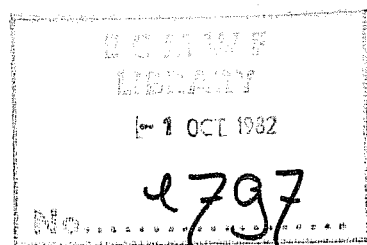# TECHNICAL REPORT No. 32

SHORT AND MEDIUM RANGE FORECAST DIFFERENCES

BETWEEN A SPECTRAL AND GRID POINT MODEL.

AN EXTENSIVE QUASI-OPERATIONAL COMPARISON

by

C. Girard and M. Jarraud

August 1982

C O N T E N T S   PAGE

Abstract

In all, 53 10-day integrations were made on a weekly basis with the ECMWF

spectral T63 model, based on ECMWF 1979-80 analyses covering a year.   These

were evaluated and compared with the ECMWF operational N48 grid-point model.

The results confirm, with an increased statistical significance, those of a

preliminary set of 7 integrations reported in ECMWF Technical Report 23

showing the superiority of T63 over N48.

Many examples of individual forecast differences are given, mainly for the

height field for the range 0-5 days, clearly demonstrating the meteorological

significance of the T63 improvements.

Finally the nature of the forecast differences is investigated in relation to

all main model differences. As far as it can be ascertained in this type of

model comparison, the most significant forecast differences that also lead to

T63 superiority come from the techniques used for horizontal discretization

in each model. More specifically, they arise mostly from errors in the

finite difference formulation of the non-linear advection operators in N48.

The predominant aspect of these errors is believed to be the inaccuracy of

derivative calculations which is very often manifested by a phase-lag of N48

systems with respect to those of T63, but may also lead to other important

forecast differences.

1. PURPOSES AND DESIGN OF THE EXPERIMENTS

The first purpose for making the set of forecast experiments reported here

was to test the medium-range forecasting ability of the ECMWF spectral T63

(triangular truncation with maximum wavenumber 63) model. Its performance

was to be evaluated in comparison with the present operational N48 as a

regular latitude-longitude ("C-grid" in Arakawa's classification and 48 lines

of points between pole and equator) grid-point model.

The integrations were made on a routine basis, from ECMWF analyses between September 1979 and August 1980, spanning for T63 the operational environment of N48. Regularly done once a week, they covered a whole year of synoptic situations (53 cases). It is worth mentioning the fairly uncommon conditions in which this quasi-operational comparison was undertaken. Such comparisons usually involve a newly developed model of more modern conception and a more ancient model in a well established operational set-up. Here neither model is more modern than the other. The comparison was initiated as a result of preliminary experiments described in Jarraud et al (1981), hereafter referred to as TR23, and started nearly as soon as ECMWF began its operational production.

In TR23, a set of 7 experiments was discussed and the main conclusions were: **T63 performs significantly better than N48. The improvement is already clear at short range, where even a coarser resolution spectral model (T40) competes successfully with N48. Beyond D+3 (day 3 of the forecasts), N48 is on average equivalent to a spectral model with resolution T53. The height field improvements of T63 over N48 are more pronounced near the surface and are concentrated in the larger scales.** But it was pointed out that more experiments were necessary to better understand the differences between spectral and grid point techniques.

This brings us to the second and main purpose of the present set of experiments: **to gather further evidence on FDs** (the expression Forecast Differences is used so often in the present report that it was found useful to abbreviate it) **and their nature in relation to theoretical technique as well as other model differences.**

Brief as well as references to more detailed descriptions of T63 and N48 have been given in TR23. In short, the two models are very similar except for the

technique used for the horizontal discretization . Development of suitable numerical techniques of integration has been one of the most important aspects of the numerical weather prediction problem since the pioneering attempt by Richardson in 1922. While grid-point methods were the first to be used and are perhaps at first intuitively more appealing, there is no doubt either that a great deal of mathematical afterthought has been necessary to make them work satisfactorily. Galerkin methods on the other hand and spectral methods in particular (Garp publication series No.17, 1979) despite their mathematical simplicity were introduced only later in numerical weather prediction. For the global (no horizontal boundaries) forecasting problem they present many advantages: spherical harmonics are solutions of the Laplace equation on the sphere and therefore make up a natural basis for a Galerkin approach. When meteorological fields are expanded in terms of properly chosen finite subsets of these functions (triangular truncation), their resolution is isotropic over the sphere (that is the truncated fields are the same, irrespective of any rotation of the coordinate axis). Their derivatives are calculated analytically and are thus exact (at the same truncation). Non-linear products can also be obtained without aliasing (in practice this will only be achieved for quadratic terms for efficiency reasons). Finally the convergence of the truncation expansion, as resolution is increased, is extremely fast: the truncation errors tend to zero faster than any finite power of $1/N$ when N (the maximum wavenumber retained) is increased (Orszag, 1974).

From a practical point of view, the development of the transform method (Eliassen et al, 1970 and Orszag, 1970) for handling the non-linear terms (including physical parameterization terms) has made spectral methods competitive with grid-point schemes in terms of efficiency. At present spectral methods are used for operational weather forecasting in Australia, Canada and the USA.

Concerning the design of the experiment there are two major differences between the present and previous sets of experiments (TR23). The first is related to verification analyses which in the present case are defined after application of a non-linear normal-mode initialization procedure (Temperton and Williamson, 1979). As a consequence, since final analyses are given on pressure surfaces and normal-mode initialization is made on σ-surfaces, resulting ECMWF analysis fields are smoother, due to both the initialization (insofar as it eliminates meteorologically genuine gravity-waves) and the vertical interpolation procedures. Furthermore, since the operational procedures used for the N48 model are used throughout this process, the final analysis fields may be biased in its favour. When post-processing the results of T63 it then becomes very desirable to minimize such biases, at least in relative terms, by using an approach that simulates the operational procedure as closely as possible. Although this was not realized at first, it was soon discovered during the course of the present experiments that relatively large and systematic differences originating from post-processing had a considerable impact on calculated objective scores, considering the close performance of the models.

The second difference in design is related to changes brought to the models' physical packages not only in the interval between the first and second sets of experiments but also during the course of the latter. These more fundamental changes were beyond our control, but all changes made to N48 were carried simultaneously into T63. Changes were however also made to the formulation of the horizontal diffusion scheme used by N48 on two occasions without parallel changes to T63.

All relevant modifications to the model, their physical and post-processing (PP) packages are given in Table 1.

Table 1. Time table of all the relevant changes made to the models, their
          physical and post-processing packages during the course of the comparison.

79/9/3   First T63 integration. The physical package is slightly different from
          the one used for the first set of 7 experiments reported in TR23. The
          post-processing package of N48 is different from the one used previously.
          (Operational post-processing = OPP). The post-processing of T63 is the
          same as before (Research post-processing = RPP).

79/11/25 First modification to the horizontal diffusion scheme in N48. The non-
          linear diffusion scheme used previously and described in TR23 is
          replaced by a linear 4th order scheme similar to that used by T63 but
          with a much larger coefficient.

79/12/20 From now on, T63 uses a post-processing package similar to the one used
          by N48(OPP).

80/01/16 Modifications to the physical package in both models.

80/02/29 Second modification to the horizontal diffusion scheme in N48. An
          implicit linear 4th order scheme with a reduced coefficient is
          implemented.

It is our belief that FDs between T63 and N48 which are related genuinely to

the use of the spectral and grid-point techniques respectively should come

mainly from large-scale dynamics terms, and can therefore be properly

investigated by looking at the forecast of the height field Z (and to some

extent the temperature field T) over the extratropical part of the Northern

Hemisphere where data are likely to be more reliable. For this reason, as in

TR23, we have limited our study of model performances (Section 2), of

synoptic differences (Section 3) and of mean differences (Section 4) mostly

to the 500 and 1000 mb height fields (hereafter referred to as Z500 and

Z1000) and to a smaller extent to the 500 and 850 mb temperature fields (T500

and T850 respectively). Having done so, we shall be in a good position to

discuss the nature and possible causes of FDs. We however acknowledge that

the evaluation of the forecasts in the Southern Hemisphere and in the tropics

(including the rainfall fields) would also be of interest, and all the results of the experiments can be made available for such extra studies. Besides the role played by pure differences of technique, the impact of the post-processing and horizontal diffusion differences will be considered. Additional (sensitivity) experiments carried out to strengthen our assertions in a number of directions will also be reported (Section 5).

The complexity of the models involved in the comparison necessitates qualitative proofs to support our conclusions (Section 6), and we hope that enough such evidence will have been given to make clear that the advantages of the spectral technique as borne out in short and medium range forecasts, may not be dismissed.

2.OBJECTIVE EVALUATION OF THE RESULTS (Northern Hemisphere)

2.1 Scores used for verification

We verify Z and T over the Northern Hemisphere between 20° and 82.5°N. As in TR23, we use the following three basic scores:anomaly correlation (AC) in %, standard deviation of the error (SE) in m and °K and predictability (P), in days or hours (h), in terms of AC. The first two are standard verification tools at ECMWF. P = P(c) represents the forecast range for which a model has been able to maintain a given level of quality, that is AC larger than c %.

To compare the individual performances of the models, we rely mainly on scatter diagrams showing AC(Z), since this score has been found to be more sensitive to FDs and to better agree with our synoptic assessments, but we also present various ensemble (monthly/seasonal) averages of height and temperature AC and SE.

2.2 Comparative performance of the models

An overall impression of the results is provided, by the scatter diagrams displayed in Fig.1 for the AC of Z1000-200 (vertically averaged score for Z between 1000 and 200 mb) respectively 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). The 60% lines are drawn as reference lines. At short range, the points are closely packed along the diagonal, indicating larger variability between cases than between models. On D+8 and D+10 a large dispersion is observed in both directions.

On D+3, the cloud of points is slightly above the diagonal on average. On D+4, the superiority of T63 is more pronounced and also quasi-systematic with only 10 points below the diagonal (less than 20%). On D+5, some very large deviations are apparent, most of them representing T63 improvements, particularly amongst the good cases (taking for example cases for which one at least of the models has a score above 60%). On D+6, there are already much fewer forecasts above than below 60%. The dispersion is still larger along the diagonal but it is increasing rapidly across it. The average difference between the models has ceased to get larger however indicating that its statistical significance is already doubtful. By D+8 and D+10, the sample size is not sufficient, according to the very low forecast skill, to consider the score differences as significant.

