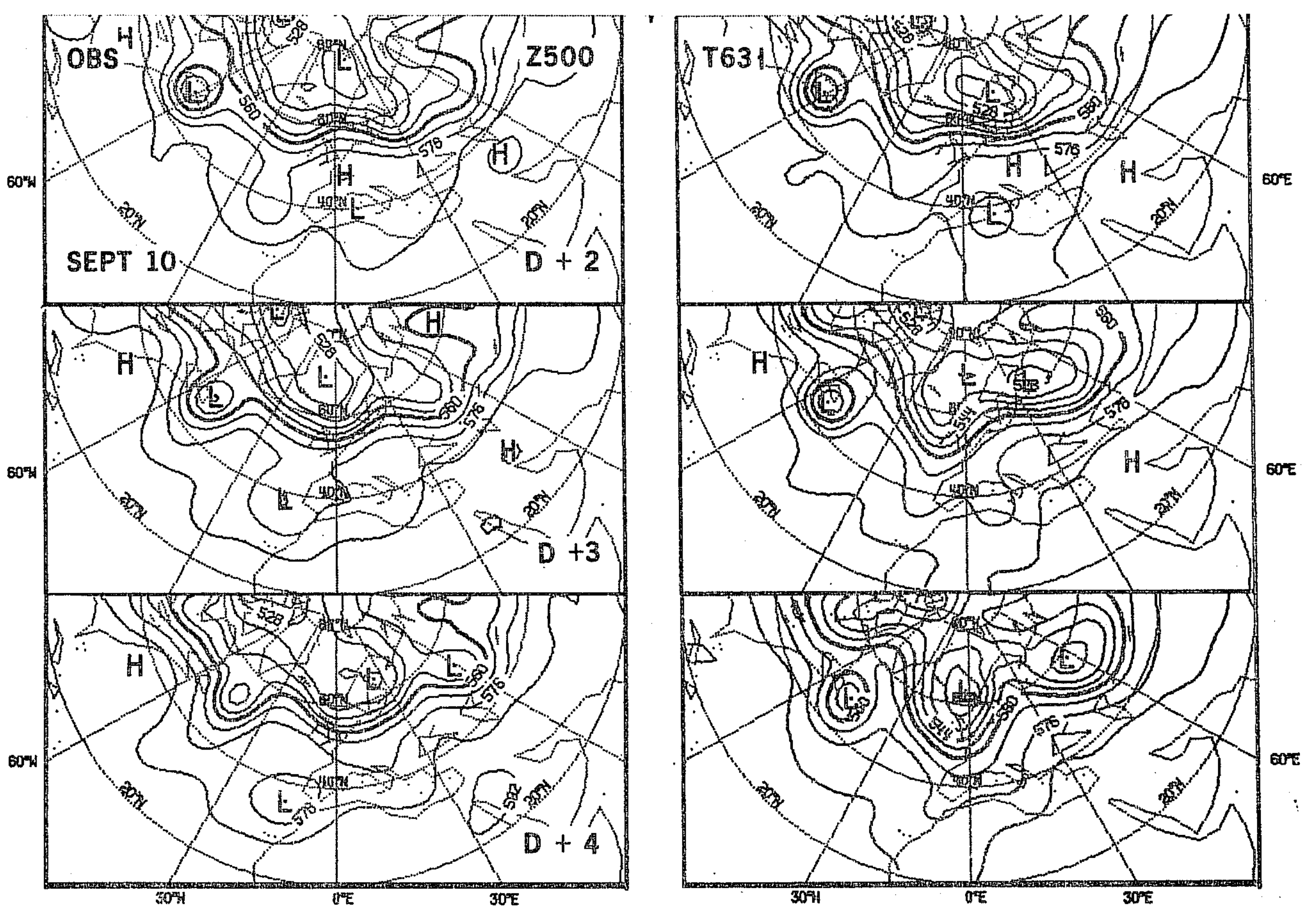


Fig. 7 SEP. 10, Z500: D+2, D+3, D+4

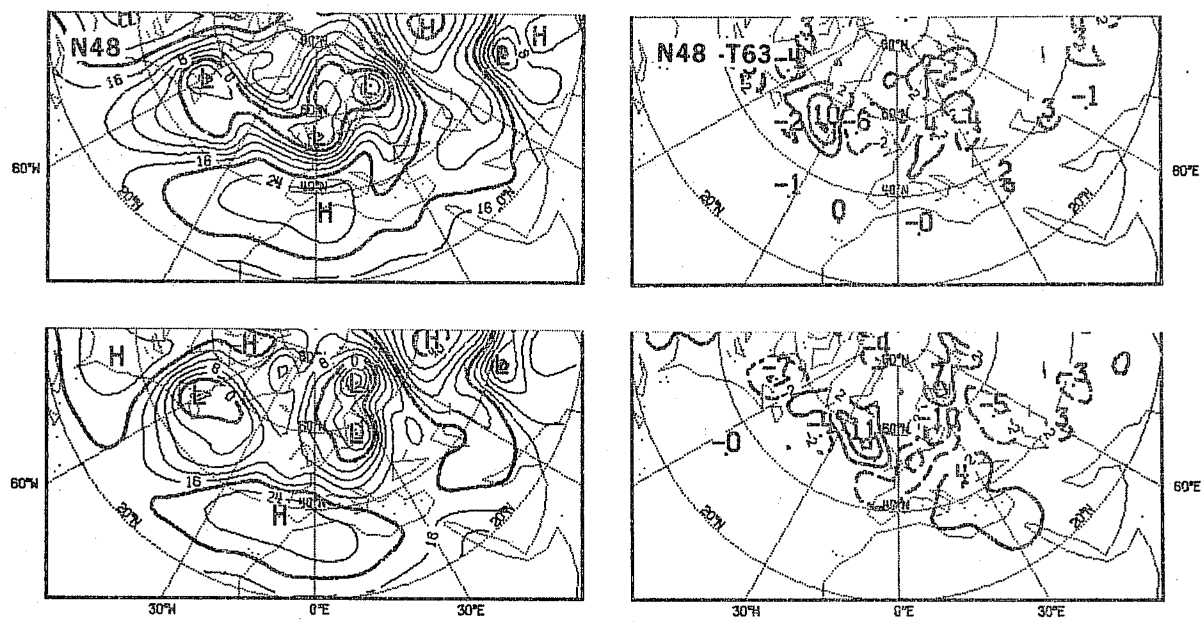


Fig. 6 (Cont.)

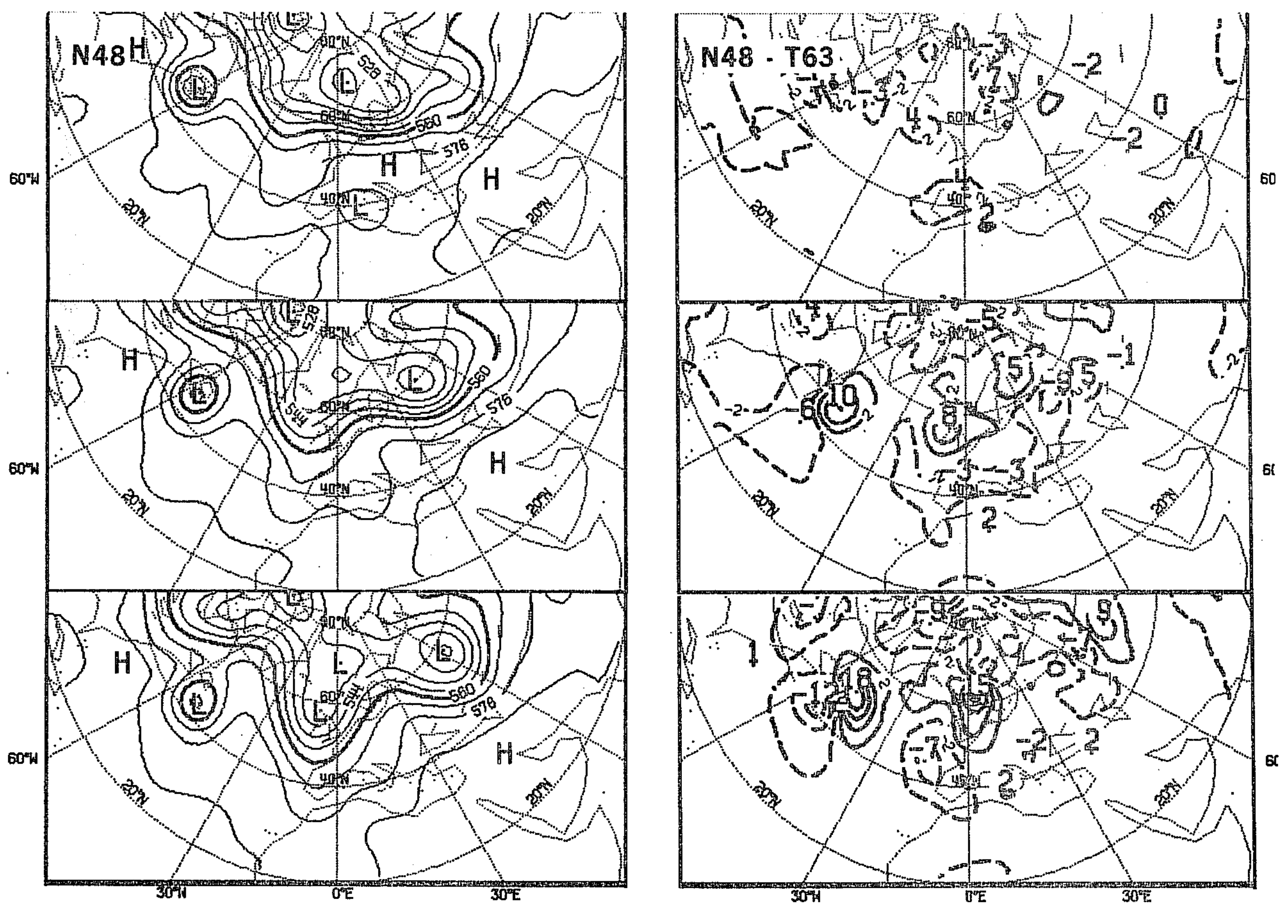


Fig. 7 (Cont.)

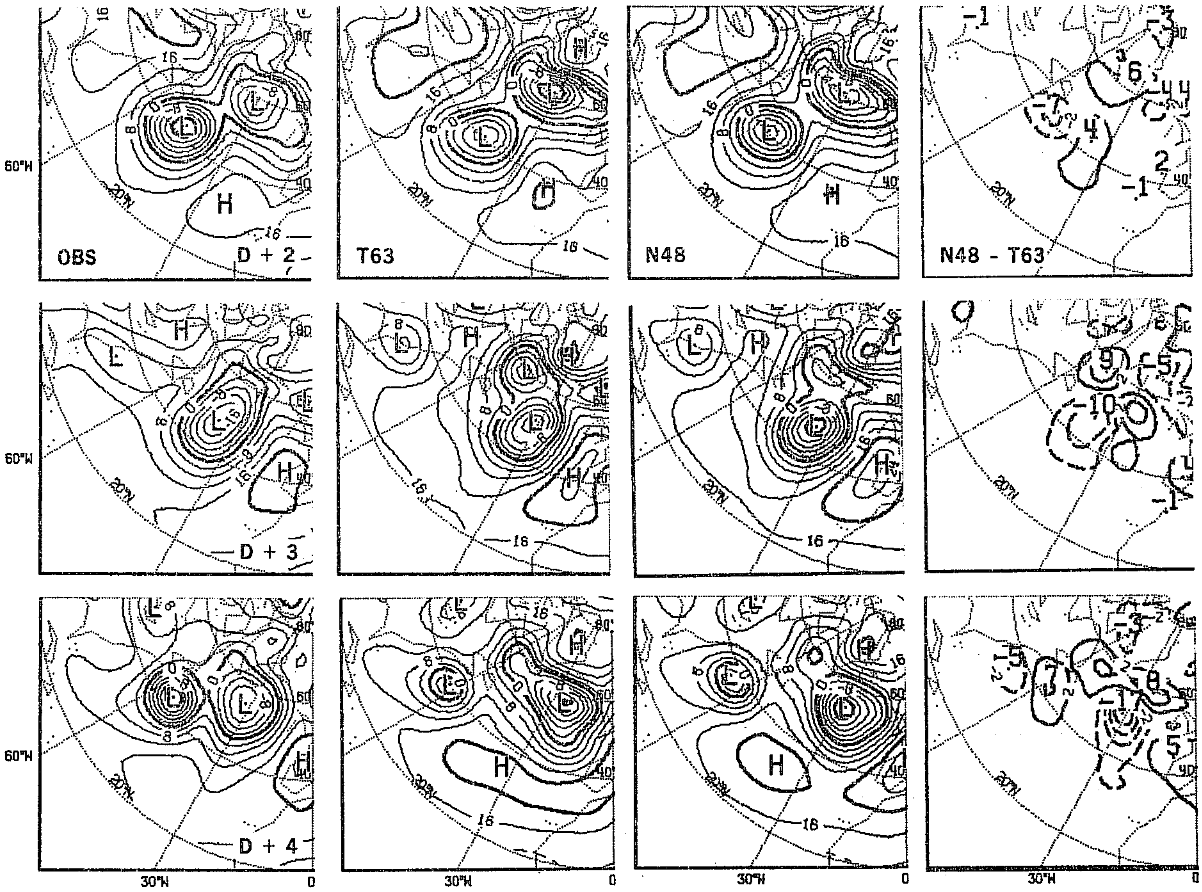


Fig. 8 FEB. 7, Z1000: D+2, D+3, D+4

mentioned. At 500 mb, however the model variance levels are too low, but T63 maintains a slightly higher level of eddy kinetic energy, in better agreement with reality.

An illustration is given in Fig.10 showing the D+2 and D+4 500 mb height forecasts for the 30 September case. T63 is more able to maintain the amplitude of the main troughs (the main differences are positive showing that T63 is deeper than N48). However this usually becomes clear only after 5 or 6 days, when the forecasts have already lost much of their predictive skill.

All these forecast differences are usually small and the corresponding score differences are relatively small. Nevertheless, as already mentioned, large score differences were observed, mostly in favour of T63, on a few more sensitive cases. These were cases where large phase and amplitude differences developed quickly, leading to very different synoptic forecasts over large part of the Northern Hemisphere by D+4 and D+5 already.

Since we could not easily classify these forecast differences into well defined categories, we think it is worth presenting several examples.

For the October 25th case, at 500 mb (Fig.11) one can observe a difference with a very large amplitude and scale related to the high north of Russia. The T63 forecast is to be preferred. Note also the better representation of the trough and ridge system over Europe by T63.

For November 1st, Fig. 12 shows the largest forecast difference observed on D+5 at 1000 mb and one of the largest at 500 mb.