Similar scatter diagrams are shown in Fig.2 comparing Z500 and Z1000 and in Fig.3 comparing, for Z1000, the large (1-3) and medium (4-9) wavenumber components of the flow on D+3, D+4 and D+5. The superiority of T63 is smaller at 500 mb than at 1000 mb where very large differences are already seen on D+3 (most of them in favour of T63). This suggests that the maximum T63 improvement occurs near the surface and decreases with altitude. The improvement is mainly concentrated in the long waves (Fig. 3) which usually are the most energetic and most predictable components of the flow. Some
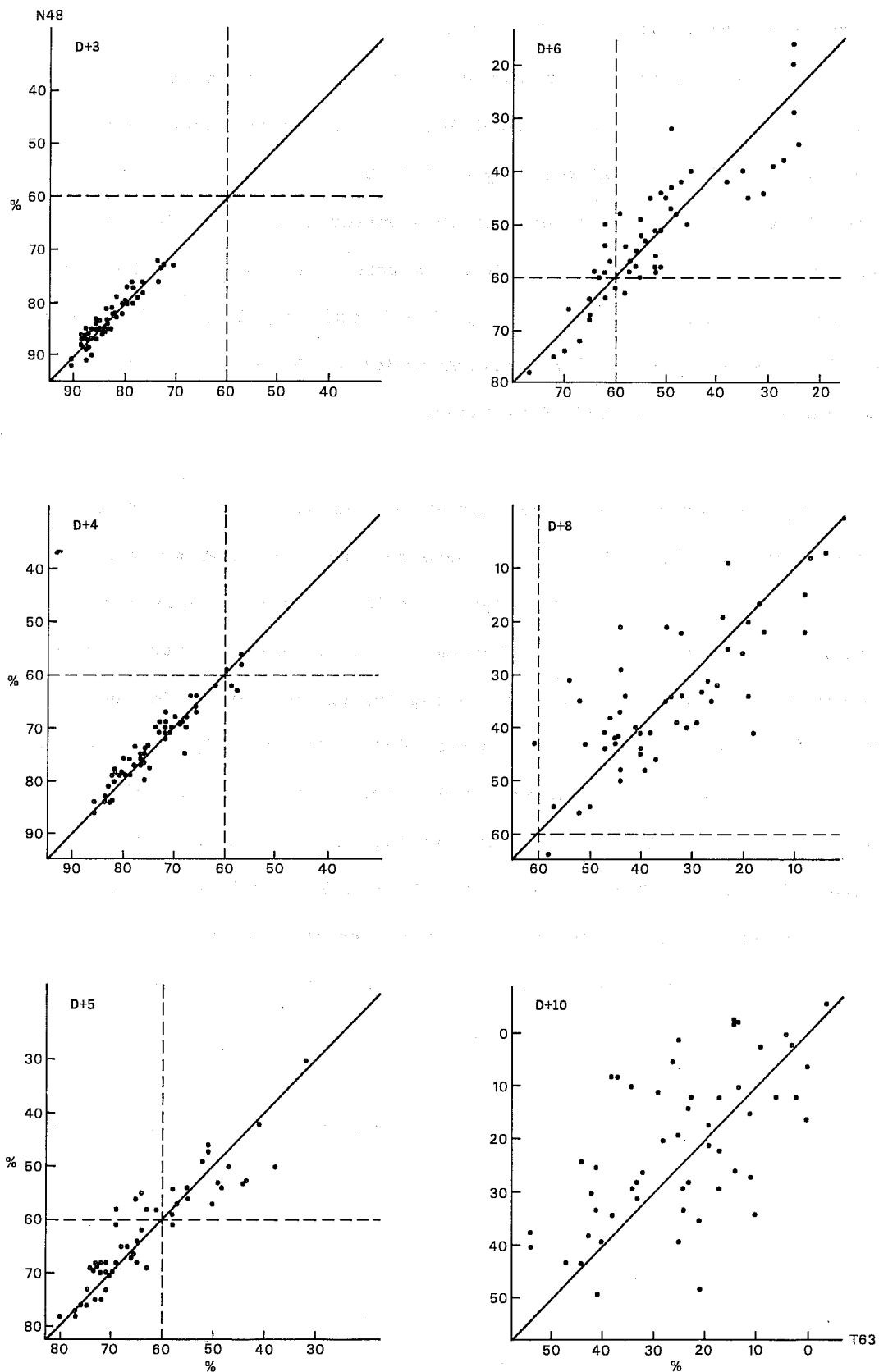
Fig.1 Scatter diagrams showing AC(Z) on D+3, D+4, D+5, D+6, D+8, D+10 averaged
       horizontally from 20°N to 82.5°N and vertically from 1000 to 200 mb,
       comparing 53 forecasts by T63 (horizontal axis) and N48 (vertical axis).

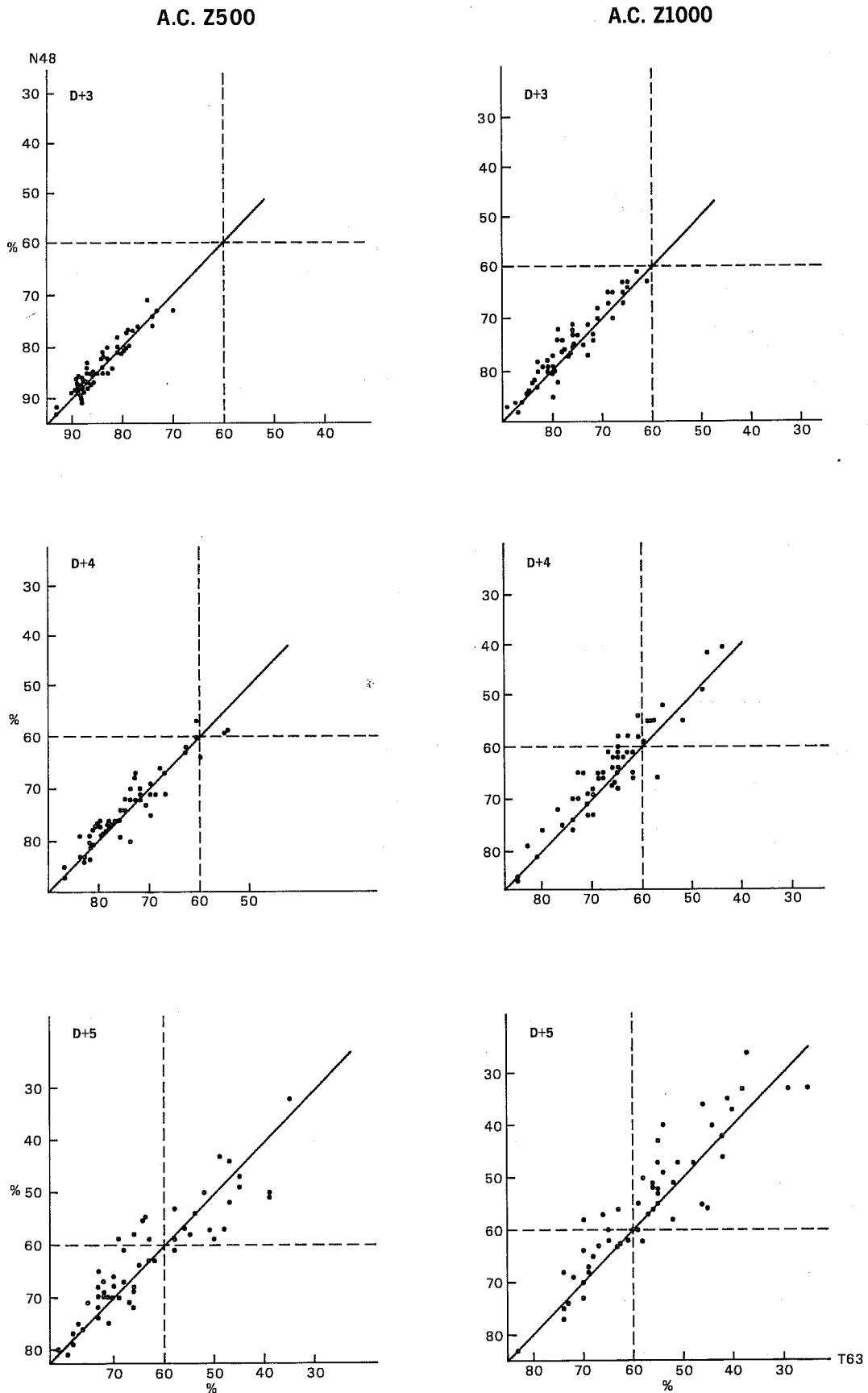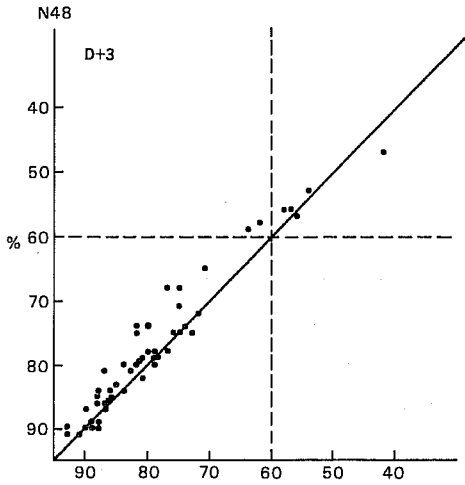8

A.C. Z500

A.C. Z1000

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).
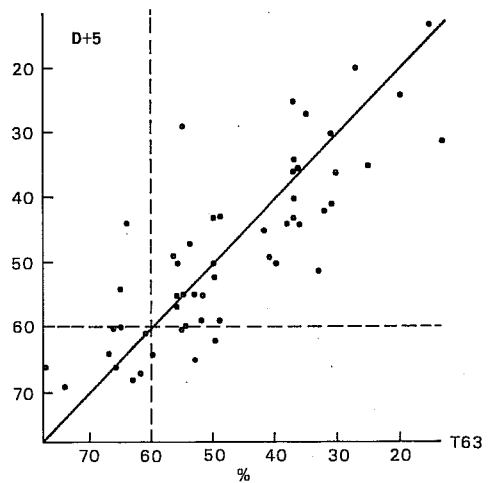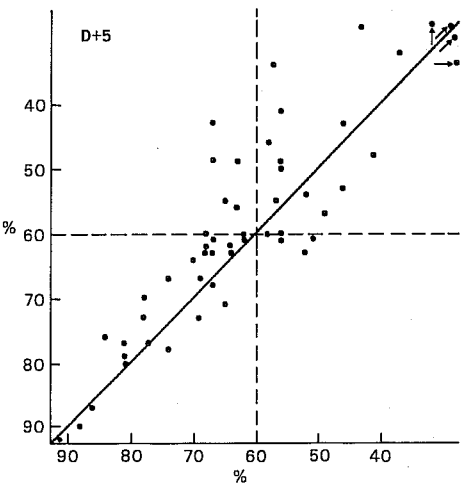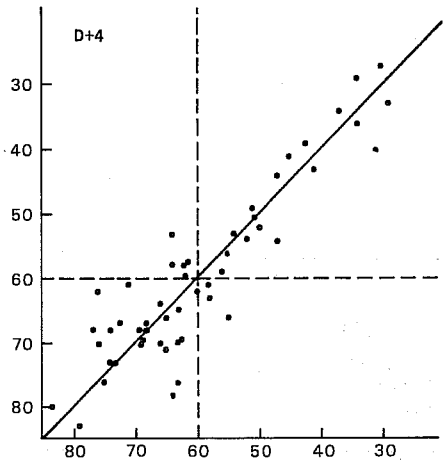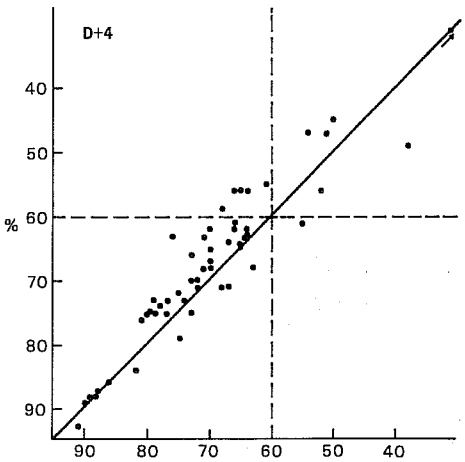
A.C. Z1000 1.3　　　　　　　**A.C. Z1000 4.9**
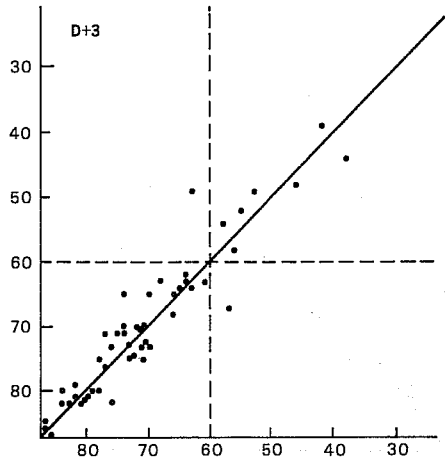
Fig.3　Same as Fig.2 showing AC(Z1000) for the long (1-3) and medium (4-9) wave components.

cases, evident on D+5 in particular, show remarkable improvements by T63. For the medium wave components, T63 is slightly better up to D+3 only.

It is interesting to isolate the best forecasts, since they are likely to be those for which the errors from sources other than technique (e.g. data, physics, ...) are smallest. In Fig.4, the distributions of the 17 best forecasts on D+3, D+4 and D+5 for Z1000 (see Fig.2) are shown and contrasted with those of the 18 or 19 worst ones, as the case may be, the selection criterion simply being a given value of AC on a given day (the best or worst scores on the various days do not necessarily belong to the same cases). The most notable feature of the distributions, for the best forecasts in particular, is their tendency to be spread out along an axis making an angle with the diagonal increasing from D+3 to D+5. The best forecasts appear thus to be mostly defined by the performance of T63: For the 17 cases on D+5, for example, T63 scores between 85 and 65% compared to 85 and 54% for N48. All the large improvements in the good cases on D+5 are in favour of T63. For the worst ones, although N48 is worst on average, the distributions are more spread out in all directions and the indication is less clear.

The mean (M) and standard deviation (S) of the complete distributions have been calculated from D+1 to D+7. Their values for each model are given in Table 2 and 3 for Z500 and Z1000 respectively.

A.C. Z1000 BEST CASES   **A.C. Z1000 WORST CASES**
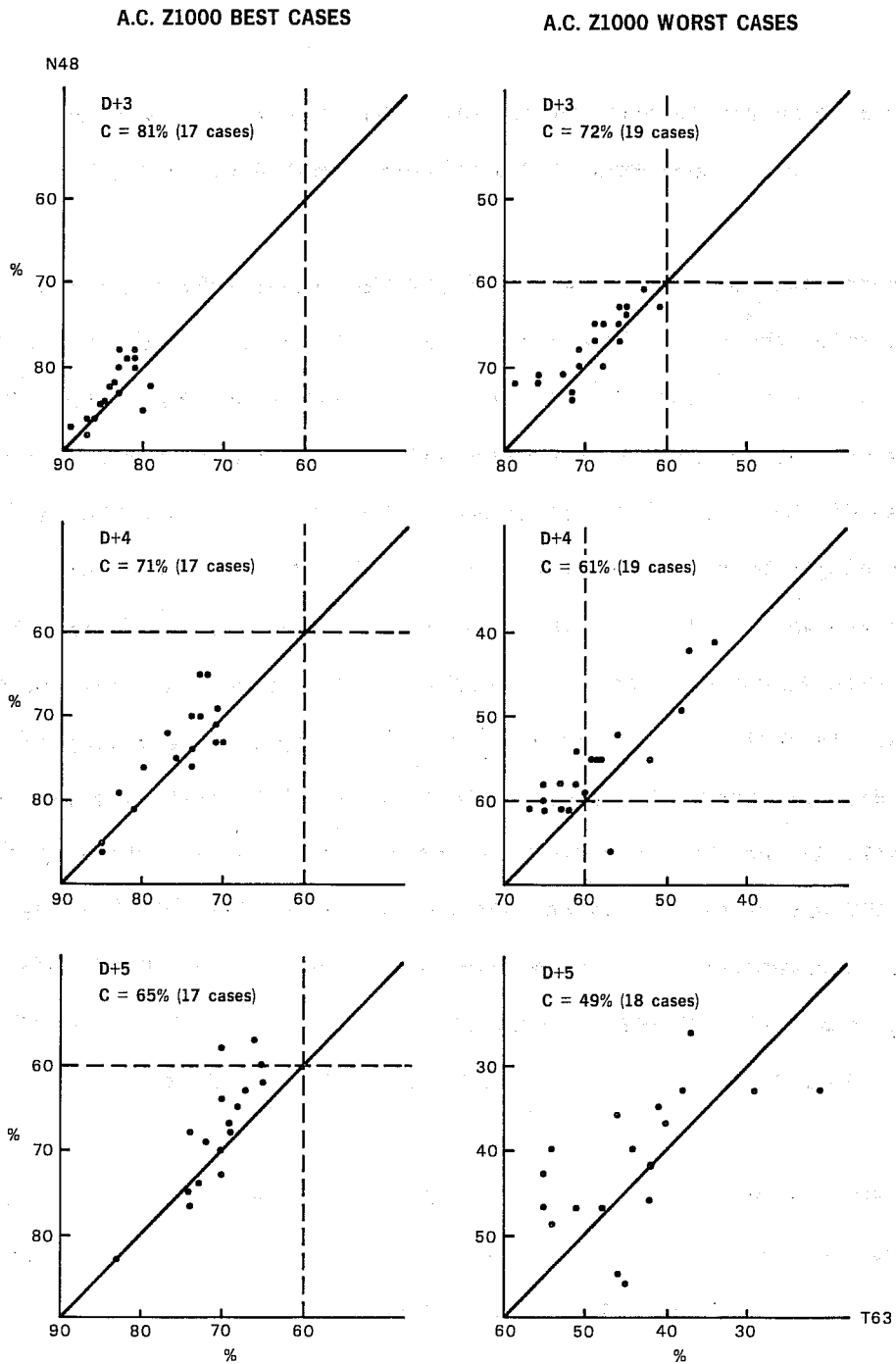
Fig.4 Same as Fig.2 showing AC(Z1000) for the best and worst forecasts. The best (worst) forecasts on a given day are determined by requiring that one at least of the models have AC $\geq$ c% ($\leq$c%), the value of c being displayed in each diagram.

Table 2. Mean (M) and standard deviation (S) of AC for Z500 (53 cases)

Z500	D+1	D+2	D+3	D+4	D+5	D+6	D+7
M(T63)	96.5	91.6	84.1	74.6	63.2	51.4	41.1
M(N48)	96.4	91.2	83.5	73.8	62.8	51.5	41.7
S(T63)	1.19	3.01	5.14	7.66	11.4	13.8	15.0
S(N48)	1.38	3.08	5.14	7.02	10.4	13.8	14.5

Table 3. Mean (M) and standard deviation (S) of AC for Z1000 (53 cases)

Z1000	D+1	D+2	D+3	D+4	D+5	D+6	D+7
M(T63)	93.1	85.4	76.5	66.6	57.1	47.3	37.3
M(N48)	93.1	84.5	75.2	64.8	55.0	45.5	37.5
S(T63)	2.22	4.81	6.80	8.70	12.4	13.3	14.1
S(N48)	2.51	5.40	6.97	9.10	12.6	13.7	14.6

As seen in the scatter diagrams T63 has better scores from D+1 up to D+5 (500 mb), D+6 (1000 mb). The mean AC differences (ΔM) are plotted in Fig.5a. They are small on average, but there are definite trends: the improvement by T63 increases up to D+4, D+5. Afterwards, the trend sharply reverses and already by D+6, D+7 it has vanished. The standard deviations well indicate the gradual dispersion of scores with forecast range. T63 and N48 are not substantially different in this respect. One should note in particular that beyond D+6, D+7, in association with the low M and large S, these curves are not stable. At 500 mb for example, the "better" behaviour of N48 was reversed by adding to our sample the two first 10 day comparative integrations done after this set of experiments. This result is an important one in terms of our second purpose since it strongly suggests that, in the study of FDs, we concentrate our efforts on D+3 to D+5 where differences are most significant and less affected by the other sources of errors, (e.g. data or physical package).
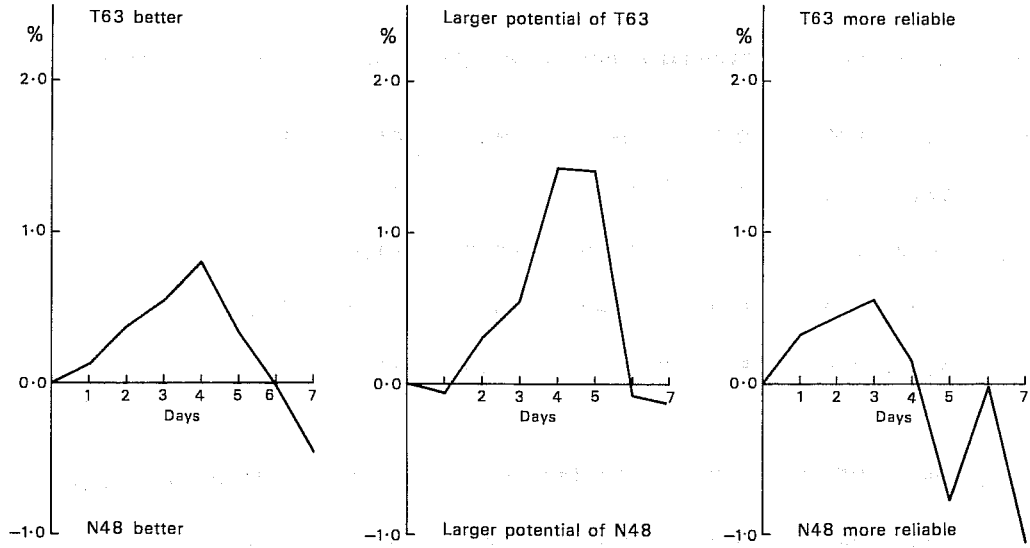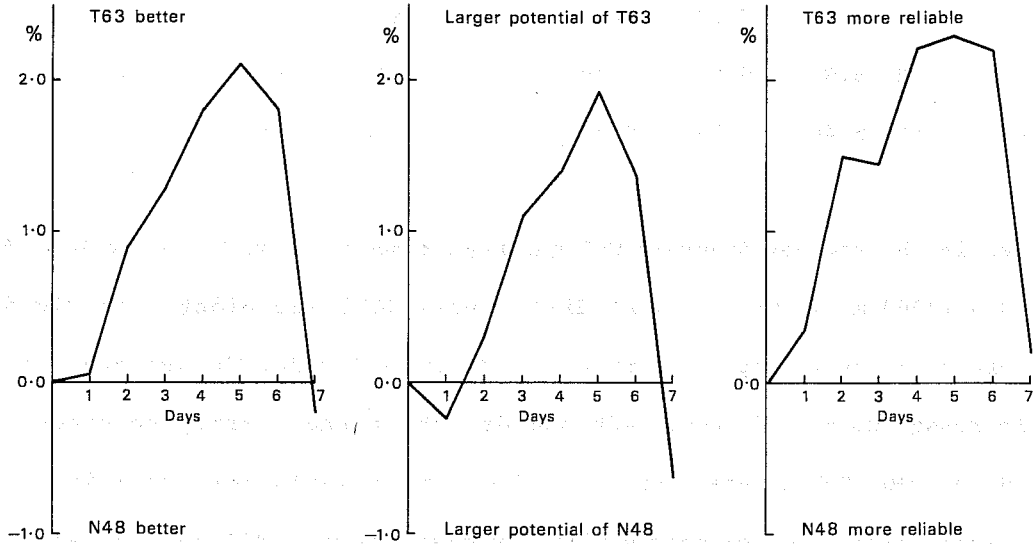
Fig.5 Mean differences in a) performance (M), b) capability (M+S) and
c) reliability (M-S) between T63 and N48. M and S are the mean and standard
deviation of the distribution (53 cases) of AC(%). Positive (negative)
differences correspond to higher values for T63(N48).

14

In Fig.5b and c we have plotted the functions Δ(M+S) and Δ(M-S). Insofar as M+S and M-S are representative statistical estimates (notwithstanding distribution skewness at short range in particular) of good and bad forecasts respectively, they may be taken as measures of model capability and reliability. In terms of these statistics, T63 improvements are systematic and relatively independent of more important causes affecting the forecast quality. T63 is not only better on average, but it is more capable of producing outstanding forecasts (Fig. 5b) and it is more reliable (Fig. 5c) since it produces fewer bad forecasts (except perhaps at 500 mb).

Finally we look at the results from the point of view of predictability (P): instead of asking what is the difference in scores at a given time range, we inquire how much further in time does one model provide the same forecast quality as the other model. We give an answer to this question in Fig.6 by comparing for a few quality levels (80, 70, 60 and 50%) the number of cases when P(T63) - P(N48) (and P(N48) - P(T63)) is higher than certain thresholds (3,6,9,...hours). Thus, for example, for a high quality level forecast of Z500 (AC $\geqslant$ 80%), P(T63) - P(N48) was more than 6 h in 25% of the cases (only 6% for the reverse) and it was more than 12 h in 9% of them (the reverse never occurred). At 1000 mb, the advantage of T63 is small for P(80%) but becomes larger for lower quality forecasts. It is in fact maximum for P(60%) and remains substantial for P(50%) whereas at 500 mb it has then essentially vanished. The mean P(60%) improvement by T63 for Z1000 is 6 h, but this mean value is not very instructive here. We must also consider the extremes. Then one notes that T63 increases P(60%) by more than 24 h 9% of the times (on average every 11th day in routine forecasting) while N48 rarely (2%) does it by more than 12 h.
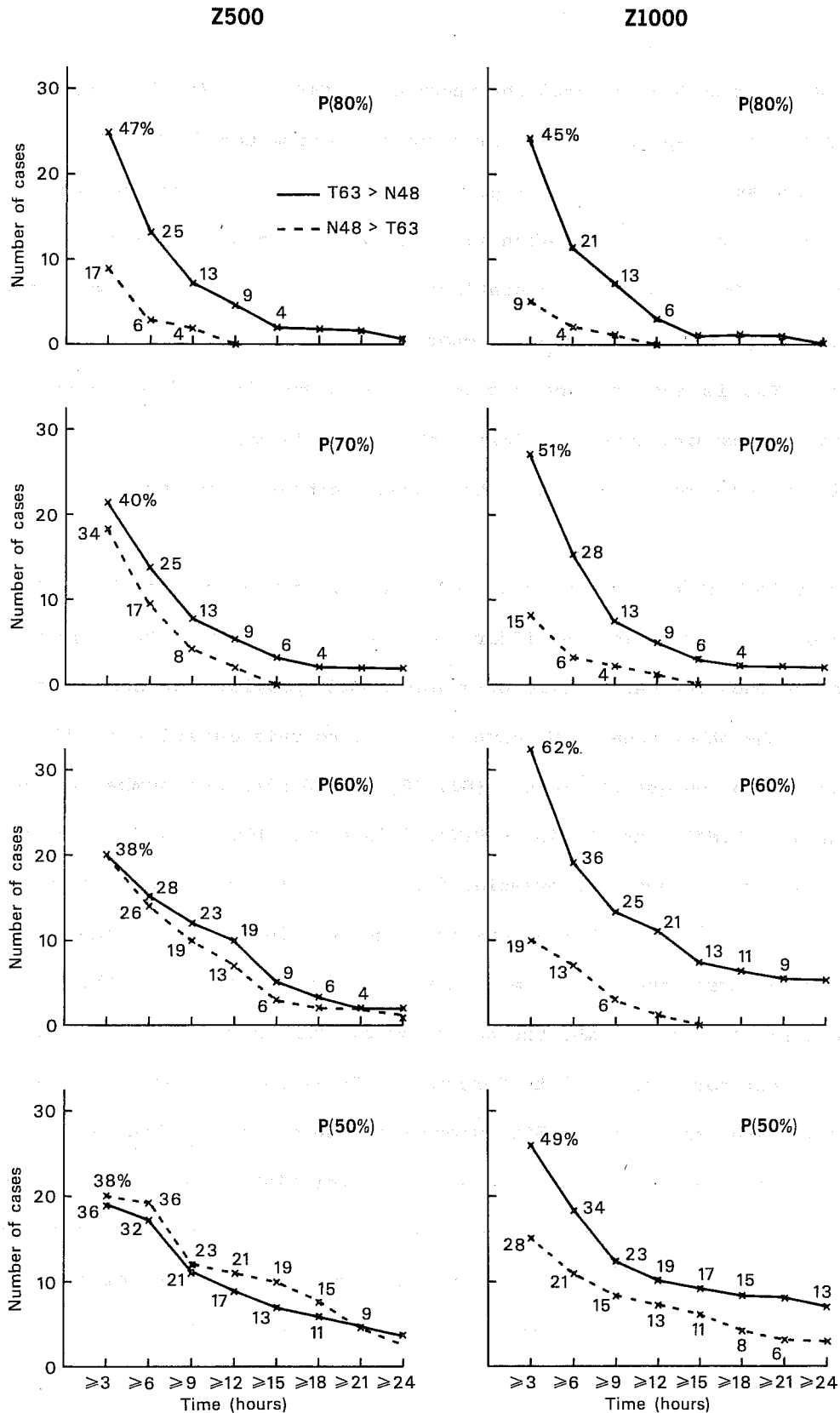
Fig.6 Predictability improvements (in terms of AC) of one model over the other
for Z500 and Z1000 at various fixed levels of quality: P(80%), P(70%)
P(60%), P(50%). Crosses on the full (dashed) lines correspond to the
number of forecasts where relative predictability improvements greater
than 3,6,9,....24 hours were observed in favour of T63(N48). Corresponding
percentages (out of the 53 cases) are also given.