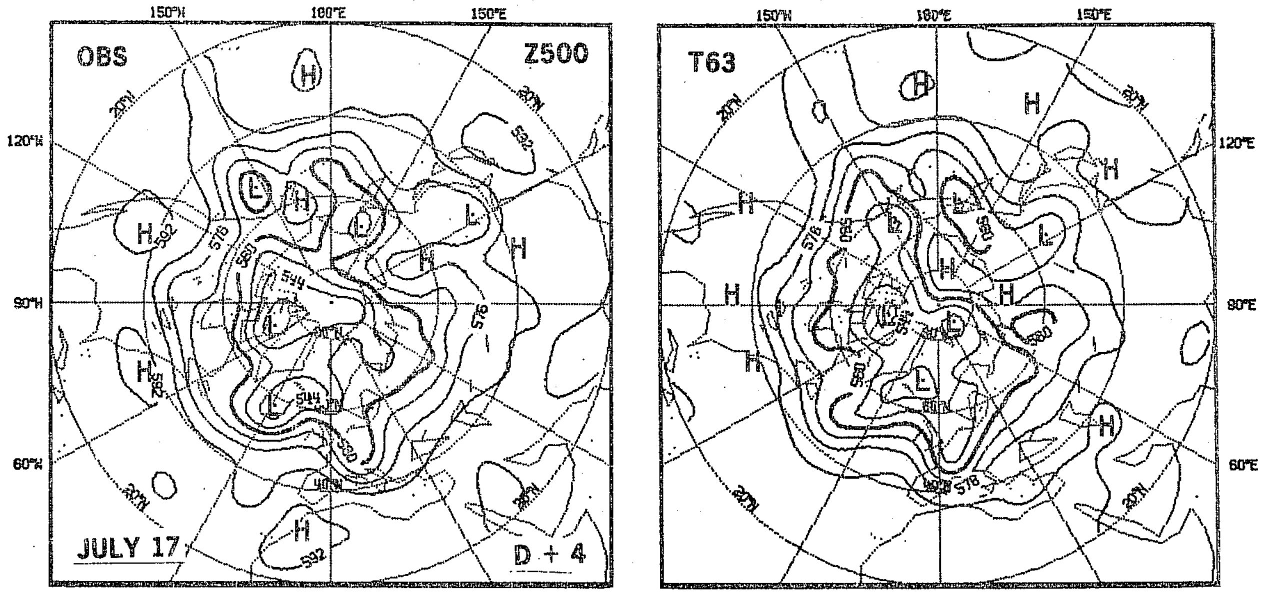


Fig. 9 JUL. 17, Z500: D+4

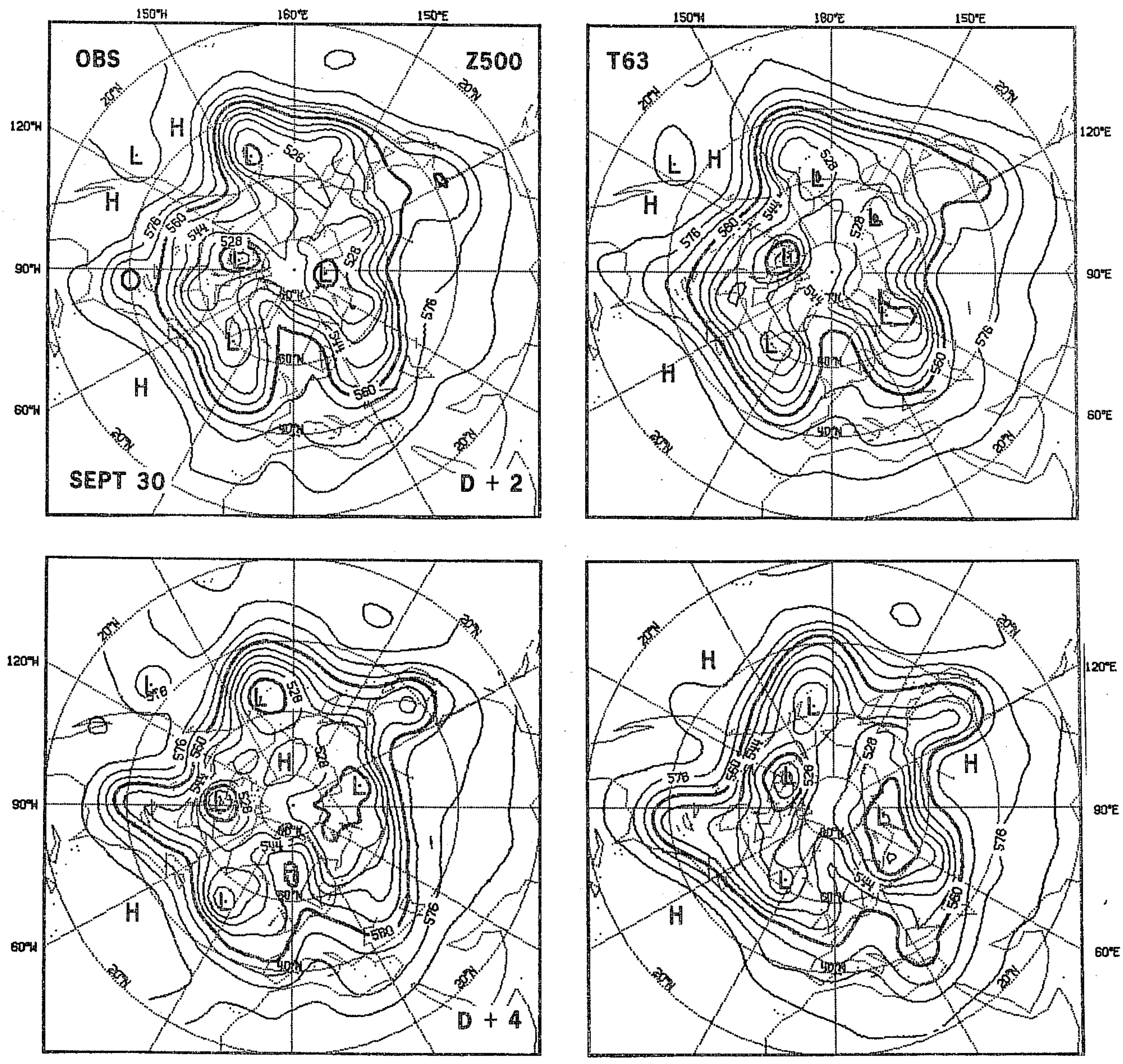


Fig. 10 SEP. 30, Z500: D+2, D+4.

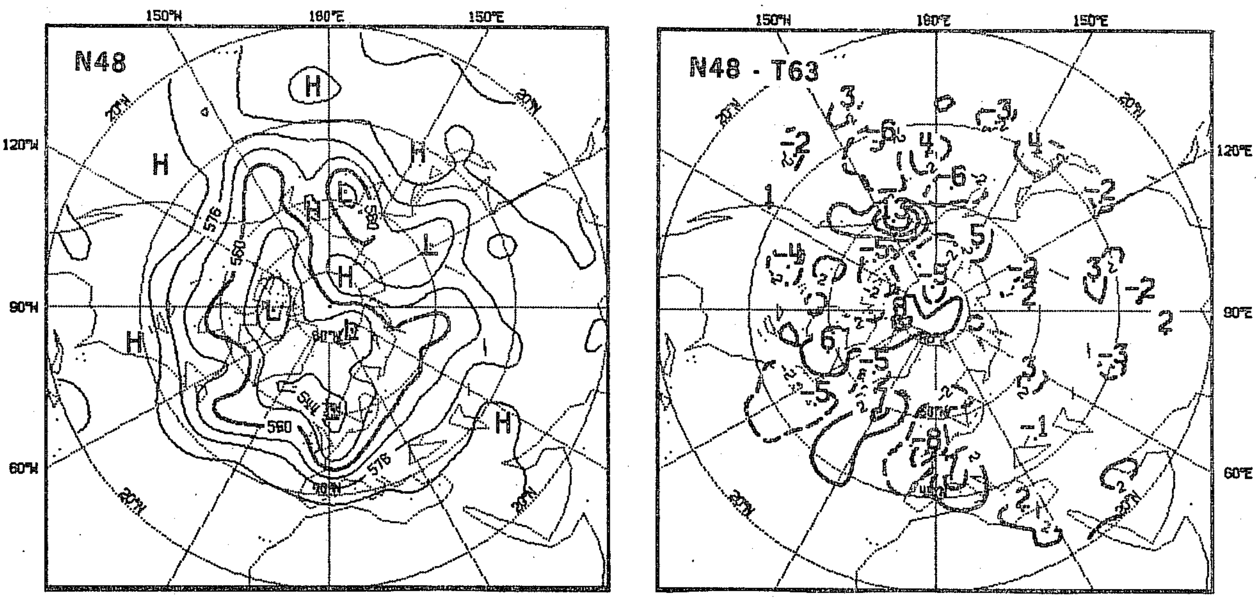


Fig. 9 (Cont.)

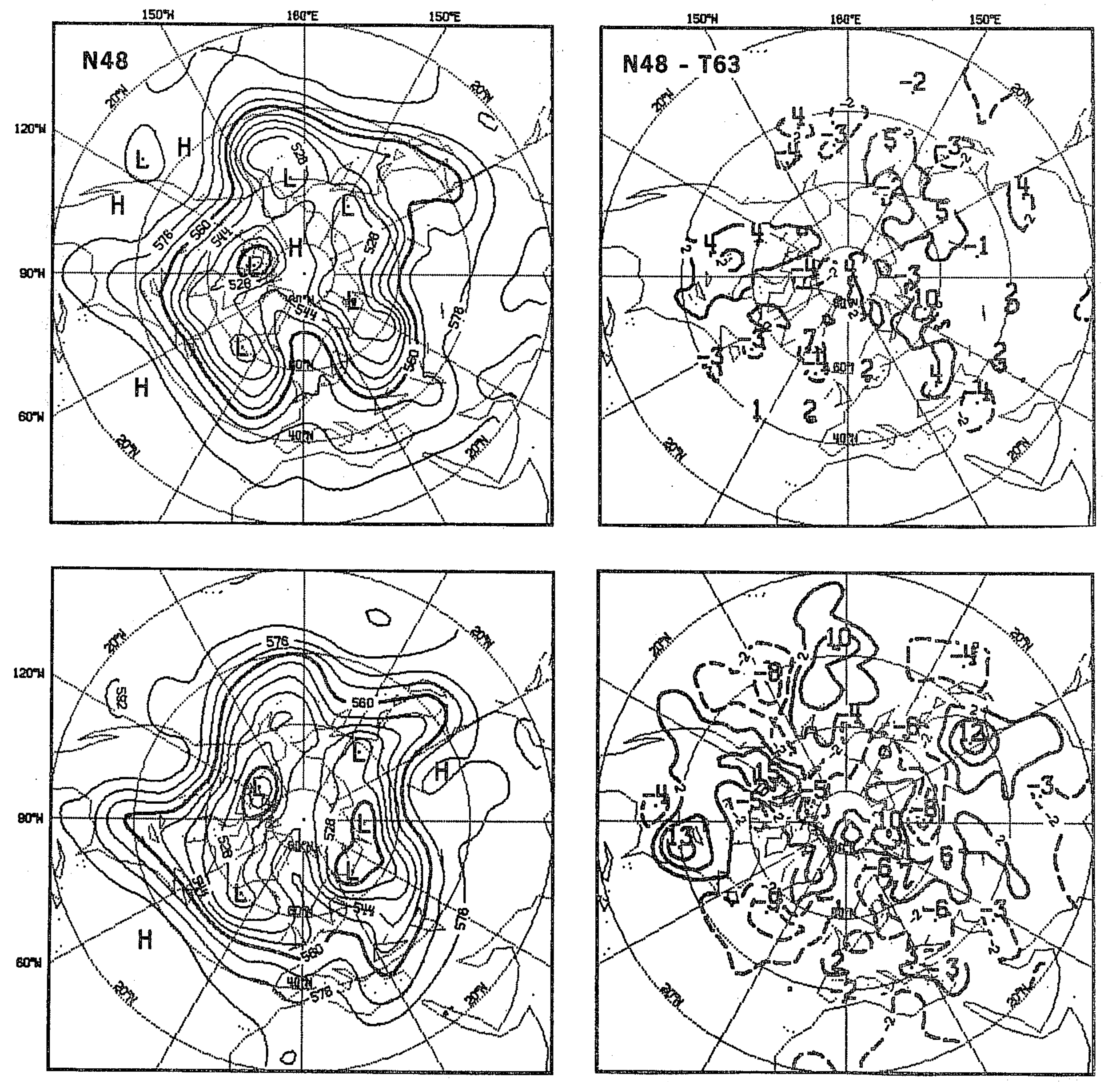


Fig. 10 (Cont.)