2.3 Monthly and seasonally averaged scores.

In order to find out more about the structure of T63 improvements we present monthly-averaged time series of AC(Z) and SE(Z), and seasonal averages of vertical and horizontal cross-sections of AC(Z) and AC(T) scores. Monthly averages contain 4 or 5 cases, seasonal averages 13 or 14 cases. The 4 seasons will be referred to here and in subsequent sections as AUT (SEP, OCT and NOV cases), WIN (DEC, JAN, FEB), SPR (MAR, APR, MAY) and SUM (JUN, JUL, AUG).

The mean monthly AC for Z500 and Z1000 on D+1 to D+7 are displayed in Fig.7. Red (green) areas correspond to better forecasts by T63 (N48). At 500 mb, T63 scores better for nearly all months in the range D+1 to D+4. At 1000 mb, this range is extended to D+6. The differences appear on average somewhat smaller in the middle (WIN) than at both ends of the period. At short range (D+1 to D+3), its advantage is highest during MAR to AUG. These apparent seasonal variations may not be significant however, taking into account the results of TR23 showing large T63 improvements (all winter cases) and the fact that, for the first 4 months, scores were biased in favour of N48 due to differences in PP procedure (at 1000 mb, these sometimes accounted for nearly one point in AC as will be shown in Appendix A).

Fig. 8 displays SE scores in a similar way. However, as will be shown in Appendix A, the effect of the different PP procedures created a considerable bias for this score, especially at 1000 mb. Therefore, at this level the mean monthly SE are presented only from JAN onwards. On average for this score neither model appears clearly superior and contrary to AC the difference at 1000 mb are not more in favour of T63 than at 500 mb. However, one should note that unlike AC, SE are not normalised. Moreover, SE is more sensitive to the short scales contrary to AC which is more strongly affected by the large scales. In Section 3 and 4, we shall find that T63 usually keeps more variance than N48 in short scales which has an adverse effect on
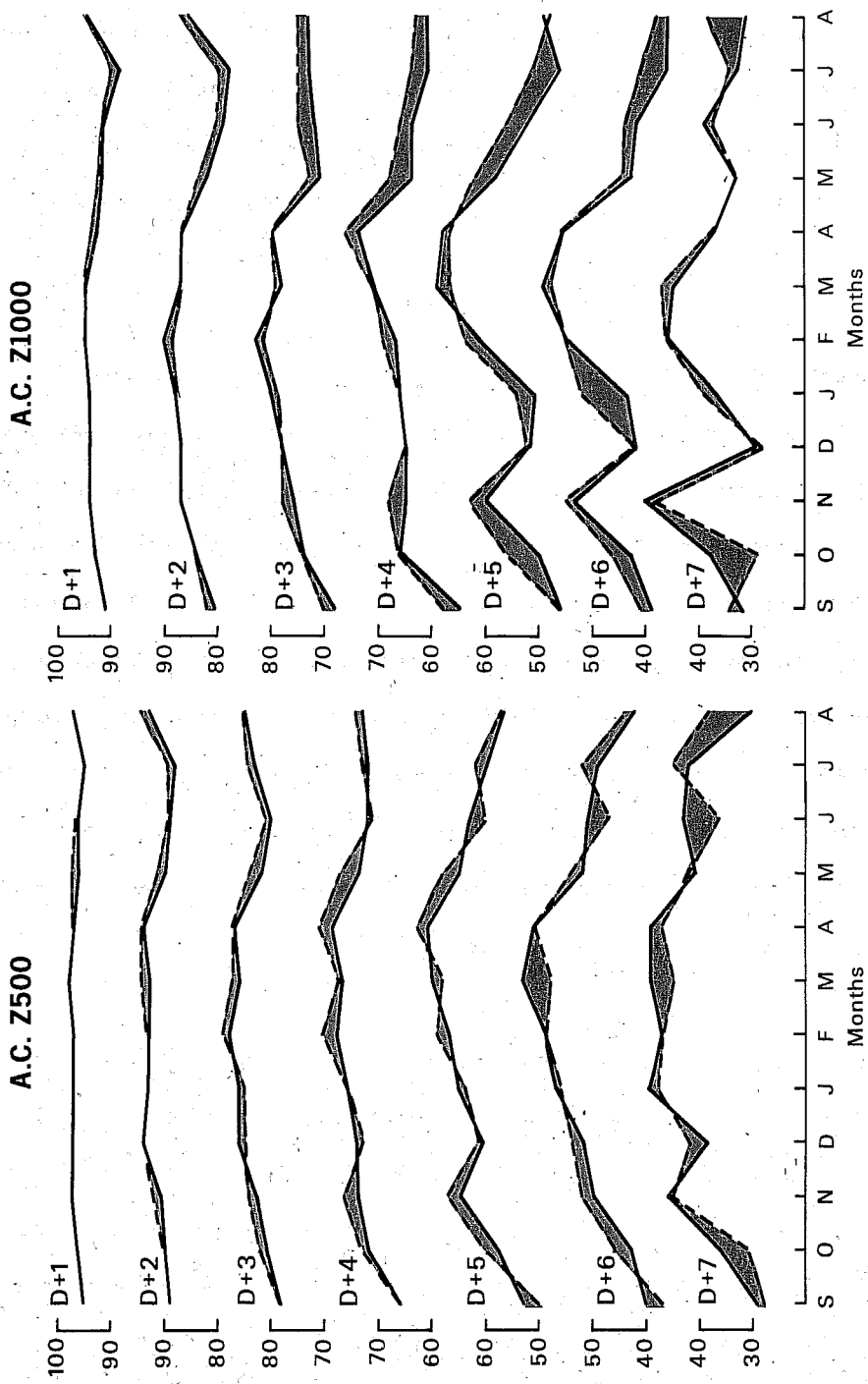
17

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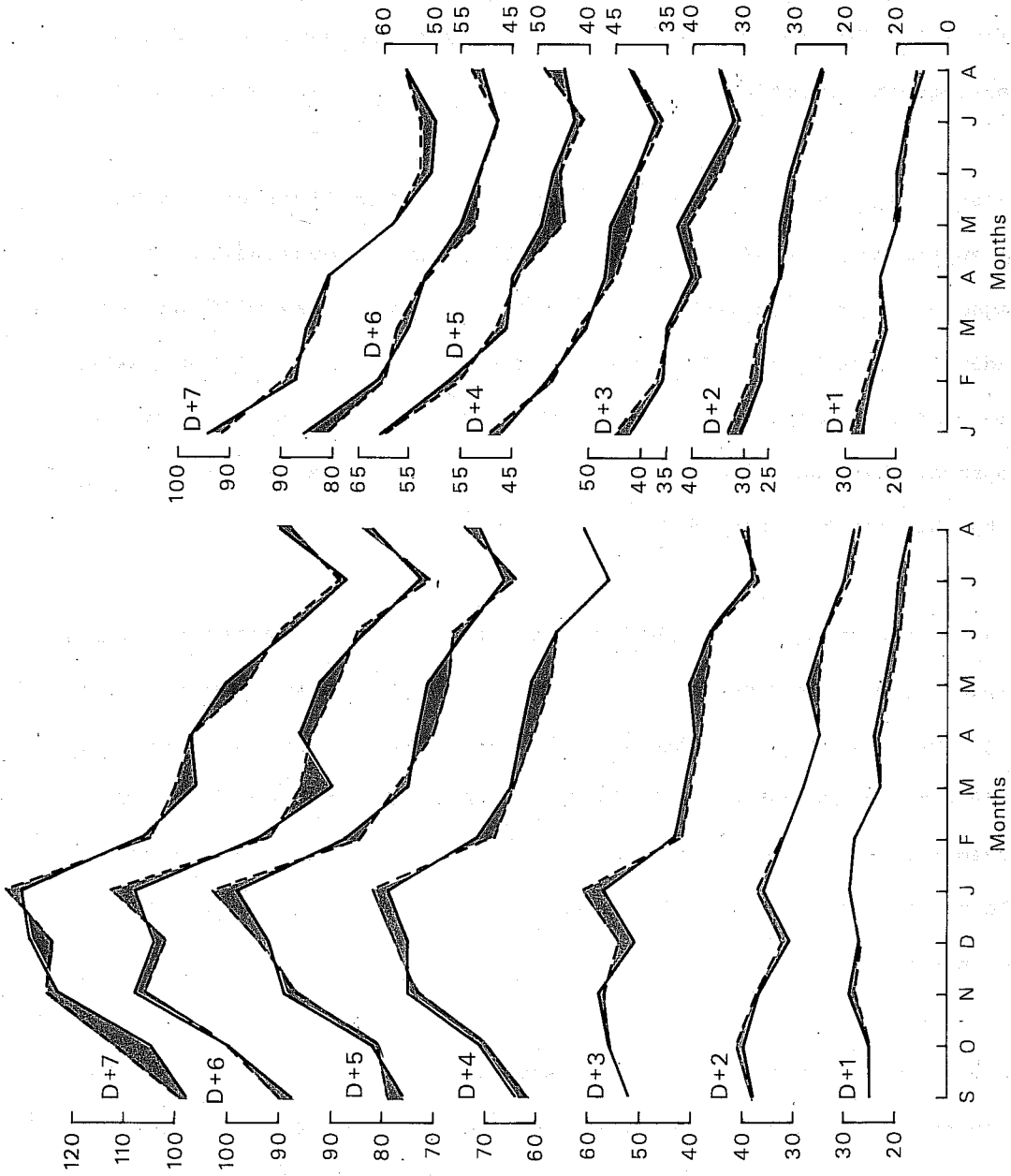82.5°N on D+1 to D+7.   (T63= ---;   N48= ———)

Fig. 8 Monthly mean values of SE(Z500) and SE(Z1000) averaged horizontally from
20°N to 82.5°N on D+1 to D+7. (T63= ----; N48= ———)

SE since in any event predictability in these scales is very small.

In Fig.9 we show pressure-time cross-sections of AC(Z) and AC(T) for the 4 above defined seasons. The structure of AC(Z) differences is as mentioned in Section 2.1. The improvement of T63 is on average maximum near 1000 mb, decreasing with altitude. It is such as to make the isolines more vertical.