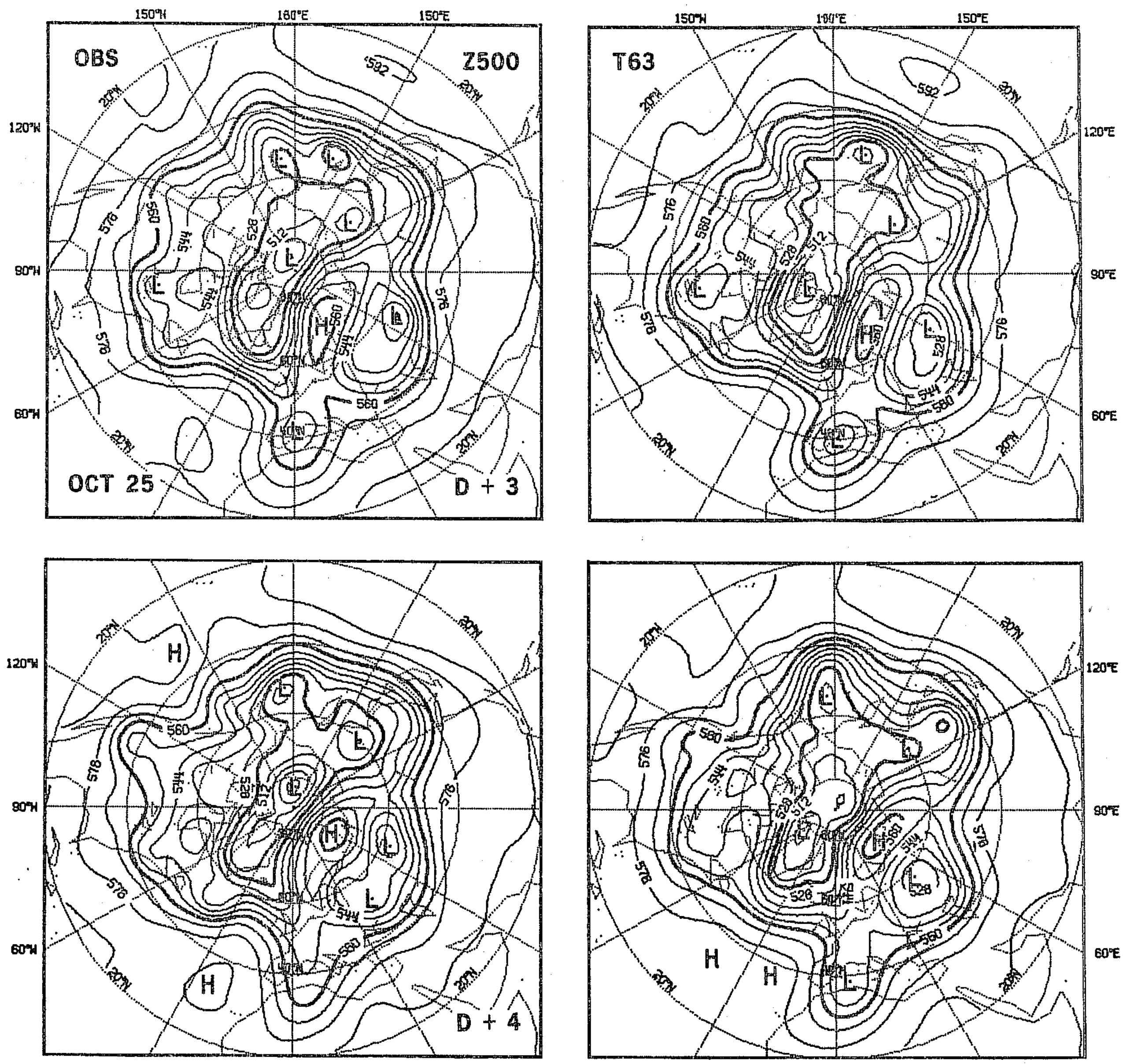


Fig. 11 OCT. 25, Z500: D+3, D+4

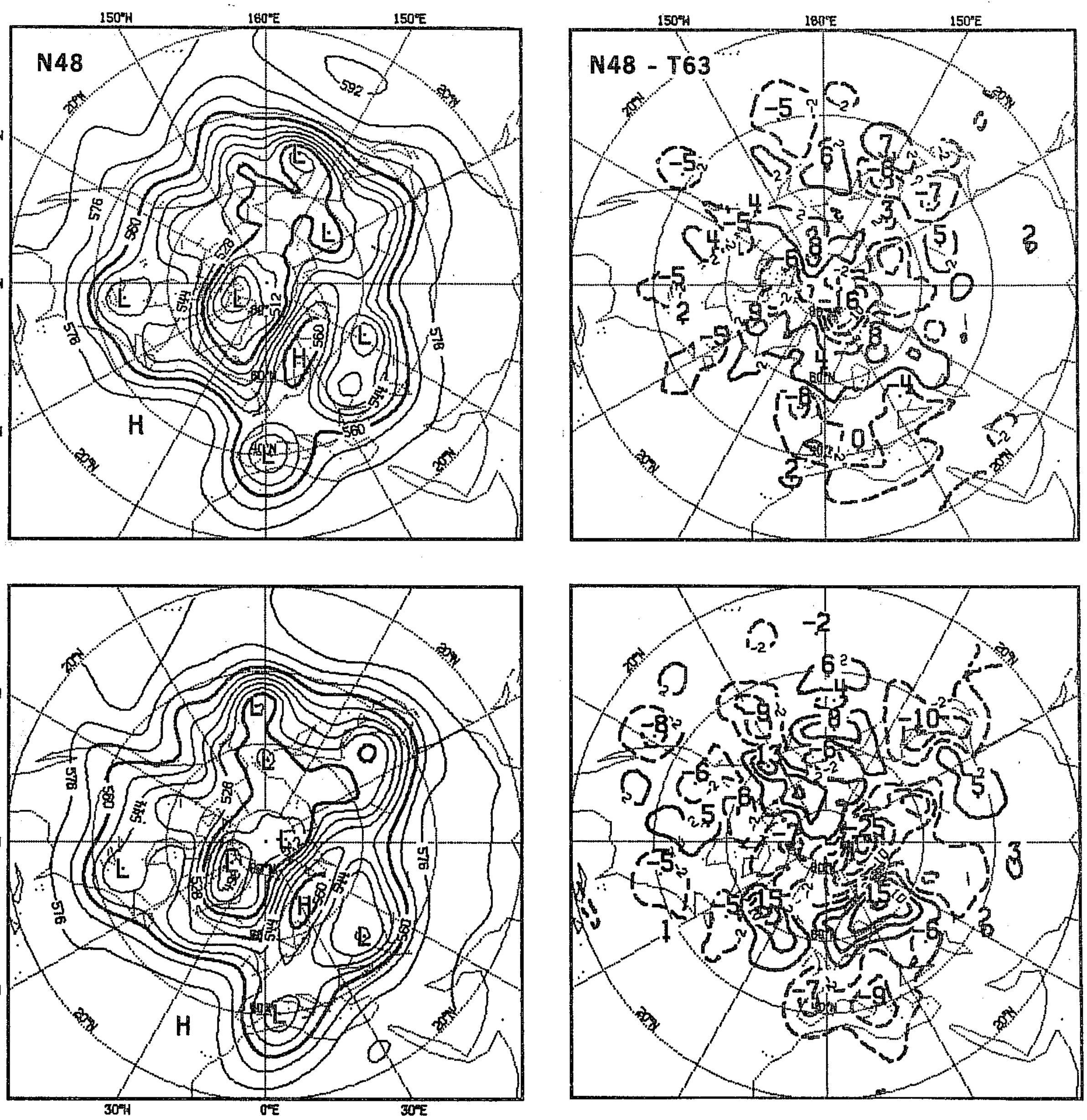


Fig. 11 (Cont.)

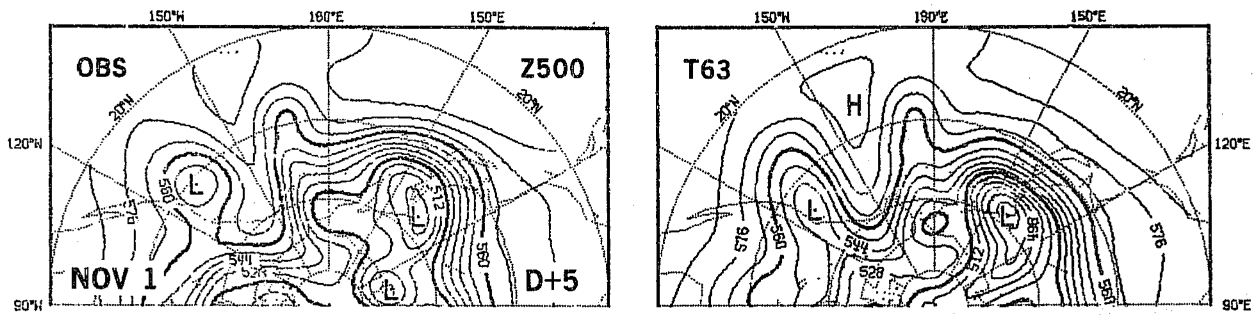


Fig. 12a NOV. 1, Z500: D+5 .

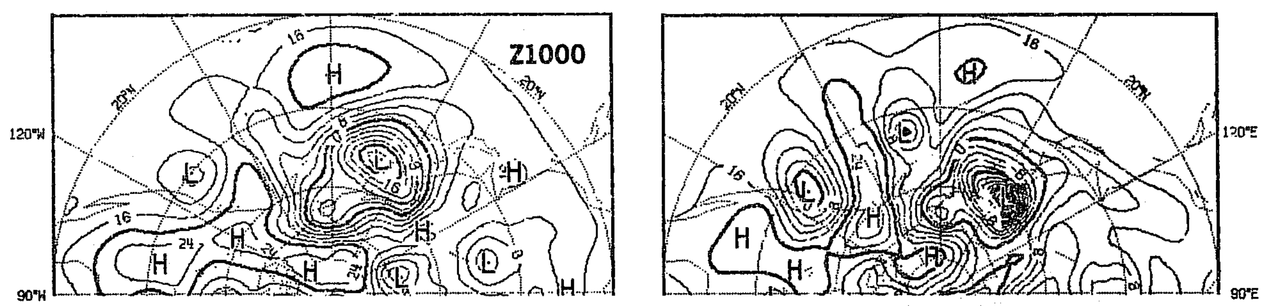


Fig. 12b NOV. 1, Z1000: D+5

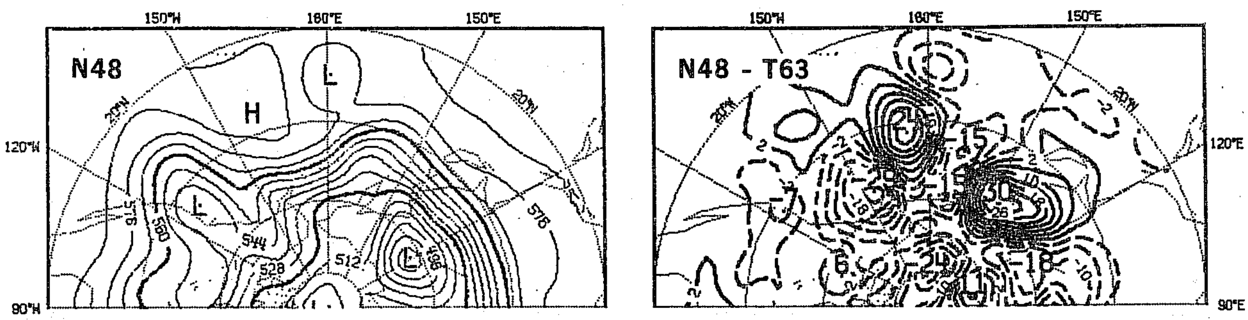


Fig. 12a (Contd.)

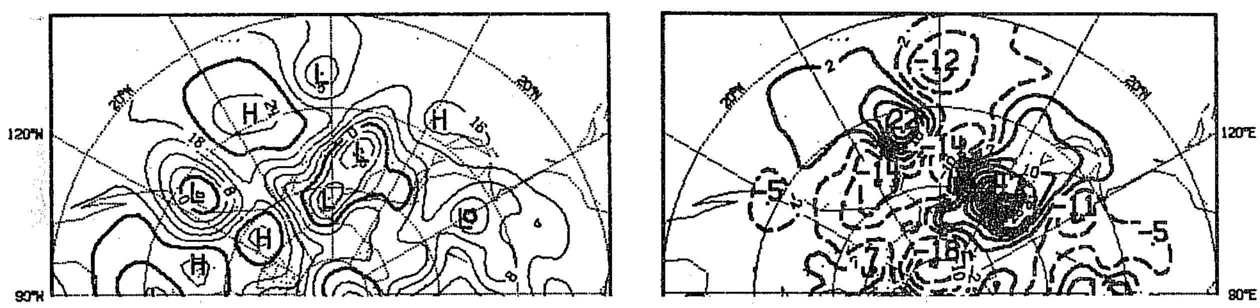


Fig. 12b (Contd.)

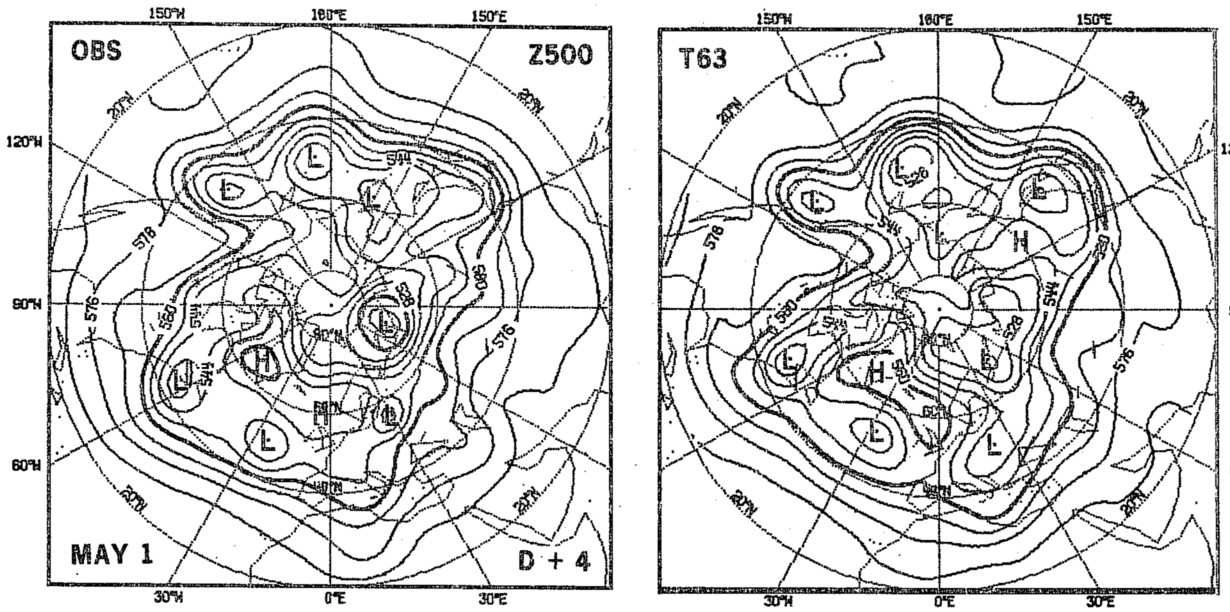


Fig. 13 MAY 1, Z500: D+4

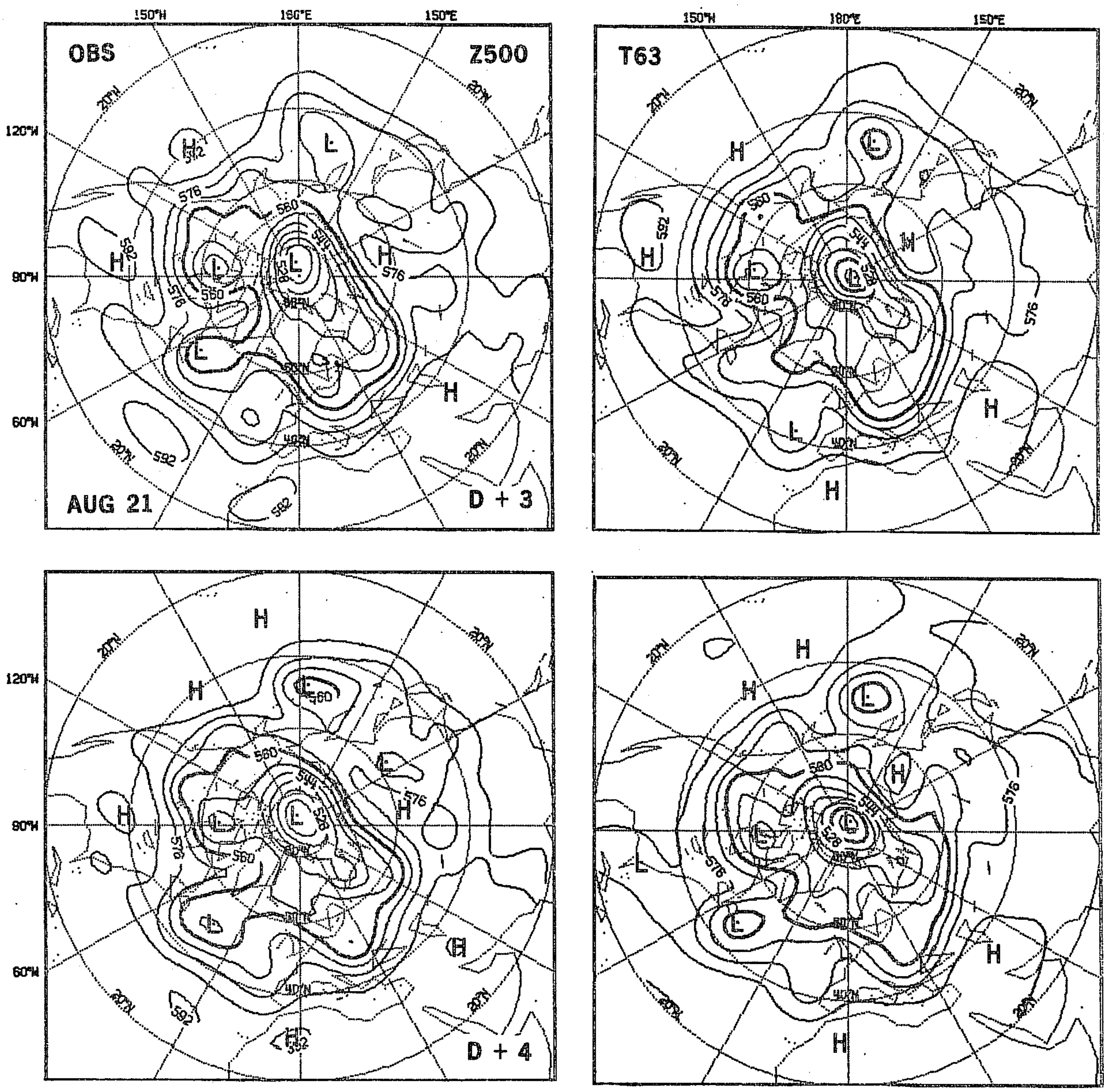


Fig. 14 AUG. 21, Z500: D+3, D+4

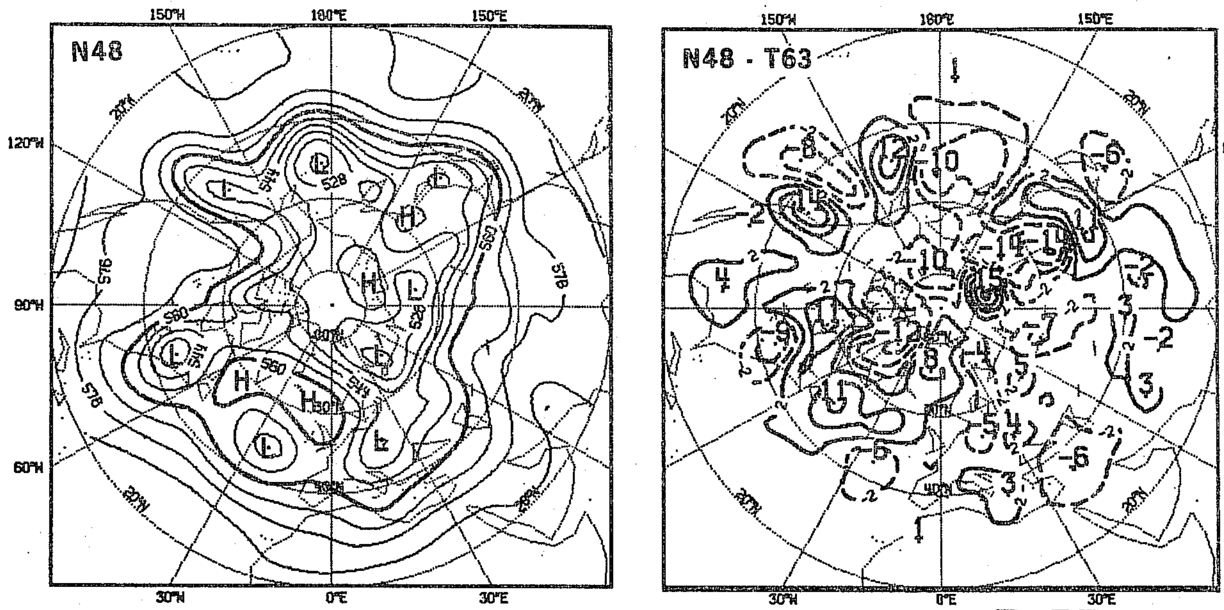


Fig. 13 (Cont.)

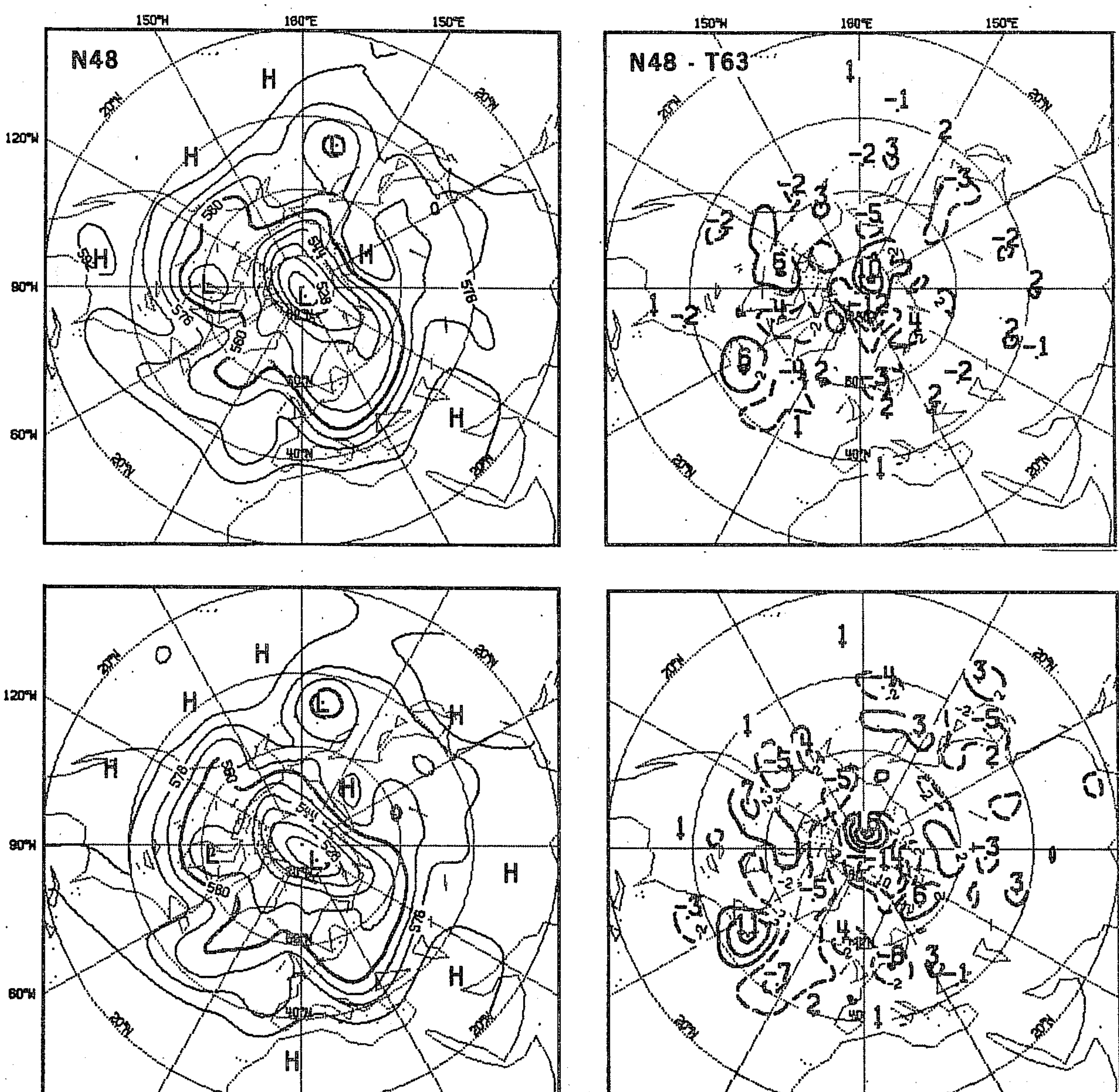


Fig. 14 (Cont.)

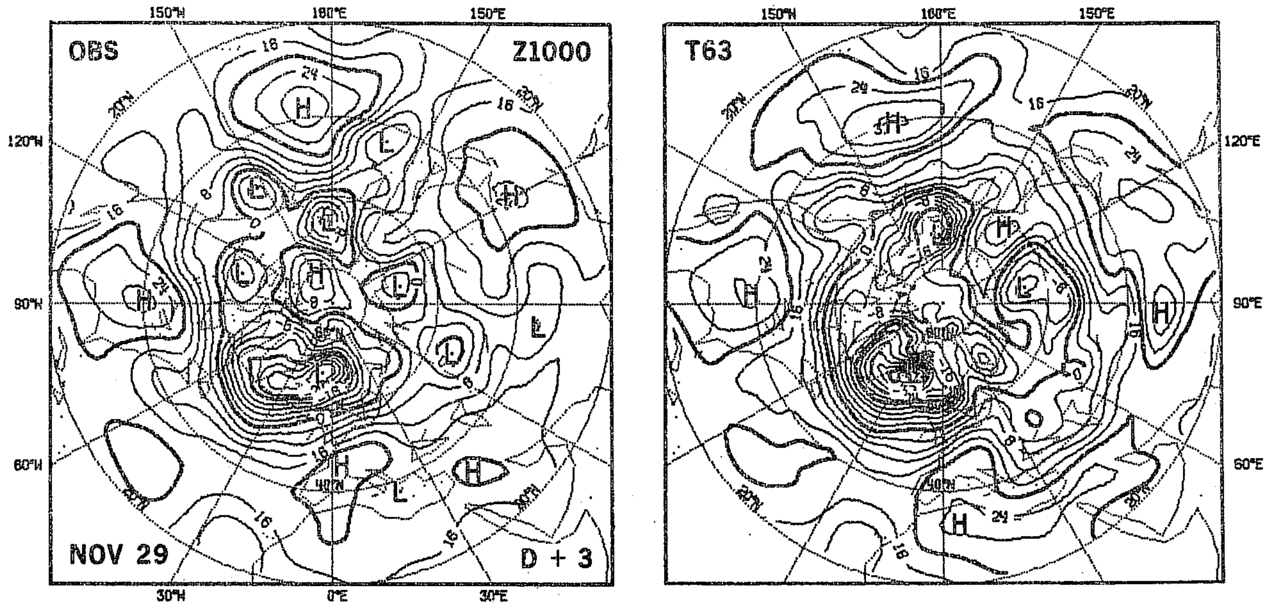


Fig. 15 NOV. 29, Z1000: D+3

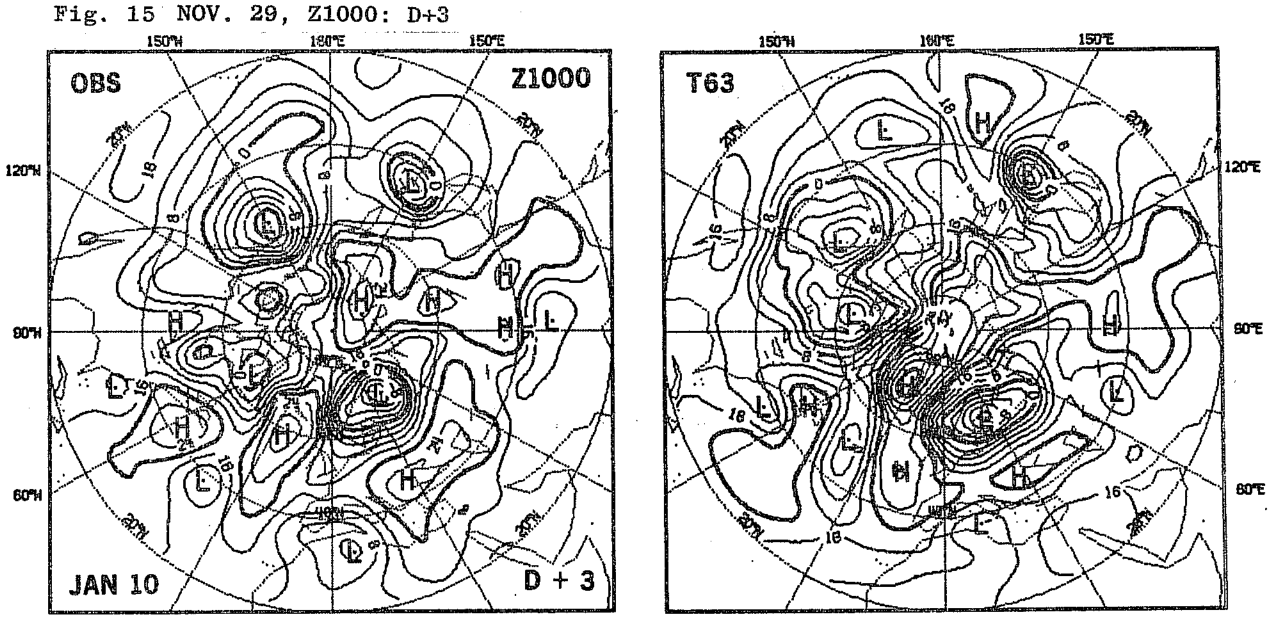


Fig. 16a JAN. 10, Z1000: D+3

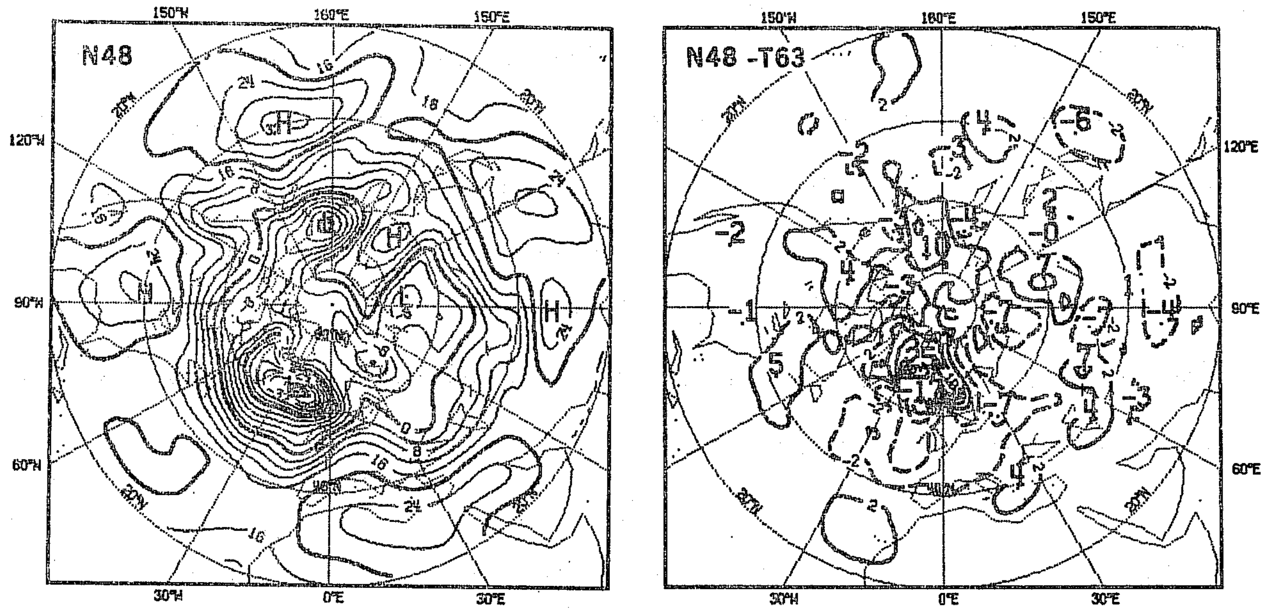


Fig. 15 (Contd.)

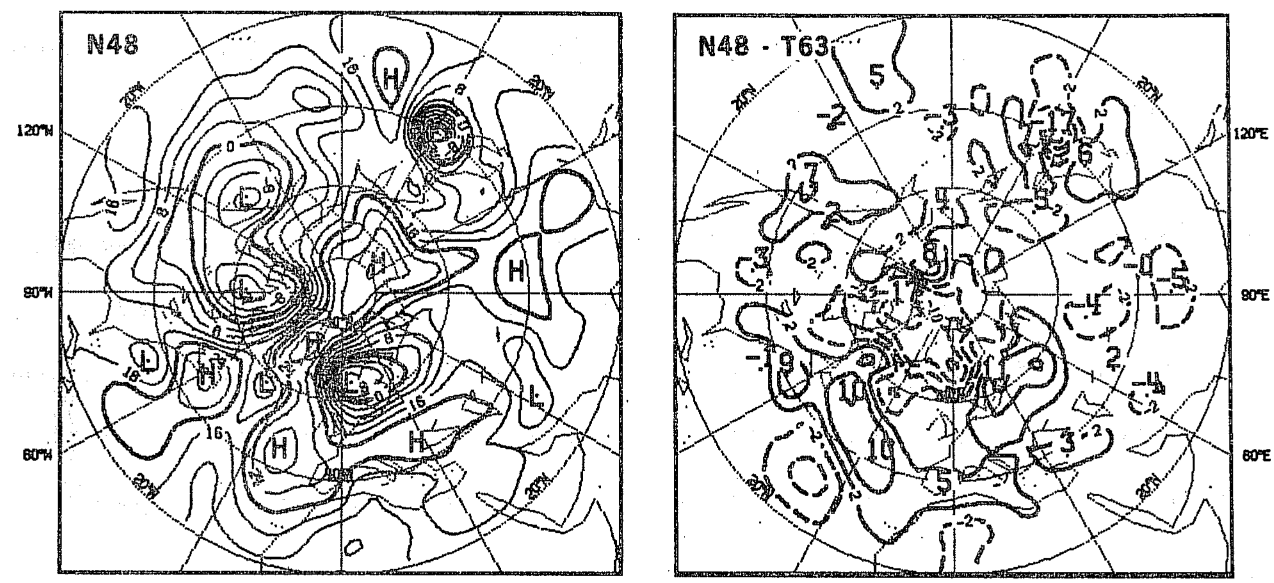


Fig. 16a (Contd.)