There are two features of interest for AC(T). The first one is the very large superiority of N48 in Autumn in the lower troposphere, which is a consequence of the different PP used by the two models, as will be shown in Appendix A. (The Winter AC(T) were also spoilt, but to a lesser extent). The second is the improvement by T63 over N48 in the region of the tropopause, which was also noticed in most individual cases. We shall come back to this point in Sect.5.3.5.

In Fig.10 the results for the long and medium wave components of Z are presented. The improvement of T63 in the long waves is systematic, but small during WIN. As for Z total,the long-wave improvement is on average maximum near 1000 mb and decreases with altitude. In fact, total and long-wave AC(Z) improvements are strongly correlated. For medium waves, the differences have little vertical structure. T63 is slightly better only on average (it is worse in WIN).

Finally Fig.11 presents meridional cross-sections for all 4 seasons confirming the validity of the conclusion made in TR23 to the effect that T63 improvements are not particularly concentrated in any latitude bands.

Z (A.C.)   T (A.C.)

Autumn
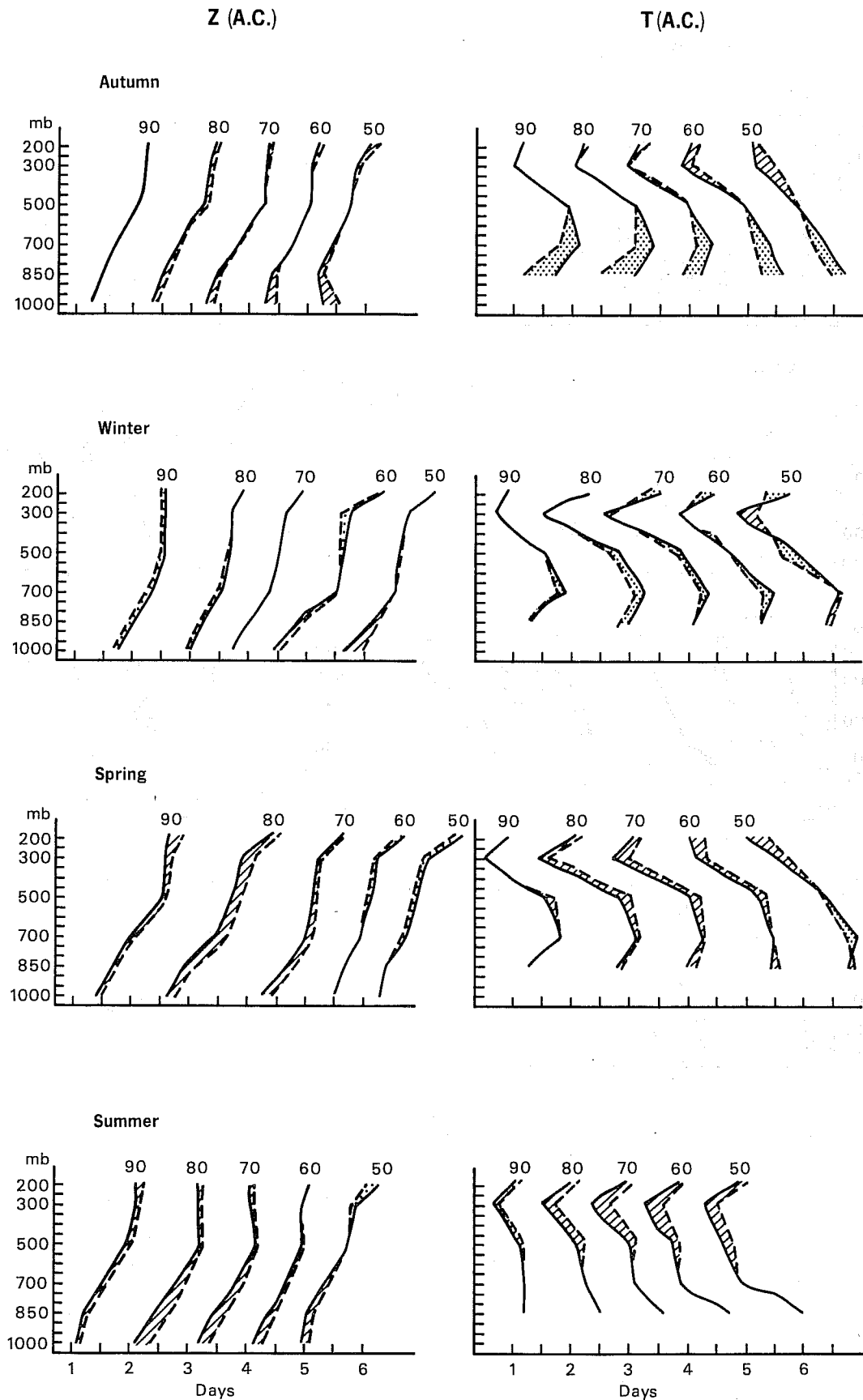
Fig.9 Time-pressure cross-sections of the 4 seasonal mean (AUT,WIN,SPR,SUM)
       values of AC(Z) and AC(T) averaged horizontally from 20°N to 82.5°N.
       (T63= ----; N48= ———). Hatched (dotted) areas corresponds to relative
       improvements by T63(N48).
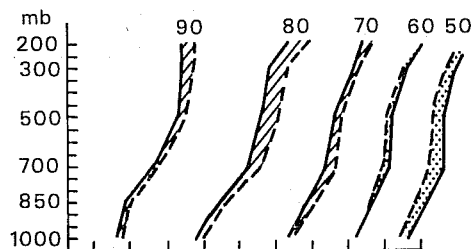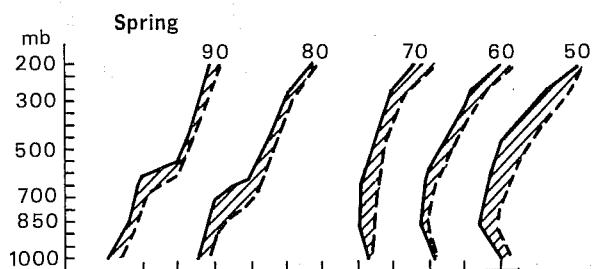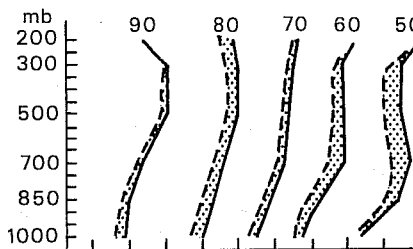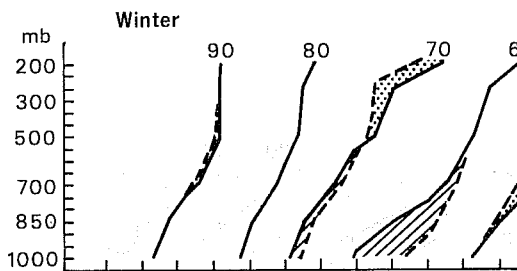
A.C. Z1000 1-3

A.C. Z1000 4-9
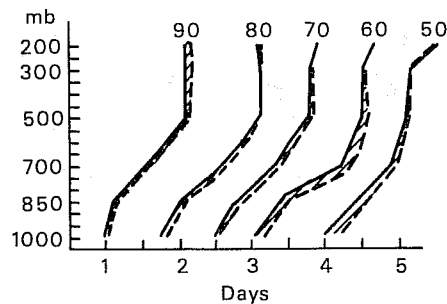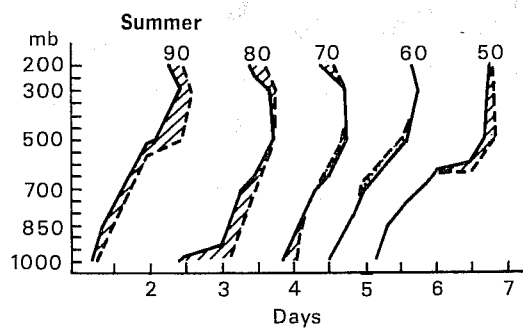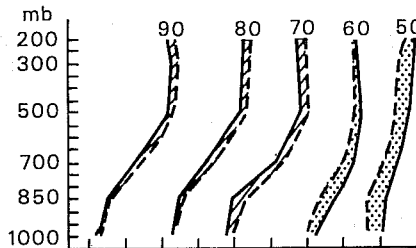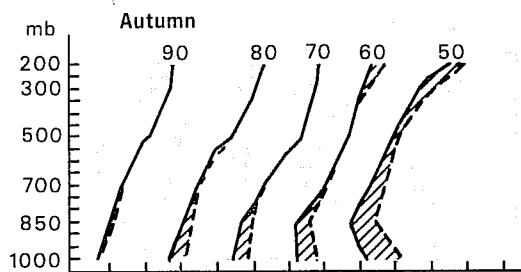
Fig.10 As in Fig.9: AC(Z) for the long (1-3) and medium (4-9) wave components.

A.C. Z1000-200
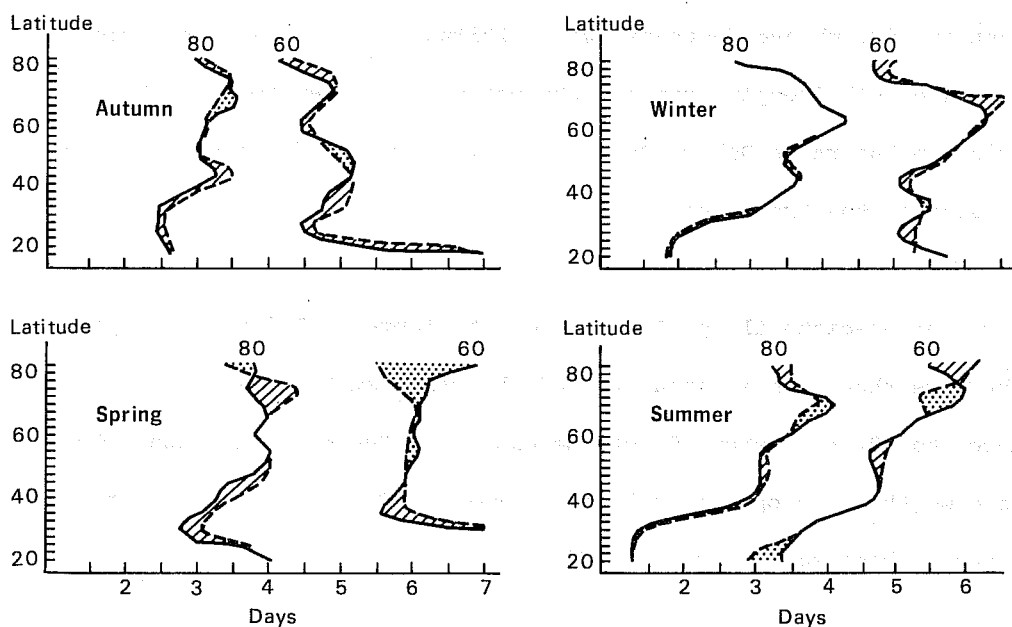
Fig.11 Time-latitude cross-sections of the 4 seasonal mean (AUT,WIN,SPR,SUM)
    values of AC(Z) averaged horizontally over longitudes and vertically from
    1000 to 200 mb (T63= ----;   N48= ——).

2.4 Summary of the objective evaluation

The main findings of TR23 on the mean performances of T63 and N48 are not substantially altered. On the contrary a strong statistical significance can be attached to them.

In terms of AC(Z), the superiority of T63 and the structure of the improvement are confirmed. Mainly concentrated in the large scales it is largest at 1000 mb and decreases with altitude. It grows with forecast range up to D+4 or D+5 dropping rapidly thereafter. Neither model is significantly superior in the range D+7 to D+10. T63 improvements for T were also noted in the region of the tropopause.

In terms of predictability, T63 is shown to increase P(60%) in comparison to N48 by more than 12 h in about 20% of the cases at both 500 and 1000 mb and by more than 24 h in about 9% of the cases at 1000 mb (which would be about 3 times a month in an operational environment). Such very large improvements were never observed in favour of N48.

Besides this higher potential to produce forecasts of outstanding quality, T63 is also more reliable (mainly for Z1000).

Finally T63 improvements are not noticeably dependent on the initial conditions and the time of the year. Their size however may depend on the type of meteorological situations. In the next section a careful examination of FDs will be made for several individual cases and their impact on the performance of the models will be closely monitored. Thus the meteorological significance of T63 improvements will become clear.