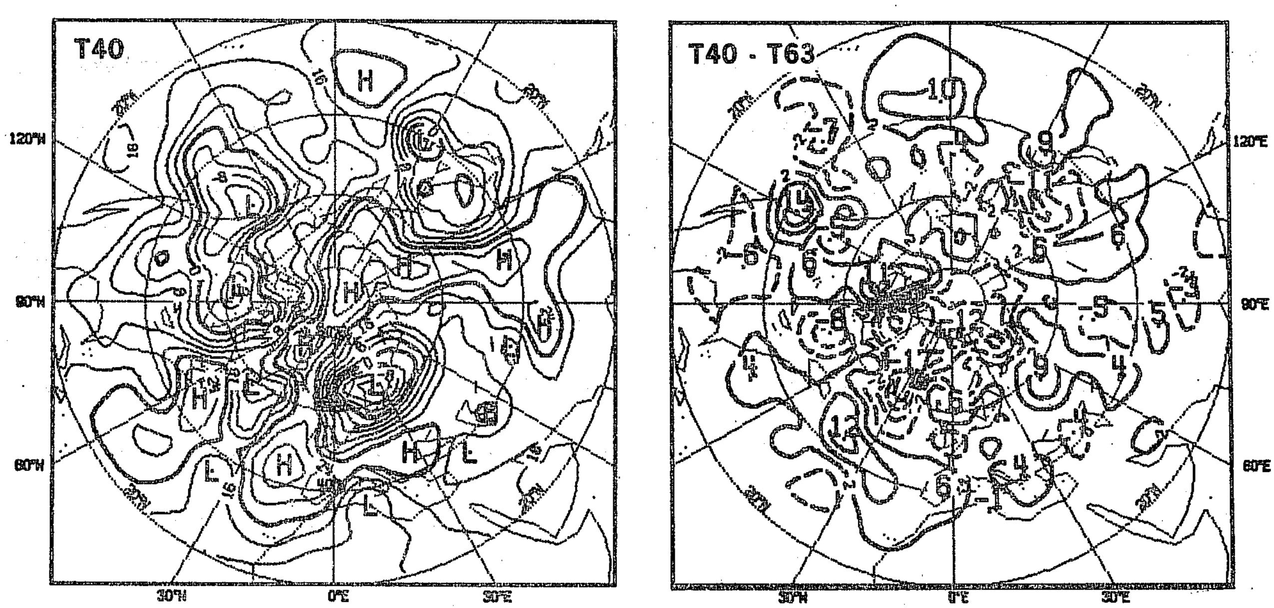


Fig. 16b JAN. 10, Z1000: D+4 (T40, T40-T63)

At 500 mb, the very intense trough ridge-trough pattern over the north eastern Pacific is much better captured by T63 (it is, in fact, almost nonexistent in N48). So also is the position and shape of the deep low over eastern Asia. The corresponding features in the 1000 mb height field show similar improvements by the spectral model. One can note in particular the more accurate representation of the multiple lows system over the northern Pacific. Most lows are however typically too deep.

For May 1st (Fig.13) besides phase differences (e.g. over the Pacific ocean), significant differences can be found in the simulation of the complex system of lows and highs extending from north east America across the Atlantic and Europe to North east Asia. Almost all of them are in favour of T63.

For August 21 (Fig.14) the largest difference occurs over the Arctic region where the position and shape of the polar low is better forecast by T63. Another large difference on D+4 is observed over the Atlantic where N48 is less successful in forecastng a cut off low (this slightly greater ability of T63 in simulating cut off was observed on several other occasions, often as here over the Atlantic region).

In the previous section it has been mentioned that the objective improvement by T63 was larger in the long waves (wavenumbers 1-3). The following example (Fig.15) is an extreme case where all the improvement came from the long waves (+9% in anomaly correlations for the group 1-3 and only +1% for the group 4-9). However, when looking at the N48, T63 and N48-T63 maps it looks more as if a local feature near the Greenwich meridian has been improved and not much elsewhere. In terms of Fourier analysis, such a sharp difference should have an almost flat spectrum and the reason why it stands out only in the scores for the long waves corresponds to the fact that these are the only ones with some skill left. So extreme care is needed when speaking of long wave improvements.

Finally, to conclude this section on synoptic evaluation, Fig.16a show the D+3 1000 mb height forecast from January 10th which corresponds to the largest score difference in favour of N48 at that range.

There is no doubt that the synoptic evaluation is also very much in favour of N48; in particular over the Atlantic in the simulation of the trough ridge trough system: T63 has merged a cut off low over the Atlantic with another low east of Canada. The high near Greenland is too intense and too much to the north and the low over Northern Europe is too broad. In all these respects N48 performs better. We shall come back to this case in Sect.2.5, but we would already like to mention that the case was rerun with a spectral model using a coarser resolution (triangular T40) which produced a forecast remarkably similar to that of N48, precisely in these areas of large differences in favour of N48. (Fig.16b).

4. COMPARISON OF THE MEAN BEHAVIOUR OF THE MODEL

For this section, the forecasts have been grouped and averaged by season (e.g. AUT corresponds to September-October-November) leading to 13 or 14 cases in each seasonal group.

Fig.17 displays the mean errors of the 500 mb height field for D+10. The similarity of the error patterns of N48 and T63 is striking in all seasons despite large seasonal variations. Thus the differences shown in the previous section were essentially transient.

Note again that both models tend to show overdeepening over Europe and the north Pacific ocean. At 1000 mb the errors and differences (not shown) are smaller, but their structure is very similar.

A more significant difference in the mean behaviour of the models is related to the level of eddy kinetic energy that they are able to maintain (Fig.18). Both models tend to loose eddy kinetic energy, in particular in the short wave, but N48 even more than T63. This supports the observation made in the synoptic evaluation that T63 tends to keep a higher level of variance. It seems to be true in all seasons more or less in the same way.

Another noticeable difference is found in the temperature field: both models tend to cool the atmosphere, mostly in the stratosphere at all latitudes, in the tropical troposphere and in the mid troposphere at middle latitudes (the global cooling in 10 days is of the order 2°K). Fig.19 shows that the tropospheric cooling is significantly reduced by T63 in particular in winter. It has been possible to relate this cooling difference to the different treatment of horizontal diffusion and it is interesting to note that the reduction of the diffusion coefficient of N48, mentioned in the introduction took place just before the Spring season and resulted in reduced differences between the two models.

This was further supported by diffusion experiments performed with T63: by increasing by a factor of 4 the diffusion coefficient of T63 it was possible to simulate almost exactly the cooling level of N48.

Other differences have also been found to be related to the horizontal diffusion, in particular in the level of zonal available potential energy. (not shown).

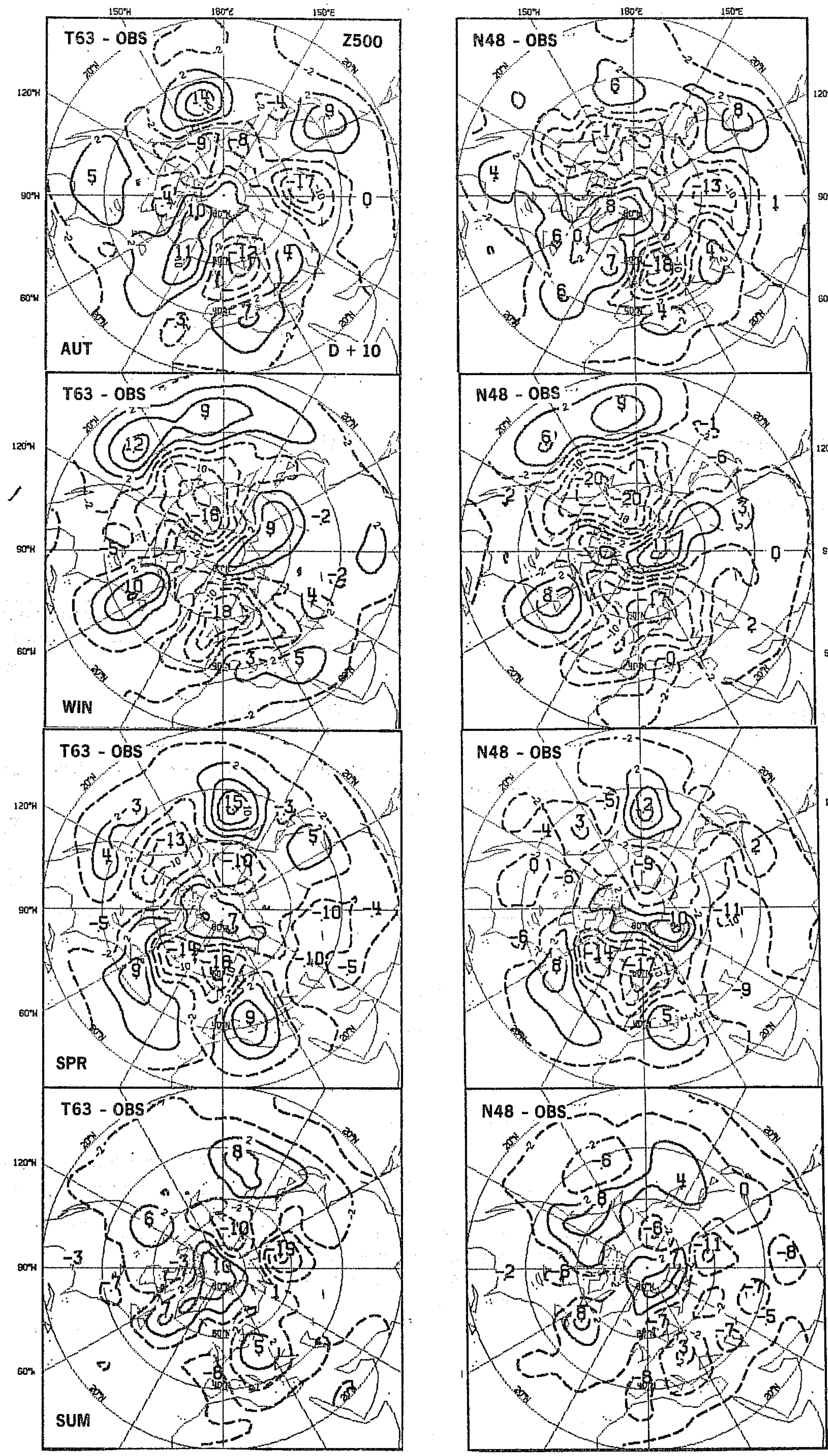


Fig. 17 Mean errors: Z500, D+10 for AUT,WIN,SPR,SUM(T63-OBS,N48-OBS)

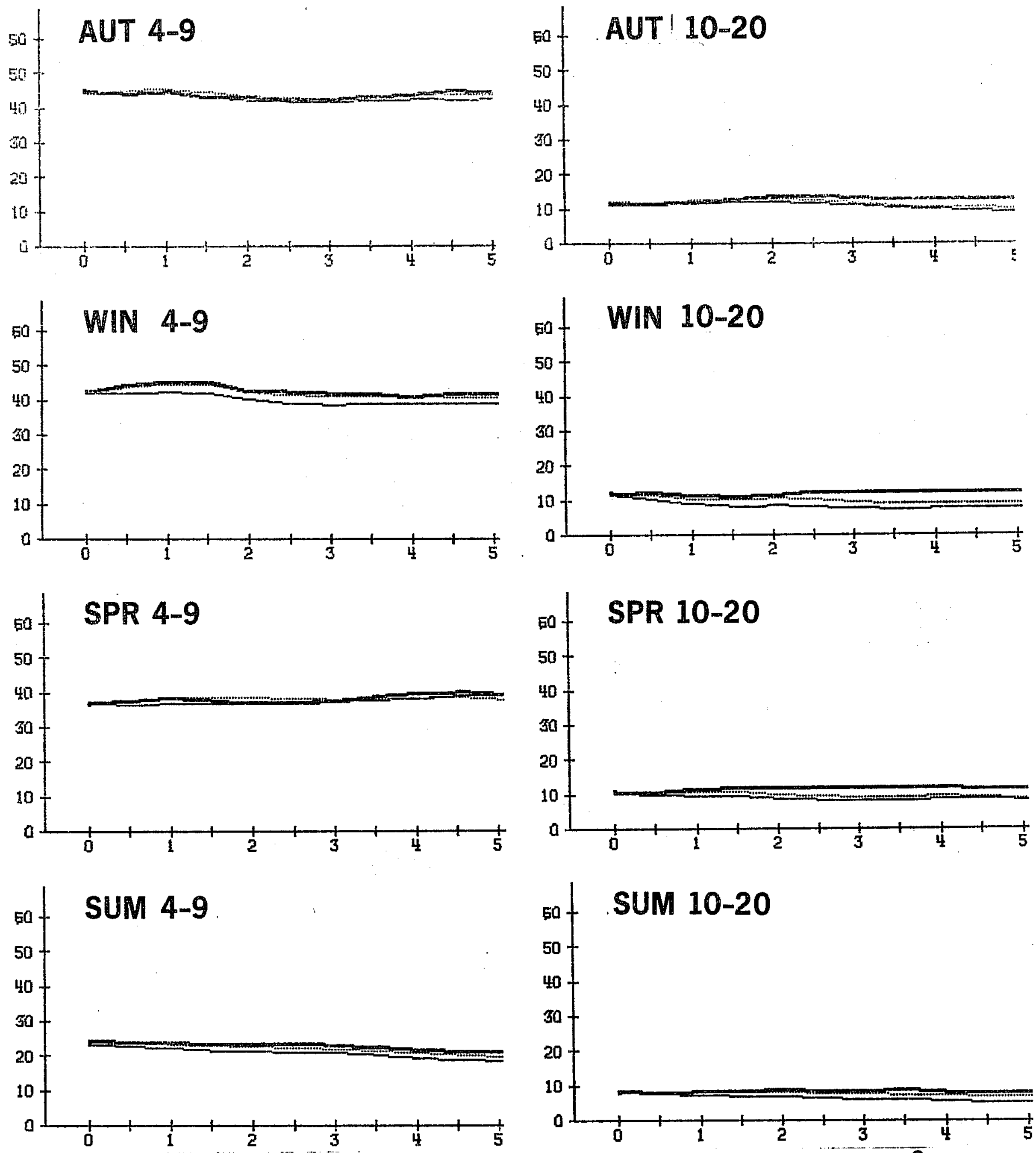


Fig.18 Time evolution up to D+5 of total averaged geostrophic KE(kJm^{-2}) in the troposphere (1000-200mb) between 20°N and 82.5°N for the wavenumber bands (4-9) and (10-20) in AUT,WIN,SPR,SUM. (OBS= ▬▬ ; N48= ———; T63= ------)

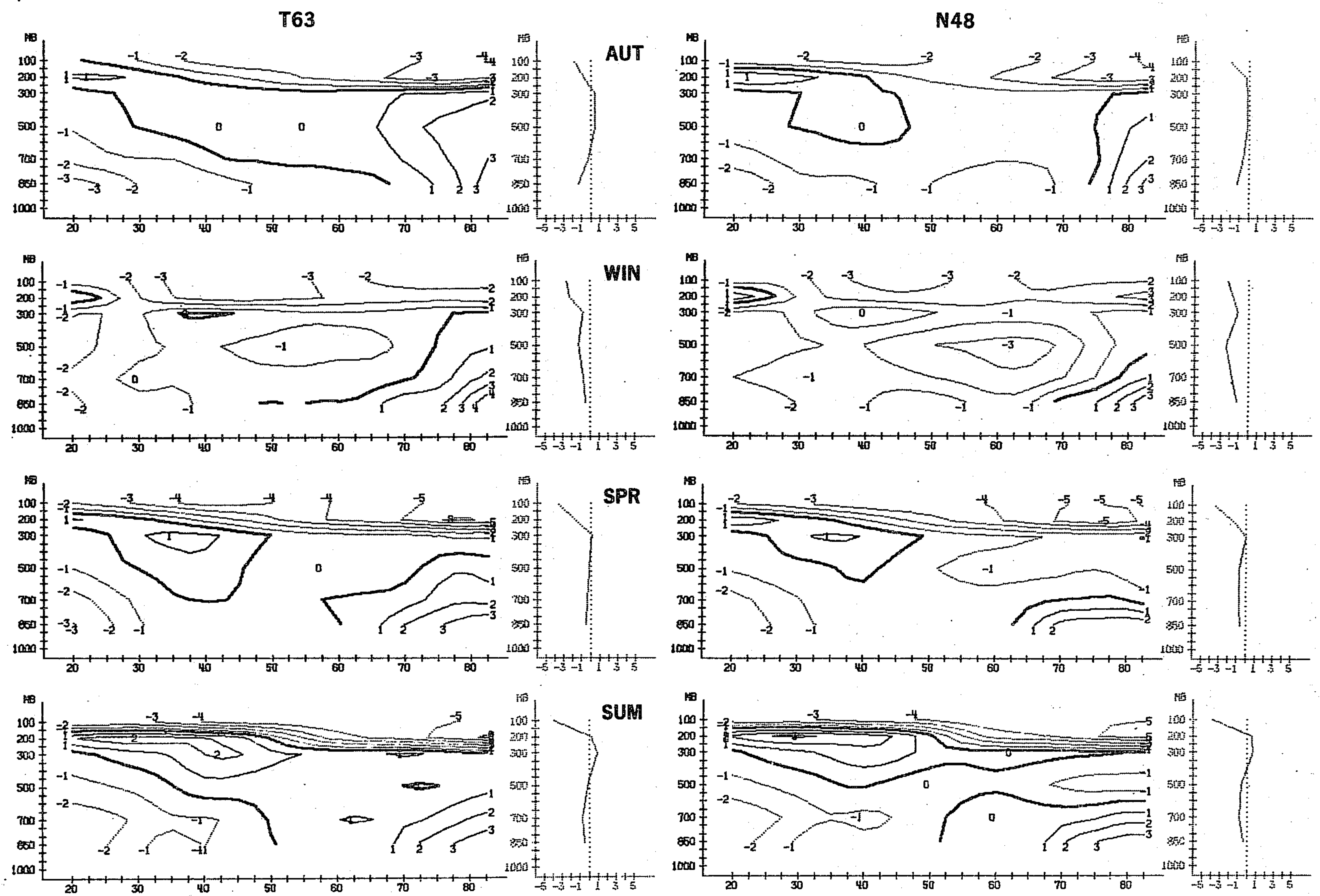


Fig. 19 Pressure-latitude cross-sections of the mean T63 and N48 errors of T(^{o}K) in AUT, WIN, SPR, SUM for the forecast period D+7 to D+10.