3.SYNOPTIC EVALUATION.   INDIVIDUAL FORECAST DIFFERENCES

3.1 Introduction

In this section we present and discuss many of the forecasts made by T63 and
N48, with emphasis on their differences. We concentrate on the large scale
flow, showing maps of Z500 and Z1000 for the Northern Hemisphere north of
20°N. The models, as stated in the introduction are very similar. FDs are
thus expected to be small, indeed much smaller (at least up to D+5, D+6) than
forecast errors themselves. For these reasons, maps of the height FDs will
be shown to complement the maps of the observed and forecast height field.
They will allow the reader to quickly identify the regions of interest. In
fact, they will prove invaluable in evaluating the FD amplitude, categorizing
their pattern, displaying their geographical distribution and evaluating
their meteorological significance.

The sampling size of the comparison (53 cases, 106 10-day integrations)
prevents any attempt at making a complete presentation. On the other hand,
it was felt that the traditional approach which consists in selecting a few
cases for detailed synoptic evaluation was not adequate in the present
situation. After looking at all 10-day forecasts, at 24-h intervals,
sufficient case to case regularity was found, depending on the season and/or
the meteorological situation to justify a preliminary classification: we gave
preference to cases and times where large local FDs developed, cases and
times where well organised difference patterns were apparent, cases showing
large differences in scores and cases with higher predictability (smaller
errors).

Root-mean-square differences grow fairly smoothly as a function of the range
of the forecasts (TR23). On D+1, differences are very small, even locally.
They do not appear well organised and are thus hard to evaluate. From D+2 to
D+5, in most cases, it is possible to observe the growth and displacement of
well defined, well organised difference patterns and to relate them to given

meteorological features, such as lows and highs, troughs and ridges. When

the forecasts do not deviate too quickly from reality it is also possible to

give a value judgement on the relative performance of the models on specific

features. By D+5, however, relatively large scale differences have developed

covering most of the surface of the maps. It becomes usually difficult to

discuss individual features in isolation, be they of very large scale.

Furthermore, beyond D+5, most forecasts have lost much of their predictive

value. It is then but rarely possible to give a meaningful interpretation of

FDs. The objective evaluation of the forecasts (Section 2) shows that

neither model is significantly more skillful beyond D+6 and D+7. This

suggests that we concentrate our attention on the forecast period ranging

from D+3 to D+5.

A list of cases is given in Table 4. The dates correspond to the starting

points of the forecasts, their initial conditions being operational ECMWF 12Z

global grid-point initialized analyses.

To give an idea of the order of magnitude of the differences, their growth

rates and the sensitivity of the models to given conditions, we also show, in

Table 4, the absolute values of the maximum height (dam) FDs at both 500 and

1000 mb on D+1 to D+5. Values greater than a given threshold (respectively

5,10,15,20,25 dam) are underlined. Note that these numbers represent local

extremes only and as such do not take into account the presence of other

extremes of similar magnitude. Numbers printed in bold characters in Table 4

correspond to cases and times for which some maps will be shown. Thus the

table also serves as a table of content for the figures presented in this

section.

In the presentation, we shall mainly follow the chronological order and use a

separate sub-section for each of the 4 seasons.

Table 4. Case enumeration giving the maximum height differences (in dam) between forecasts at 500 and 1000 mb on D+1 to D+5 (values greater than certain thresholds are underlined). Cases and times where figures are presented in Section 3 are printed in bold characters (the corresponding figure numbers are given in the right columns).

CASES	Z500 D+1	D+2	D+3	D+4	D+5	Z1000 D+1	D+2	D+3	D+4	D+5	Figures Z500	Z10000
SEP 3	4	7	10	18	16	5	6	10	15	12		
10	4	**7**	**10**	**18**	18	7	10	11	19	18	17	18
17	3	12	16	18	21	5	9	15	22	26		
23	4	9	14	15	26	6	10	11	18	25	15	16
30	3	10	14	15	14	10	12	14	15	20	19	20
OCT 4	5	9	16	15	20	9	11	16	19	25		
11	4	9	9	15	21	4	12	13	19	20		
18	4	9	14	17	21	4	11	10	15	18	21	
25	5	7	16	25	33	4	8	11	13	21	22	
NOV 1	4	9	15	22	30	5	14	19	32	47	23a	23b
8	12	11	16	17	29	6	8	10	15	22		
15	6	12	18	14	16	7	11	28	25	26		24
22	6	9	9	14	18	4	9	11	14	19		
29	6	10	23	21	21	5	10	25	15	15		25
DEC 6	4	6	7	13	19	6	8	13	15	20	26	
13	5	11	13	20	21	4	10	18	15	15		
20	7	11	15	17	17	5	9	17	21	22		
27	4	6	9	17	18	5	7	10	11	17		
JAN 3	3	6	17	22	22	3	5	11	15	20		
10	4	9	17	28	31	4	7	12	17	25	27,28	29,30,62
16	3	7	10	17	19	3	8	14	14	19		
24	4	8	15	18	26	6	7	8	17	21		
31	4	10	10	13	21	5	10	11	14	20		31
FEB 7	3	6	9	14	21	3	7	10	19	16		32
14	3	6	7	13	13	5	4	10	11	12		33
21	3	7	18	17	18	4	8	14	12	25		34
27	3	8	13	11	15	6	6	15	11	17		35
MAR 6	2	5	8	13	19	4	4	9	19	19	36,58	37,59
13	3	6	9	13	14	3	9	6	9	10		
20	3	5	10	15	14	3	5	17	12	12		
27	3	6	8	14	18	3	7	8	13	23		
APR 3	2	7	12	15	22	3	8	9	13	12	38	
10	3	8	12	18	19	3	6	9	14	17		
17	4	7	8	11	15	4	8	7	14	17		
24	3	6	8	13	17	4	6	7	22	26	39	
MAY 1	4	7	12	15	18	3	7	10	13	15	40	
8	4	6	9	11	15	5	7	9	9	10		
15	6	7	10	11	16	4	6	15	20	17		
22	3	6	14	13	14	4	6	7	10	13		
29	2	5	10	10	10	3	6	6	10	11		
JUN 5	2	6	10	14	25	3	5	10	13	15	41	42,60
12	2	5	11	15	22	2	7	11	12	15		
19	2	4	8	8	11	4	4	7	9	12		
26	2	5	7	6	11	3	6	8	11	7		
JUL 3	2	4	6	9	11	2	4	5	9	10		
10	2	4	7	8	9	2	4	7	8	12		
17	3	5	8	13	13	3	4	9	9	12	43a	43b
24	2	8	9	10	13	2	8	7	9	11	44a	44b
31	2	3	10	15	20	2	5	10	11	12	45	
AUG 7	3	6	10	13	15	3	6	7	8	11		
14	2	4	9	16	26	2	5	12	12	15		
21	3	9	12	16	24	4	8	12	8	17	46	47
28	3	5	11	17	19	3	8	12	12	13		

Mainly to serve as reference, Fig. 12 and 13 compare AC(Z500) and AC(Z1000) on D+3, D+4, D+5 of all individual forecasts by T63 and N48. Red (green) areas correspond to better AC for T63 (N48). As discussed in Section 2, the superiority of T63 is small and true only on average at 500 mb while it is much larger and nearly systematic at 1000 mb.

In the following we shall have to refer to various regions of the Northern Hemisphere. For convenience, we have divided the polar stereographic map projection used to display the fields into 7 regions: the polar region (ARC) and 6 sectors (tilted) of equal area defined and labelled as shown in Fig.14 (ATL,EUR,WAS,EAS,PAC,WAM).

3.2 Autumn cases

Period from SEP 3 to SEP 29

At the beginning of SEP 1979, the large scale flow is weak but with strong meridional components and numerous short waves travelling slowly eastwards. During the month, the circulation gradually speeds up and the scale of travelling waves increases (The observed (OBS) Z500 and Z1000 on SEP 12 to 14 can be seen in Fig.17 and 18). In the second half of the month, we note the formation of a planetary scale wave number 2 pattern with the main troughs located over the oceanic (ATL and PAC) regions. The short waves often become unstable and amplify (e.g. Fig.15 and 16, OBS).

FDs (as seen from Table 4) are fairly large in general during this period at both 500 and 1000 mb on D+1 and D+2. The average performance of the models (as given by Fig.12 and 13) is similar for the first 4 cases with 2 cases in favour of T63 and 2 against it.

Largest differences are often found over ATL which is also where the 500 mb jet (loosely defined as the region of tight north south height gradient) is generally strongest. A typical example of the type of FDs arising over ATL