5. NATURE OF THE FORECAST DIFFERENCES

5.1 The resolution problem

The main difference between the two models lies of course in the techniques used to discretize the horizontal part of the equations, and our purpose is to try to analyse its effects. Other differences have also been found to play a role and have been analysed in more detail in Girard and Jarraud (1982): post processing of forecast fields (in the first 3½ months of the comparison), horizontal diffusion (as already shown for the temperature field), and coast line definition. This last item has been found to have only a small impact up to D+5 and therefore was not extensively investigated. It proved however to be important beyond D+5 on one case discussed in the following. When comparing two techniques it is also important that the resolutions used by the two models are comparable, in order to ensure that the truncation error differences are small. In our case N48 and T63 have fairly high resolution, significantly higher than in previously reported comparisons. More precisely we can consider that in terms of number of degrees of freedom N48 is equivalent to a spectral model with a rhomboidal truncation at wavenumber 96 (R96) and thus the differences N48-T63 can be regarded as the resultant of differences N48-R96 (technique differences) minus R96-T63 (resolution differences).

Results of comparisons between N48, T63 and T40 (the spectral model with a coarser triangular truncation) showed that T40 could compete successfully with N48 up to D+2 or D+3, the extra truncation errors becoming an important contributor to the total error only beyond D+3. Thus T63 truncation errors being significantly smaller than the ones of T40, due to the fast convergence of the spherical harmonics expansion, are also likely to play a lesser role than the technique differences. This is further supported by some higher resolution experiments (T79 and T96) not reported here. Furthermore these different truncation errors should favour N48 and therefore contribute to underestimate the contribution of the technique improvements by T63 in the total difference N48-T63.

5.2 Systematic phase differences N48 "linear" phase error

A crucial component of atmospheric dynamics is advection, and thus technique differences should be analysed accordingly. In particular as mentioned elsewhere in these proceedings, discretization of derivatives with finite differences usually leads to systematic linear advection errors depending on the scale of the phenomena and on the mesh size. Atmospheric circulation from a mathematical point of view is highly non linear. However, in practice over large regions of space and time the mean wind is reasonably constant and linear concepts may be used to characterize finite difference errors in advecting small scale systems.

It is fair to say that overwhelming evidence has been observed in the synoptic evaluation that in the majority of cases well defined system in N48 are lagging behind those of T63 (e.g. see Figs.7,8,9,10,13).

Table 1 gives some statistics for all well-defined lows between D+1 and D+2: for slow moving systems both models are too fast and N48 slightly more than T63. (This is associated with the general failure of the models to slow down sufficiently the cyclones in their mature stage).

In addition, both models are clearly too slow for fast moving systems, but N48 substantially more than T63.

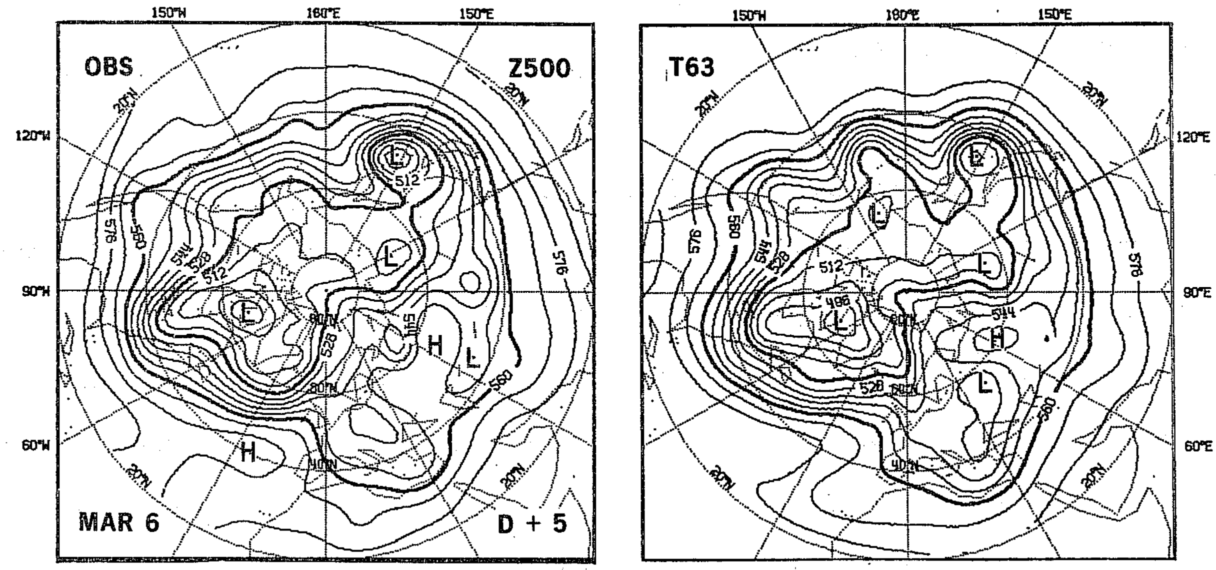


Fig. 20a MAR 6, Z500, D+5

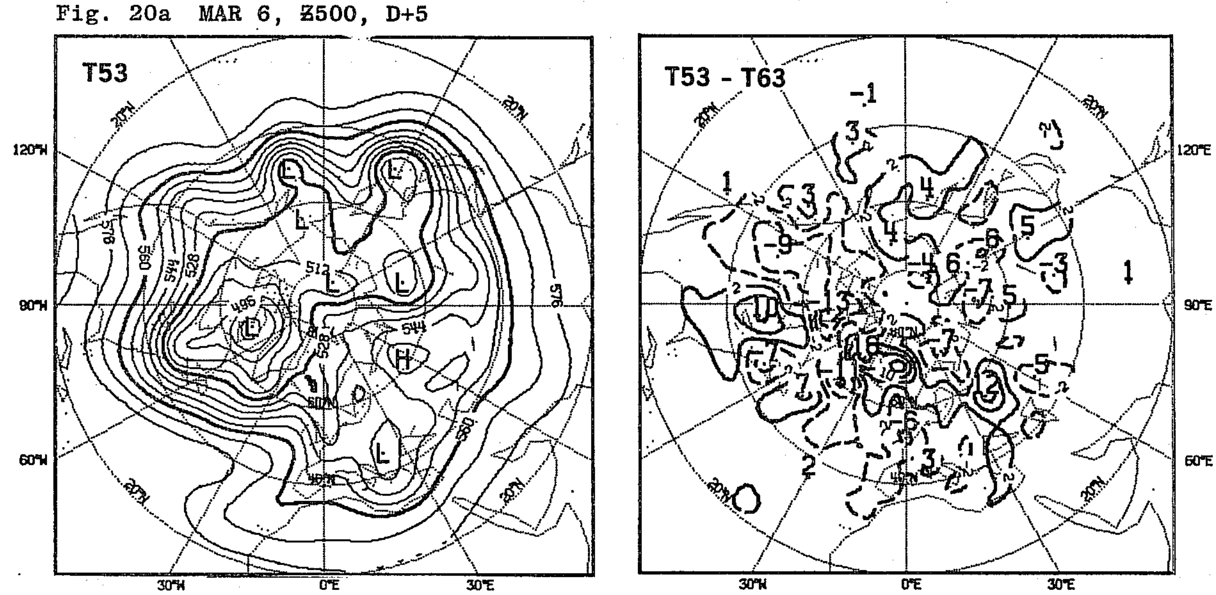


Fig. 20b MAR 6, Z500, D+5 (T53, T53-T63, N48-T53)

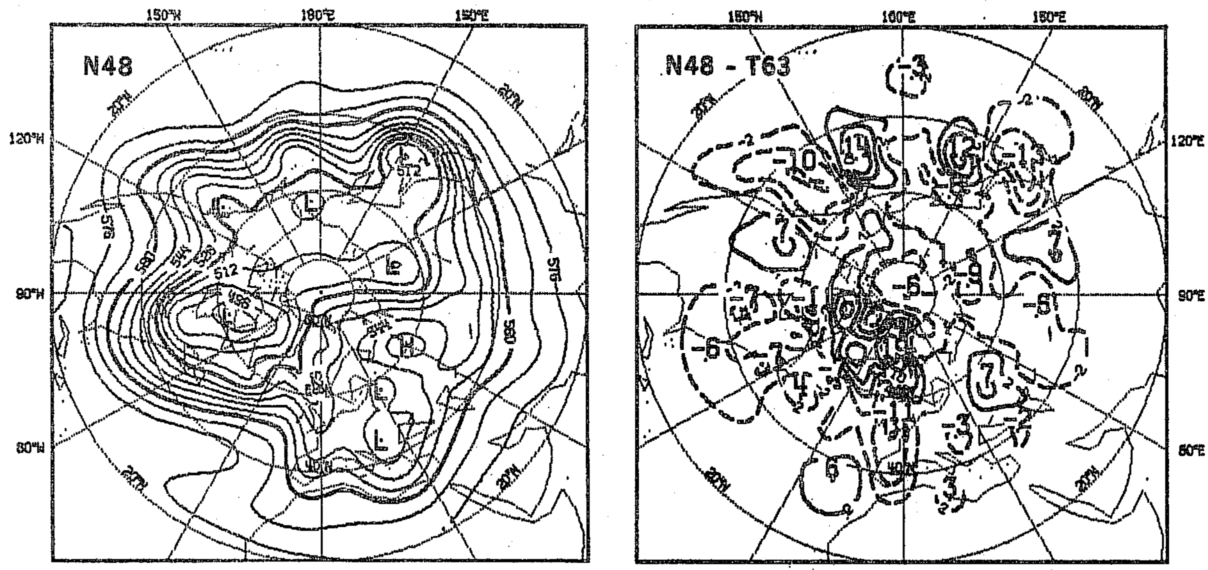


Fig. 20a (Cont.)

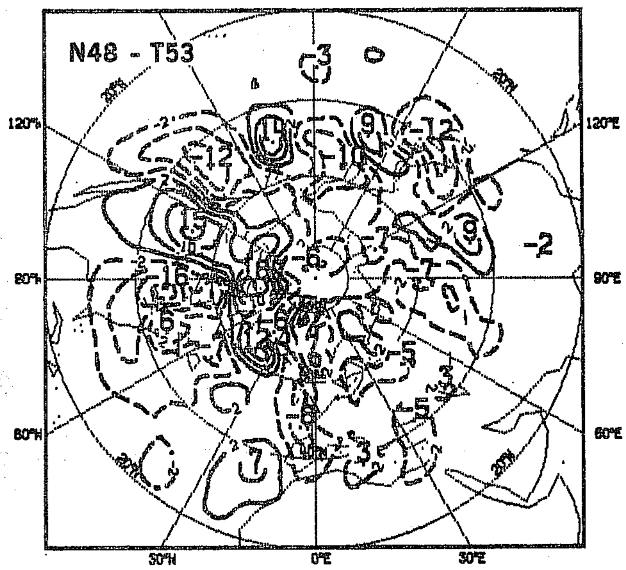


Fig. 20b (Cont.)

Table 1 Phase error statistics

D	Number of lows considered	E(degrees) T63	E(degrees) N48	L(km) T63	L(km) N48
all lows	192	-.55	-.82	-43	-64
D < 5°	64	+.64	+1.02	50	80
5° ⩽ D < 10°	49	+.33	+.23	26	18
10° ⩽ D	89	-1.80	-2.60	-142	-204
15° ⩽ D	44	-1.82	-3.32	-143	-261
20° ⩽ D	16	-2.94	-4.50	-231	-354

D = eastward displacement (in°) of the analysed low between D+1 and D+2

E = mean error (in °)

E > 0 if system was forecast to go too fast eastwards

L = mean equivalent error at 45°N (in km)

$$\left(L = \frac{2\pi\ R\ \cos 45°}{360}\ E \qquad R = \text{radius of the earth/km}\right)$$

A good illustration of almost pure linear phase errors is given in Fig.20 for the 6th March case (D+5 forecast). It displays forecasts by T63, N48 and also T53 (the spectral model with a triangular T53 truncation): the N48-T63 difference shows a string of positive-negative centres aligned along the 500 mb jet from eastern Asia to western Atlantic which is a typical signature of these differences. When comparing N48 and D53 the phase difference signature is even more typical, while the differences between T53 and T63 (smaller on average) have a much more complex structure.

There can be many causes to the fact that for fast moving systems, both models are too slow, e.g. parameterization errors, inaccuracies in the vertical, and time discretization schemes,... However, phase differences must be related to technique differences, and therefore to differences in the treatment of the dynamical terms. In order to prove this point, the March 6th case was rerun with both N48 and T63, but using a limited physical package (just surface friction and dry adiabatic adjustment). The corresponding 500 mb height difference maps (not shown) exhibited almost exactly the same signature along the jet from eastern Asia to north America.

Besides pure phase differences (along a cyclone track) there are also track differences. Most tracks curve northward, but in a majority of cases, N48 cyclone tracks are to the south of those of T63. (e.g. Figs. 7,8,12 and several other figures in Girard and Jarraud - 1982). Among the possible contributing factors, two can be directly related to the linear phase errors in N48. First the wave packets are not isotropic and therefore different errors arise in the EW and NS direction, but these should not be such as to curve the tracks less to the north on average. Second and more important, the N48 grid is not isotropic and for a single wave moving north eastwards at about 45°N, the phase error will be about twice as large in the NS than in the EW direction leading to a more southern track for N48. (assuming a second order scheme with phase errors proportional to the square of the grid size).

It is also worth noting that no such tendency for track differences was observed between spectral models of various resolutions.

These phase like differences have been found to be very similar at 1000 and 500 mb, although often larger at 1000 mb (in relative terms, compared to the total variance) which is consistent with the larger score improvements near the surface seen in the objective evaluation.

5.3 Non linear errors

A classical non linear error associated with finite difference techniques is the aliasing error occurring when non linear interactions give rise to scales which cannot be resolved by the finite set of grid points. However for resolutions such as N48 it is unlikely to be large, due to the very small amount of energy in the smallest scales (further reduced by the strongly selective horizontal diffusion used) compared to that in the longer waves. In spectral models, aliasing is generally avoided for linear and quadratic terms.

Besides aliasing, other errors which can be given the general name of coupling errors (cf the general presentation on spectral method earlier in these proceedings) occur as a consequence of the finite difference representation of non linear differential operators. A number of schemes have been designed in order to satisfy some integral constraints and to minimize the average errors. But the individual interactions always remain somewhat misrepresented.

We have already demonstrated the impact of "linear" errors in N48. We would like here to emphasize the importance of the more non linear ones. The distinction made here between linear and non linear is rather arbitrary: most of the linear phase differences shown previously are in fact likely to correspond to erroneous non linear interactions between small scale systems and the very large components of the flow, such that the flow can be regarded as locally constant and where linear theory can provide a reasonable first order approximation.

Thus the presence of these "linear" errors guarantees the presence of non linear ones which should be of a similar order of magnitude and are more likely to be observed in regions (of space and time) where the circulation changes abruptly. As for linear errors they should lead on average to improvements of T63 over N48, and if we accept that resolution and other differences between N48 and T63 play a lesser role they must account for the "non linear" forecast differences observed in the synoptic evaluation: e.g. Figures 10,11,12,16 and 20 (over the Atlantic).

These differences, whether or not they are in favour of T63 often appear as a lack of effective resolution for N48. An example was shown in Sect.3 (Fig.16). Fig.21 further emphasizes this point by showing the D+5 1000 mb height forecast from January 10th case for N48, T63 and T40. The similarity between N48 and T40 over the Atlantic is remarkable and is clearly demonstrated also by the difference maps. A similar tendency again over the Atlantic can be seen between T53 and N40 for the March 6th case (Fig.20).

The differences between T63 and N48 mentioned so far should all favour T63, except resolution. Therefore in order to investigate whether resolution could be responsible for the cases when N48 performed better than T63, these cases were rerun with a coarser resolution spectral model (T40 or T53).

Most of the time (there was only one noticeable exception discussed below) T40 (or T53) did as well and often better than T63 with many synoptic features developing more like those of N48. This, we believe, further strengthens the assertion that N48 non linar interaction errors contribute to reduce its effective resolution. A similar conclusion had been reached in a previous series of experiments (Jarraud et al. 1981) where N48 had been found to be on average equivalent to a spectral T53 model.

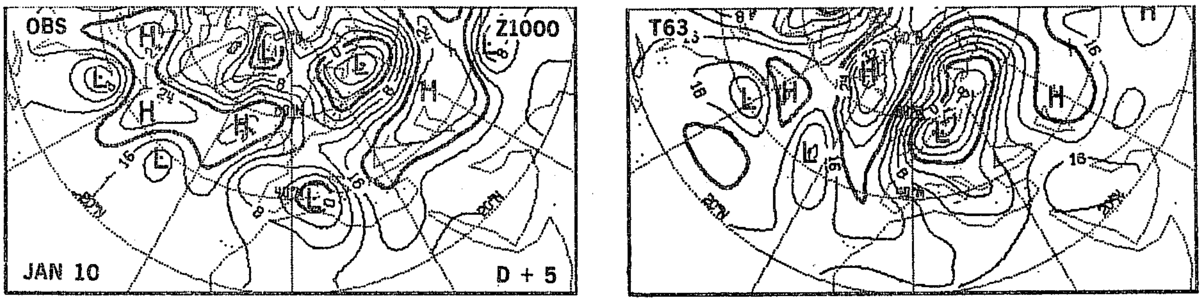


Fig. 21a JAN 10, Z1000, D+5.

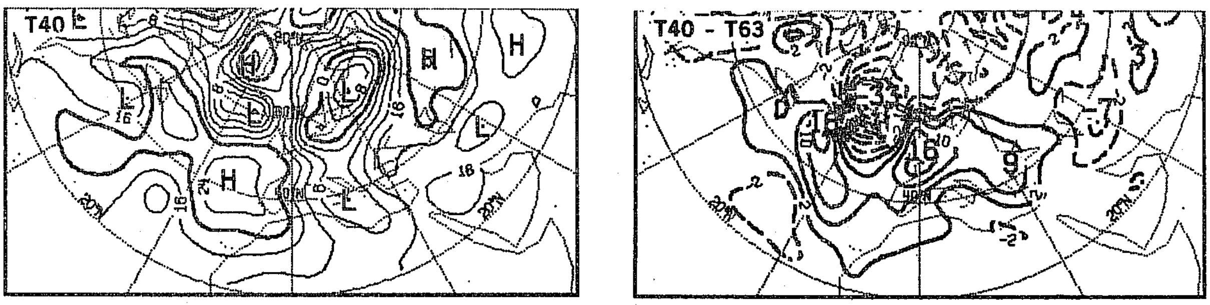
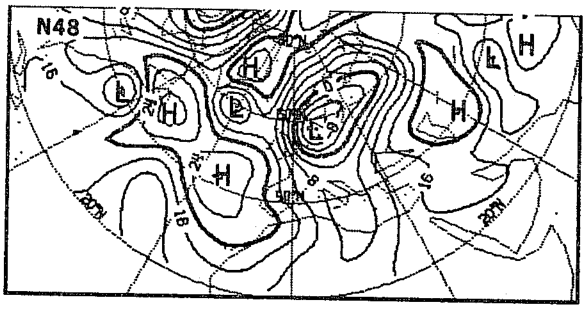


Fig. 21b JAN 10,Z1000, D+5 (T40, T40-T63, T40-N48)



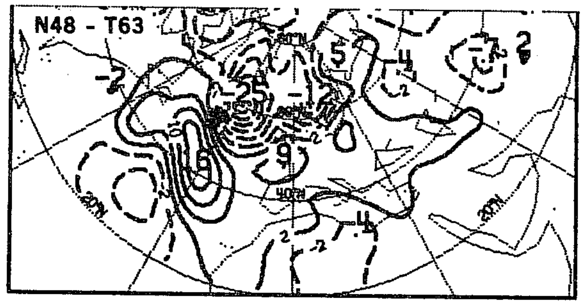


Fig. 21a (Cont.)

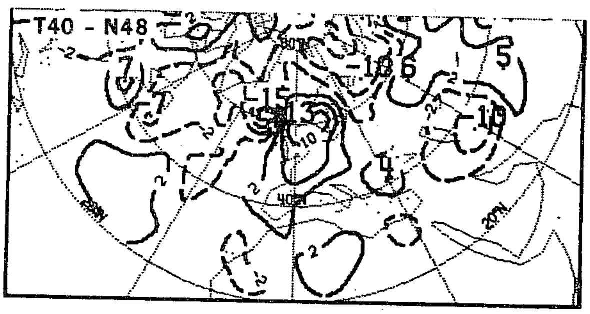


Fig. 21b (Cont.)

5.4 Other differences

In the previous paragraph we mentioned an outstanding case in this comparison. For the 14th February case, N48 and T63 produced a very good forecast, but N48 even better than T63. An experiment with a spectral T40 model produced a substantially worse forecast (more than a day less predictability). The main difference between N48 and T63 was in a low spuriously developed by T63 over South eastern Europe (Fig.22). Experiments with different horizontal diffusion coefficients did not yield any conclusive evidence. The question was therefore: could it be that we see in this case a positive impact from the higher resolution of N48? There was however a last difference between the models which had been found to be only marginal up to D+5: the land sea mask was not defined in the same way for both models, leading to non-negligible differences in particular for the Mediterranean Sea. An experiment was therefore carried out with N48 and the same coast line definition as for T63. The results (labelled C48 in Fig.22) are much more similar to T63 over Europe and the large score difference in favour of N48 has more or less vanished.

6. FURTHER EXPERIMENTS

Since the end of the comparison reported here many new experiments have been carried out with the spectral model. They basically confirm the previous findings and bring them stronger statistical significance. However they show larger (on average) deviations than the ones reported above. This may partly be due to natural atmospheric variability but also to improvements brought in the meantime to the ECMWF forecasting system (analysis, physical parameterizations,...). By reducing the other sources of errors they allow the impact of technique differences to stand more clearly out. In particular it is worth saying that the strong overdevelopments mentioned earlier (and stronger for the spectral model) have been considerably reduced.

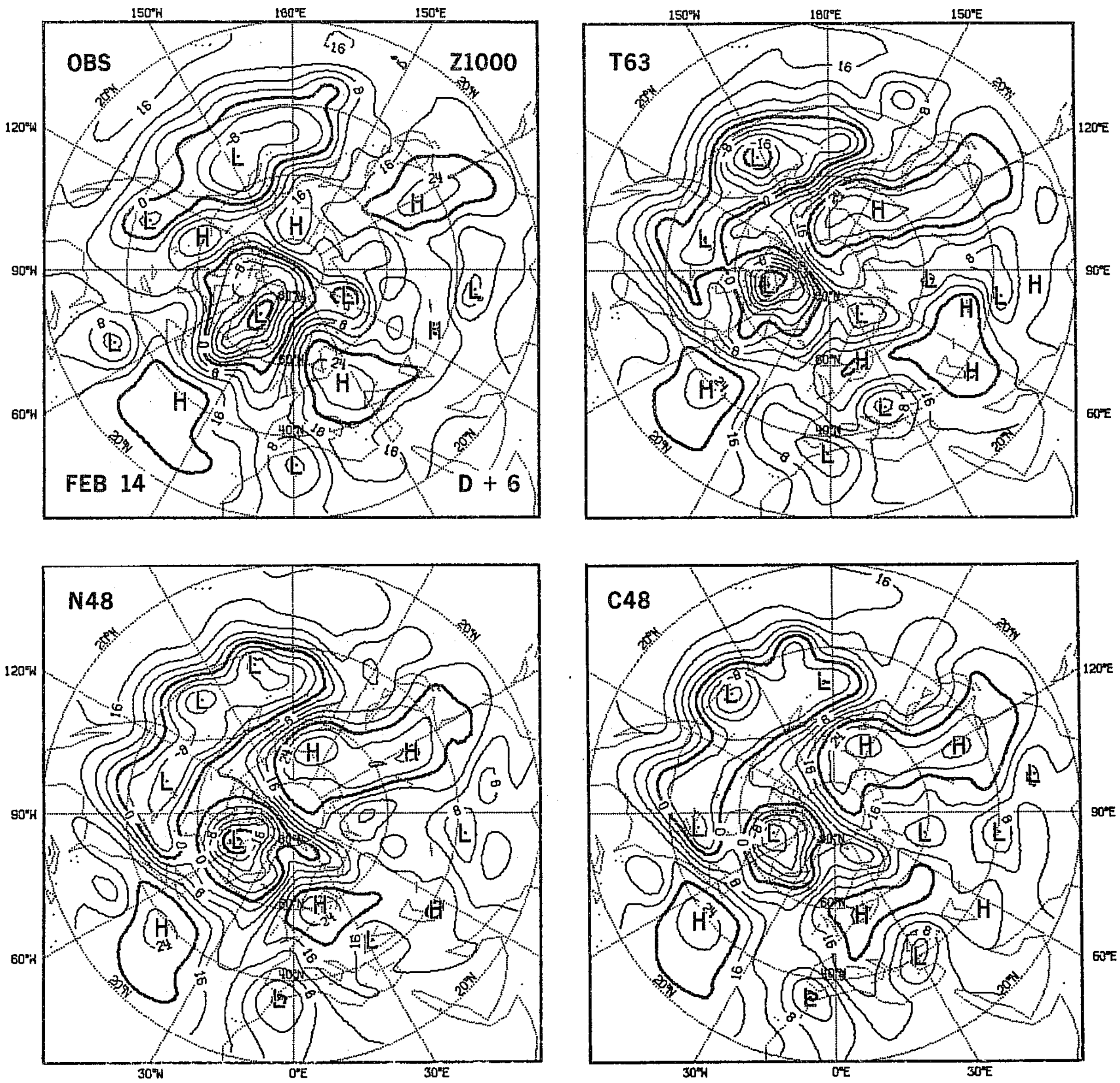


Fig. 22 FEB 14, Z1000, D+6 (OBS, T63, N48, C48)

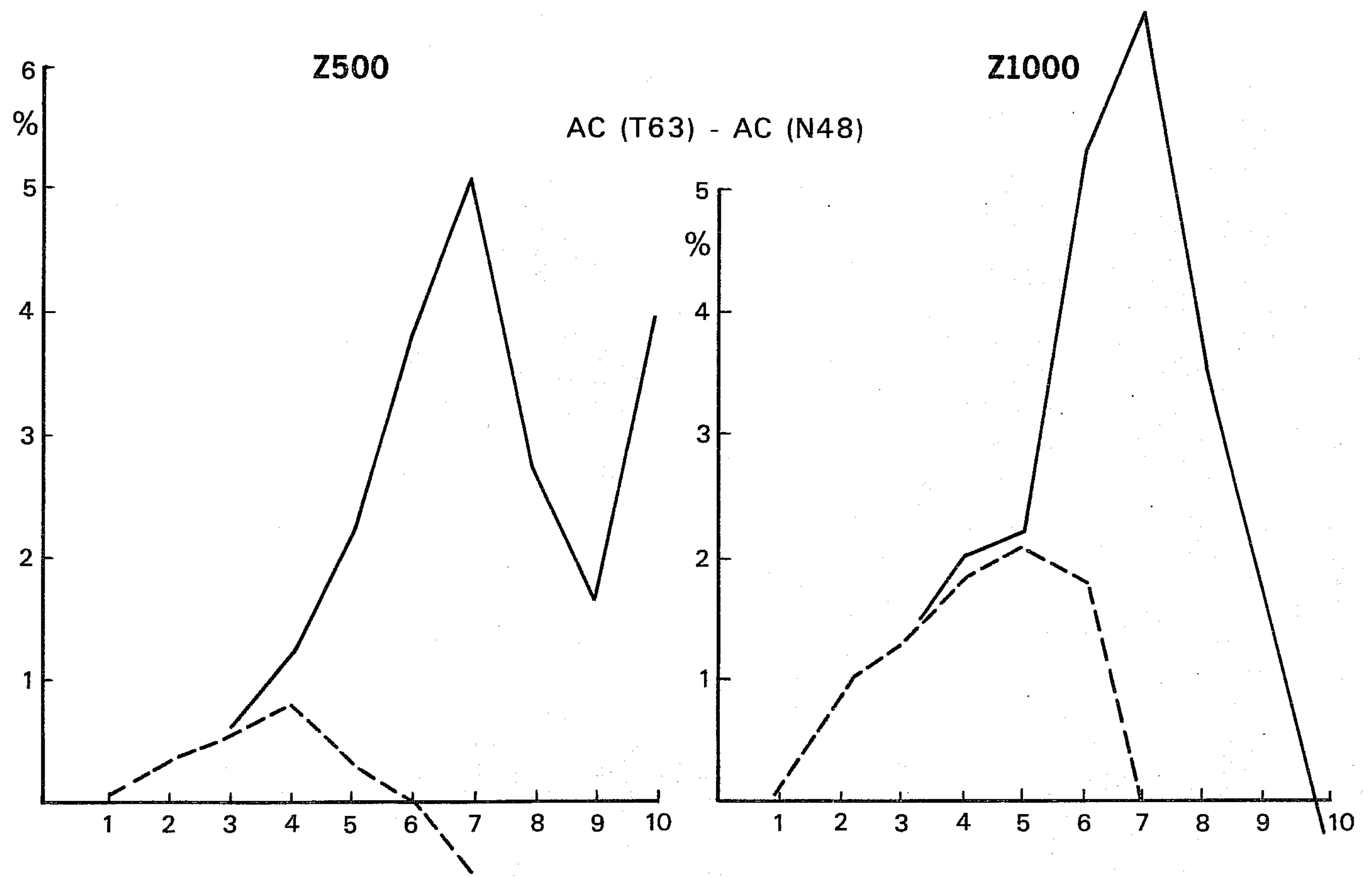


Fig. 23 ——— =13 WINTER 82/83 cases - - - - =53 cases comparison (79-80)