29
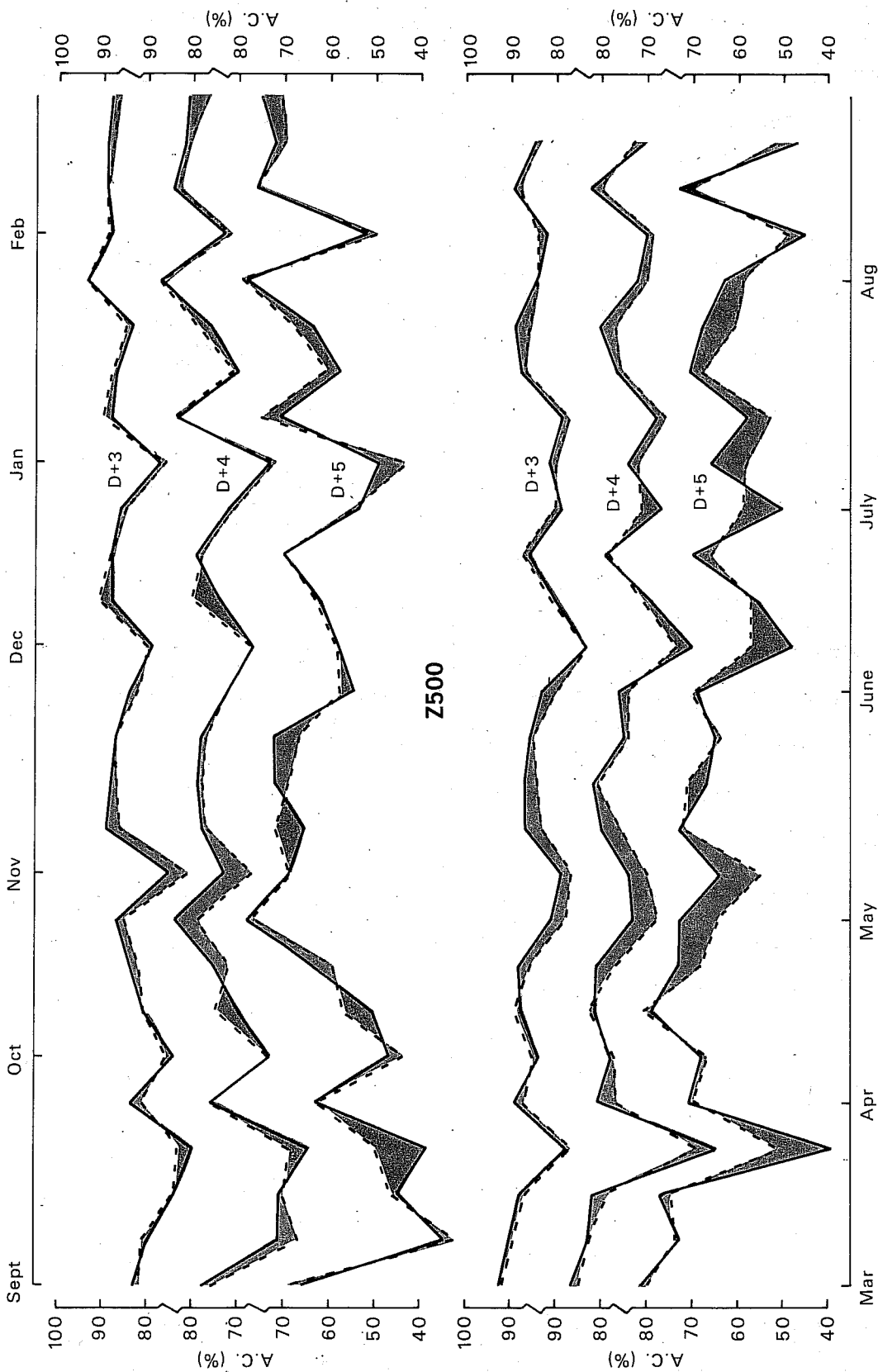
Fig.12 Individual forecast values of AC(Z500) on D+3, D+4 and D+5 (N48= ----; T63 = ——). Red (green) areas correspond to relative improvements by T63(N48).
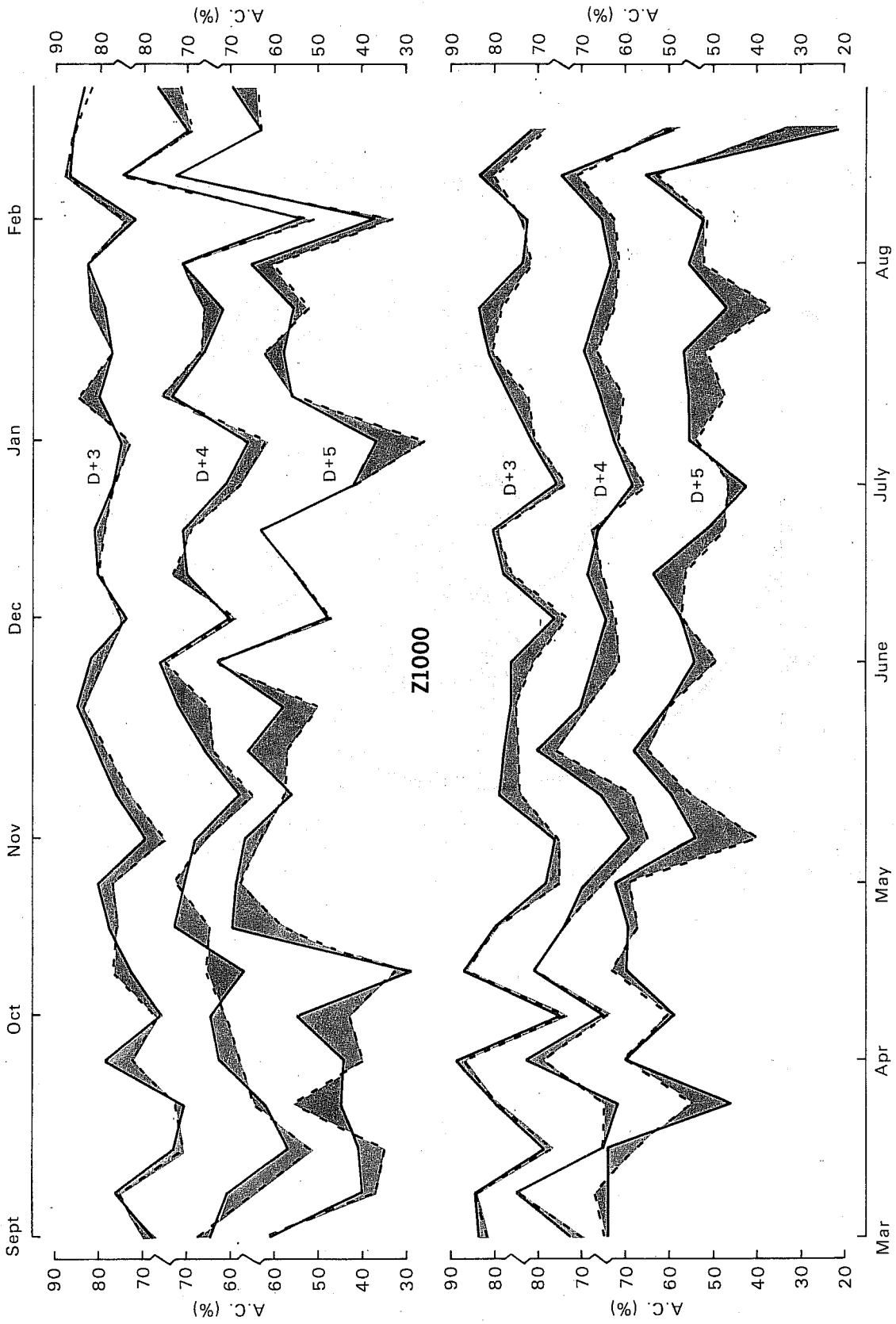
Fig.13 Same as Fig.12 for AC(Z1000).
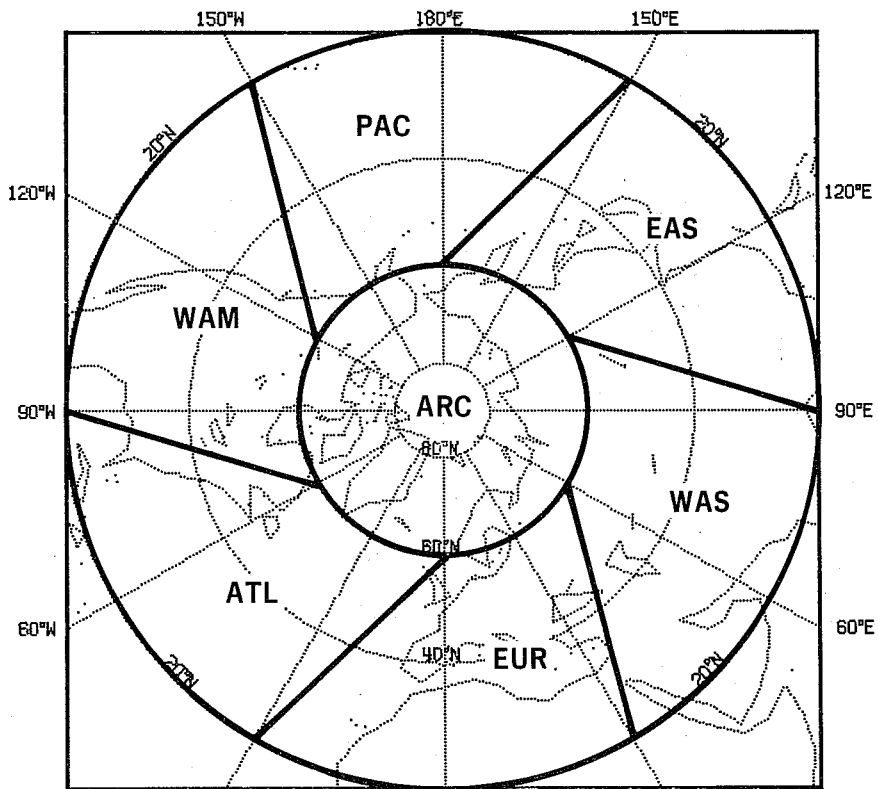
Fig.14 Polar stereographic map of Northern Hemisphere north of 20^ON divided into
7 labelled regions. The labels are often used in the text to locate
approximately forecast features of interest.

during this period will be shown for the SEP 10 case. On SEP 23, T63 performed substantially worse than N48. Nevertheless, it appeared to us to illustrate the growth and the nature of some of the most common FDs between the models (phase differences in particular) and thus worth a detailed discussion. Breaking the chronological order, it will be presented first in order also for the reader to get more easily accustomed to our procedure.

Case of SEP 23

Fig. 15: Z500, D+1, D+2, D+3, D+4, D+5, D+6

Fig. 16: Z1000, D+1, D+2, D+3, D+4, D+5, D+6

In the above figures, as in all subsequent ones in this section, the following panel arrangement has been used. Forecast time increases from top to bottom. From left to right are shown the ECMWF analyses (OBS), T63, N48 forecasts and forecast differences (N48-T63). All Z values are given in dam. Contour intervals are at every 8 dam at 500 mb and every 4 dam at 1000 mb and for FDs which are positive when T63 is deeper than N48 and vice versa.

We would like first to compare the forecasts of T63 and N48 to OBS, then make a detailed discussion of the difference maps .