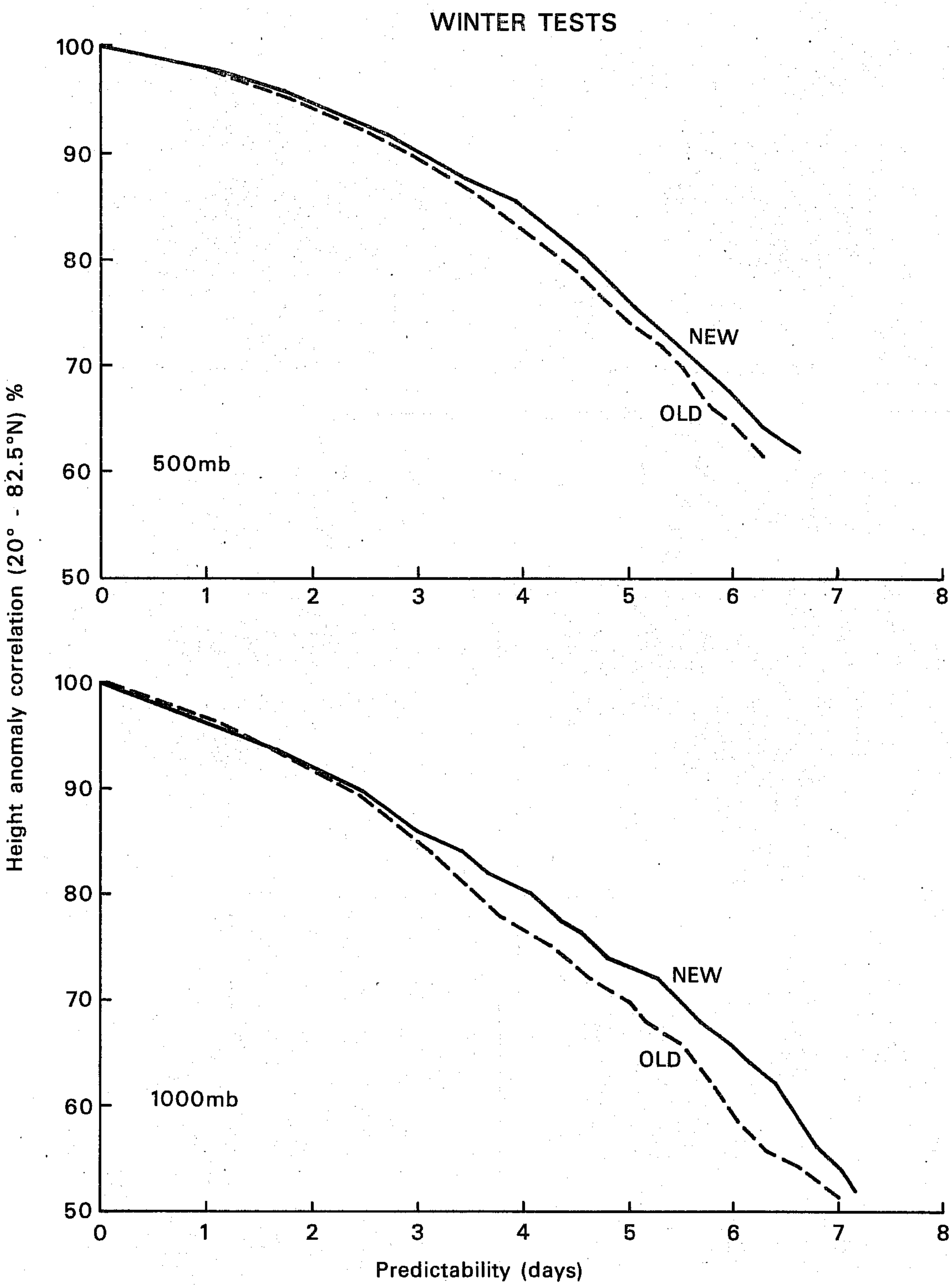


Fig. 24 Predictability as a function of anomaly correlations at 500 and 1000mb between old and new operational model.

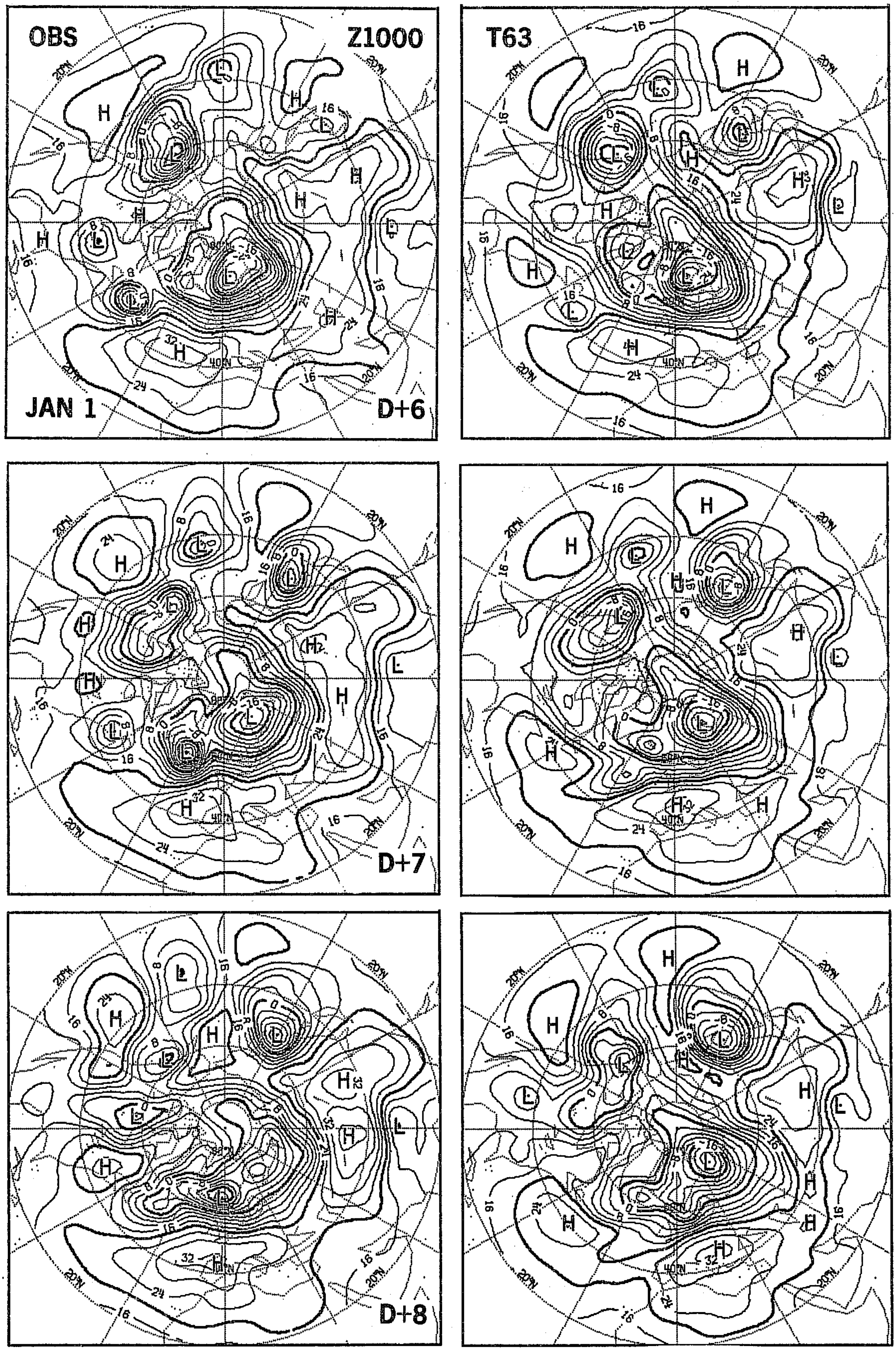


Fig. 25 Jan 1, Z1000, D+6 to D+8 (OBS, T63, N48, N48-T63).

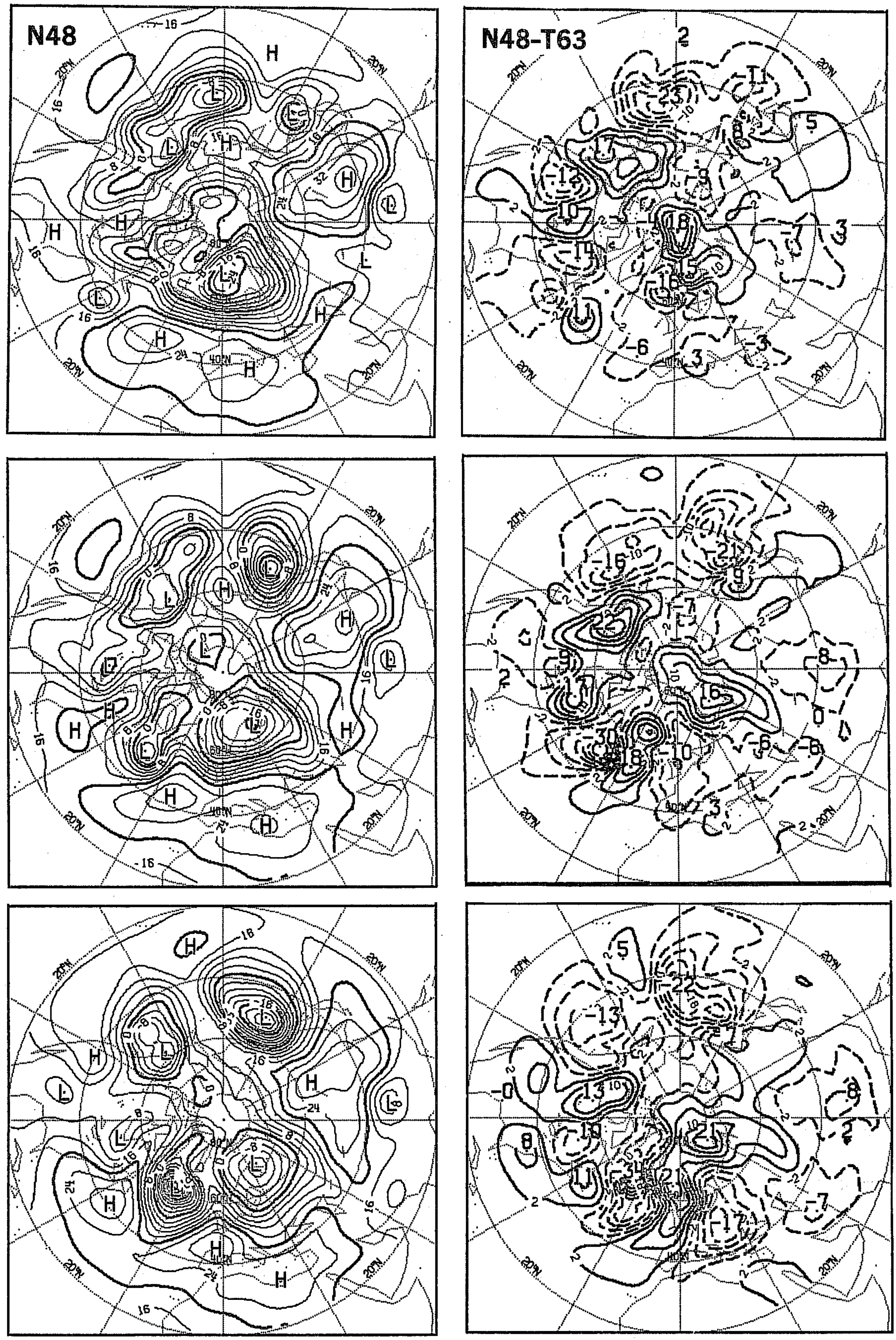


Fig. 25 (cont.)

Most of these comparisons were as clean as the main one. This was not the case for a set of 13 winter cases run when testing the new spectral forecast model. This had 16 levels and a hybrid vertical coordinate (Simmons and Strüfing, 1982) and slight changes to the formulation of the thermodynamic equation and to the physical parameterizations. However separate testing of these various differences has shown them to have a smaller impact than the N48/T63 technique differences. They can therefore safely be added to our sample to bring more weight to the conclusions.

As can be seen from Fig.23 the improvements in terms of A.C. (anomaly correlations) are still larger at 1000 mb than at 500 mb. At both levels they are much larger than in the main comparison and the advantage of T63 no longer vanishes after 6 days but rather after 9 to 10 days, which tends to strengthen the hypothesis that the sources of errors other than resolution and technique have been reduced.

In terms of predictability the improvement was very large with P(60%) improved by T63 by more than 24 hours in 4 cases out of 13 at both 500 and 1000 mb (the reverse being as in the main comparison never true). The resulting average improvement is displayed in Fig.24: in the range 60 to 80% it is more than 12 hours at 1000 mb and more than 6 h at 500 mb. A synoptic example is given in Fig.25. These differences are substantially larger than the ones found in the winter of the main comparison and one should also note that the average predictability was quite high, in agreement with the fact that T63 improvements are larger in situations of high predictability (see Sect. 2).

7. SUMMARY AND CONCLUSIONS

The results have shown T63 forecasts to be superior to those of N48 at all atmospheric levels, with a larger impact near the surface for the height field. It was also true at all scales but predominantly for the long waves. The improvement increased up to D+6, D+5 and then decreased to become statistically insignificant beyond D+7.

Despite their arguably small average amplitude, the T63 improvements proved significant both from an objective and a synoptic point of view. At 1000 mb for example the average 6 hours increase in predictability of T63, corresponded to a 24 h increase in almost 10% of the cases, the reverse being never true. Some examples of their synoptic relevance have been shown in Sect.3.

Since the models had been kept as close as possible to each other, it has been possible to show that the main source of forecast differences must have been pure technique differences, that is errors in the calculation of non linear terms in N48 (coupling errors) this including the common "linear" phase error, the principal component of which being due to the finite difference approximation of horizontal derivatives. This was supported by the type of forecast differences observed on daily height maps. The most common type was characterized by a phase lag of N48 systems with respect to those of T63, preferentially and in a more pronounced way in regions of strong steady rectilinear jets. They were often accompanied by track differences.

Non systematic differences developed on many occasions when there occurred larger than usual intensity or direction changes, in either space or time, of the large scale flow pattern. Often in such cases N48 behaved on a local basis like a spectral model of resolution lower than T63 and did not really provide for a different solution. In most such cases T63 performed better.

Besides their differences the models shared a number of common deficiencies: both underestimated the speed and deepening of young fast moving lows and tended to overdevelop more mature systems and not to simulate properly their decay phase. They also tended to develop similar "systematic" errors.

The results of the comparison were judged by ECMWF to justify the adoption of the spectral technique for its operational model in April 1983. More details of the design, implementation and performance of this new model are given in the following lecture.

Acknowledgements

We are particularly thankful to U. Cubasch for his decisive help in running the experiments, to A.P.M. Baede for his strong contribution in the design of the first ECMWF spectral model, to A Simmons for his invaluable support, advice and comments, to D. Burridge for many helpful discussions and to L. Bengtsson for his continuous encouragements during the course of the project.

References

Baede, A.P.M., Jarraud, M., Cubasch, U. 1979: Adiabatic formulation and organization of ECMWF's spectral model. ECMWF Tech.Rep.No.15, pp.40.

Burridge, D.M. and Haseler, J. 1977: A model for medium range weather forecasting, adiabatic formulation. ECMWF Tech.Rep.No.4, pp.46.

Daley, R., Girard, C., Henderson, J., and Simmonds, I. 1976: Short-term forecasting with a multi-level spectral primitive equation model. Atmosphere, 14, 98-134.

Girard, C. and Jarraud, M. 1982: Short and medium range forecast differences between a spectral and grid point model. An extensive quasi-operational comparison. ECMWF Tech.Rep.No.32, pp.178.

Jarraud, M., Girard, C. and Cubasch, U. 1981: Comparison of medium range forecasts made with models using spectral or finite difference techniques in the horizontal. ECMWF Tech.Rep.No.23, pp.96.

Simmons, A.J., Hoskins, B.J. 1975: A comparison of spectral and finite difference simulations of a growing baroclinic wave. Quart.J.Roy.Met.Soc., 101, 551-565.

Simmons, A.J. and Strüfing, R. 1981: An energy and angular momentum conserving finite-difference scheme, hybrid coordinates and medium-range weather prediction. ECMWF Tech.Rep.No.28, pp.67.

Temperton, C., and Williamson, D.L. 1979: Normal mode initialization for a multi-level gridpoint model. ECMWF Tech.Rep.No.11, pp.91.

Tiedtke, M., Geleyn, J.-F., Hollingsworth, A. and Louis, J.-F. 1979: ECMWF model parameterization of sub grid-scale processes. ECMWF Tech.Rep.No.10, pp.46.