At 500 mb (Fig. 15), both models tend to exaggerate the meridional component of the flow: the troughs over EAS and PAC and the ridge over WAM are already too pronounced on D+3. The trough over east ATL is also too intense but its SW-NE tilt is noticeably different: T63 forecasts it about right while N48 moves the trough too slowly in its northern part (as readily seen from the N48-T63 maps. On D+4, the models have failed to deepen the trough over west ATL, developing instead the one upstream over WAM. By D+5, the forecast 500 mb pattern do not display anymore the predominant wavenumber 2 pattern as seen in OBS. On both D+5 and D+6, the large trough-ridge-trough systems developed by the forecasts across WAM are in complete disagreement with the
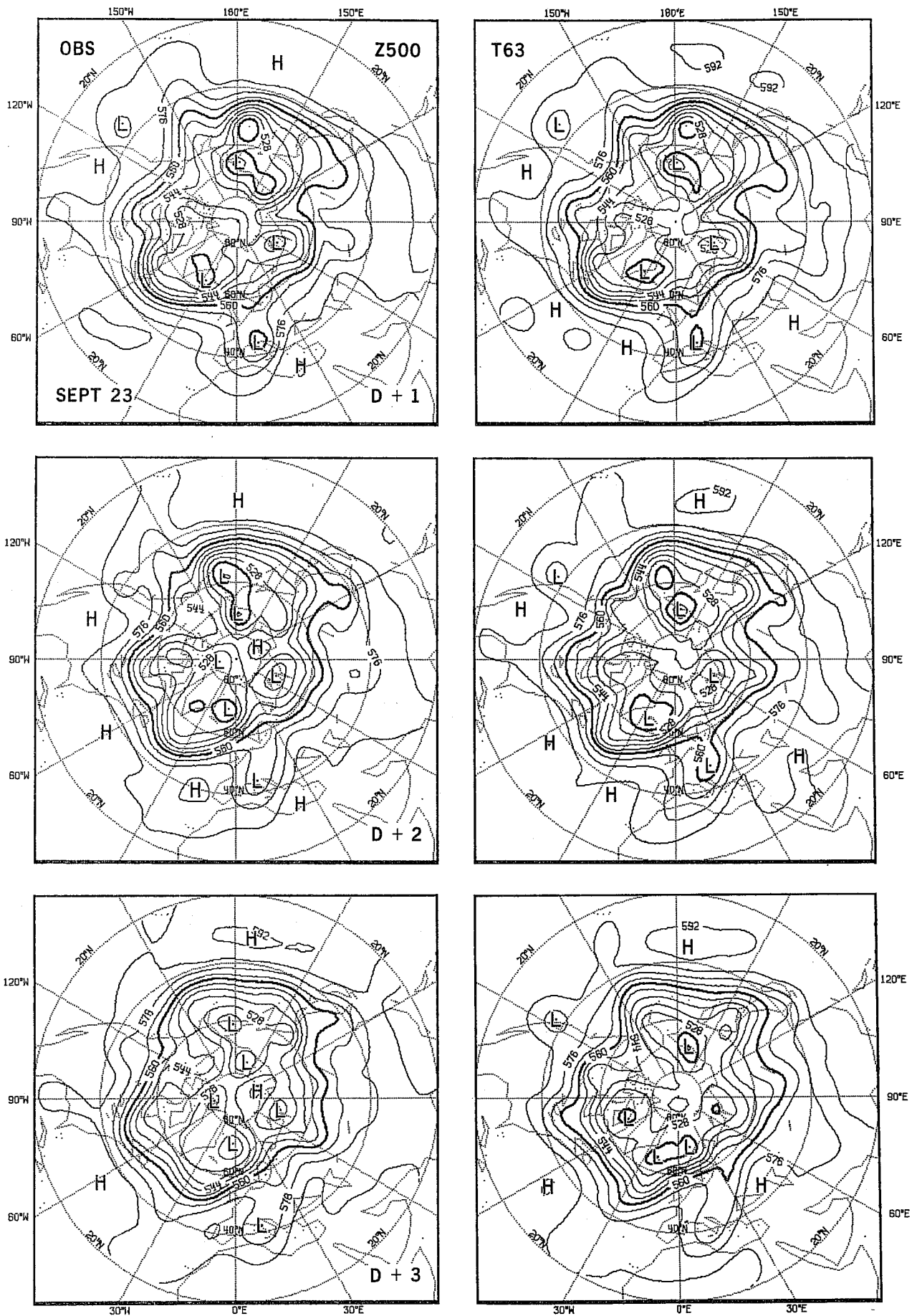
Fig.15a SEP. 23, Z500: D+1, D+2, D+3
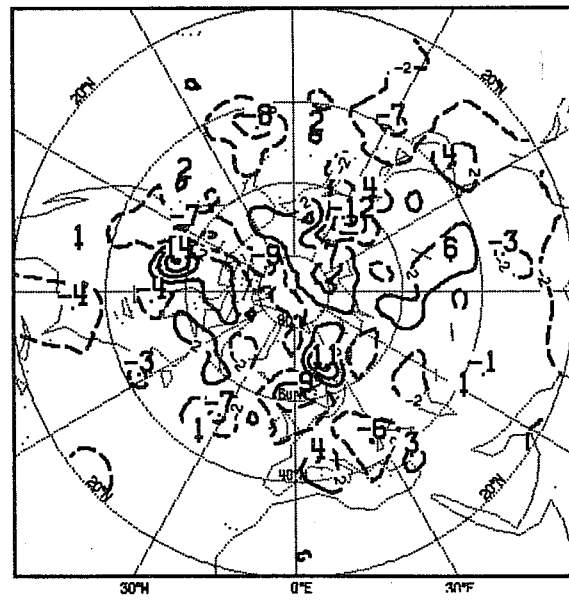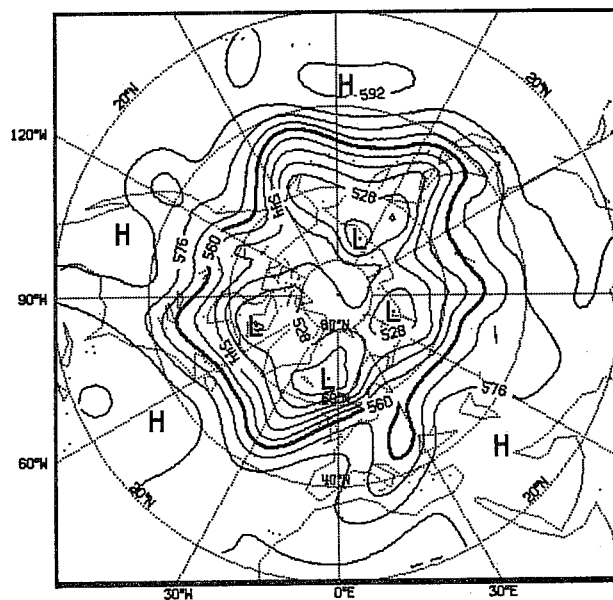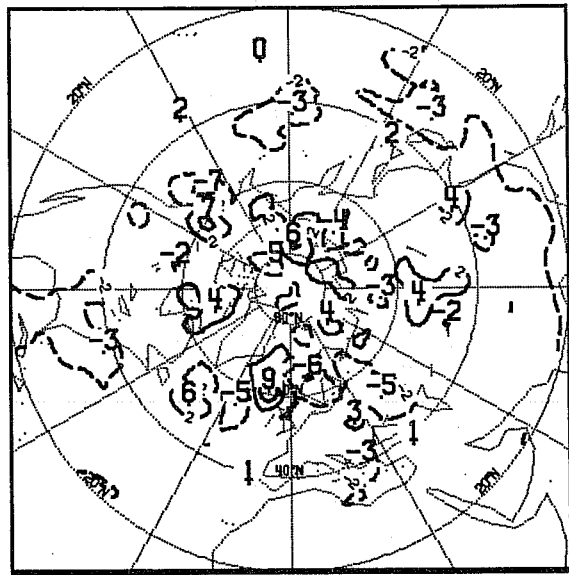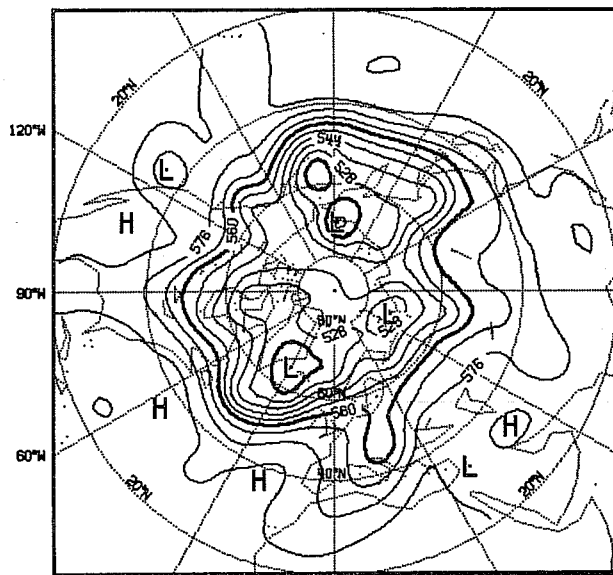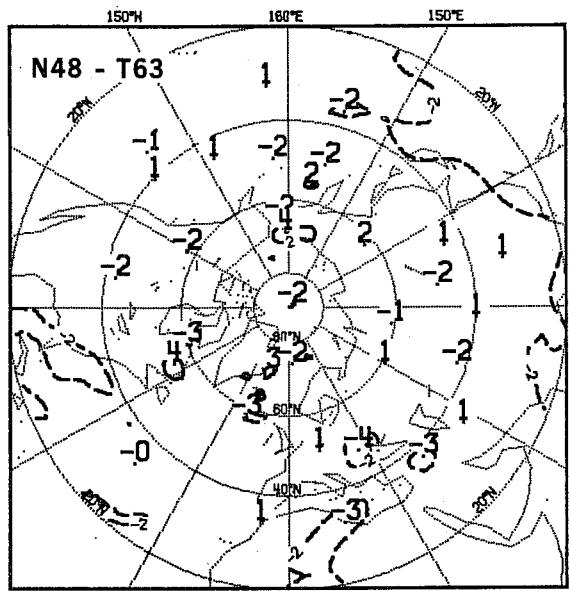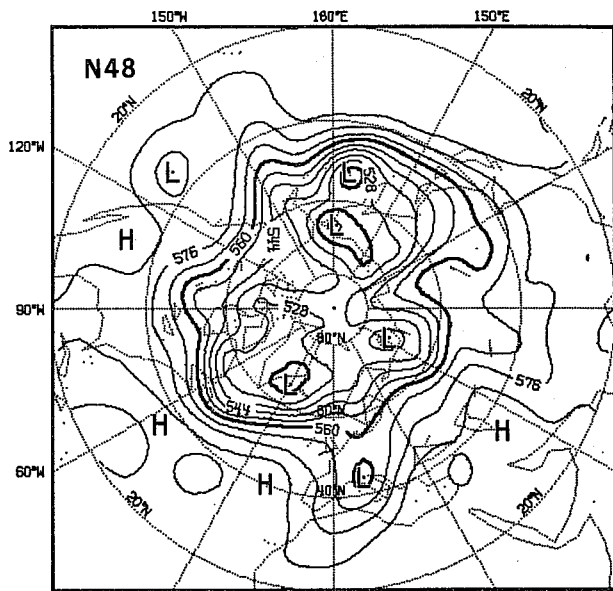
Fig.15a (Cont.)
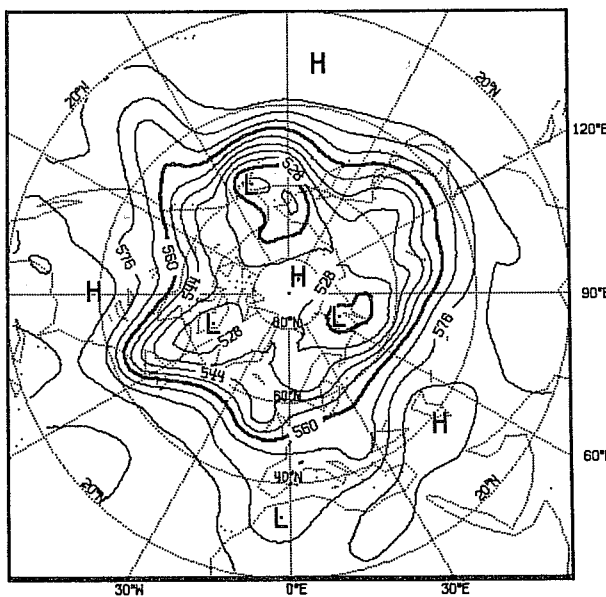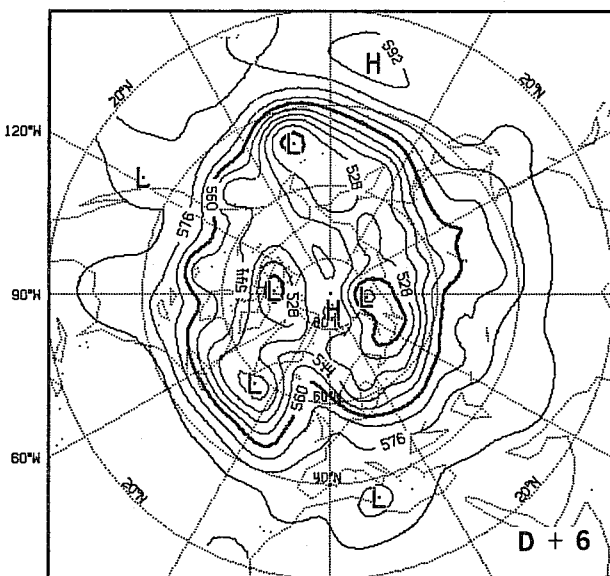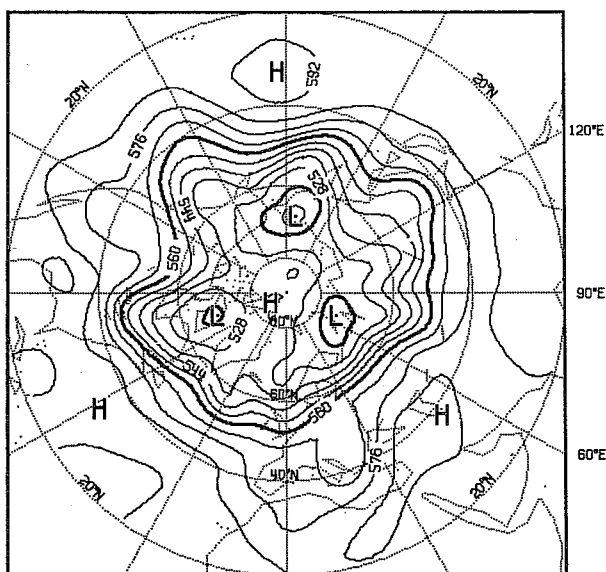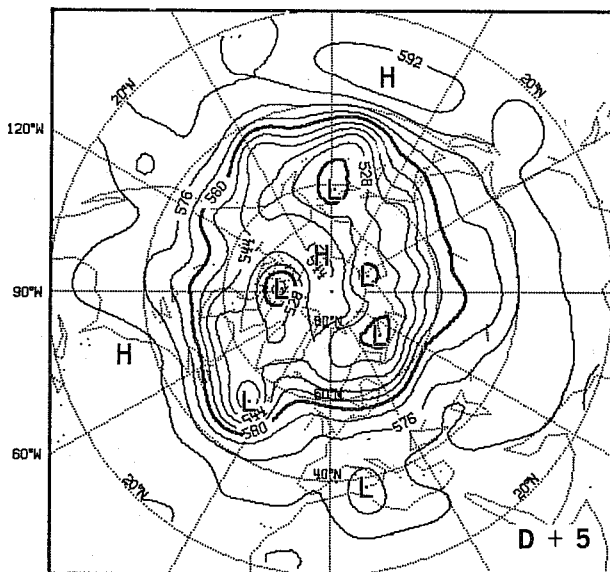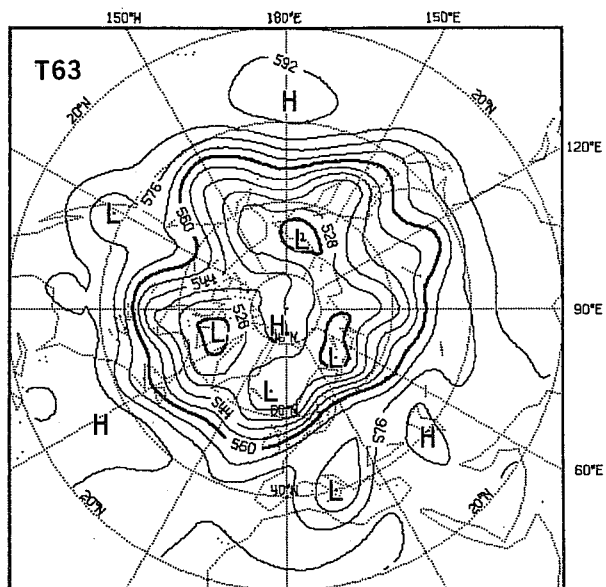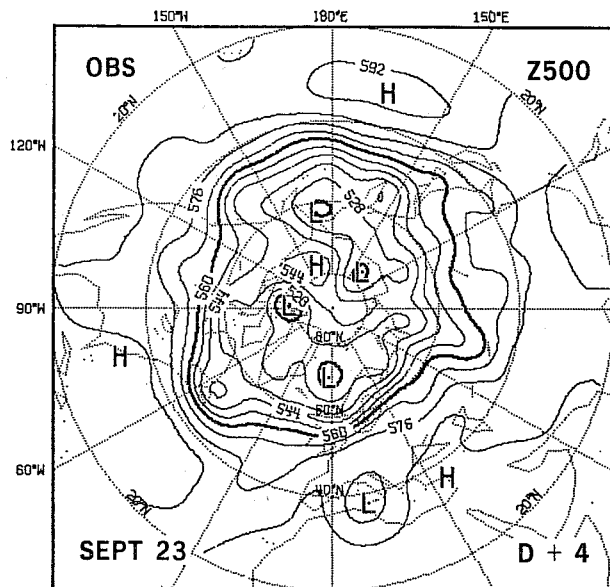
Fig.15b SEP. 23, Z500: D+4, D+5, D+6
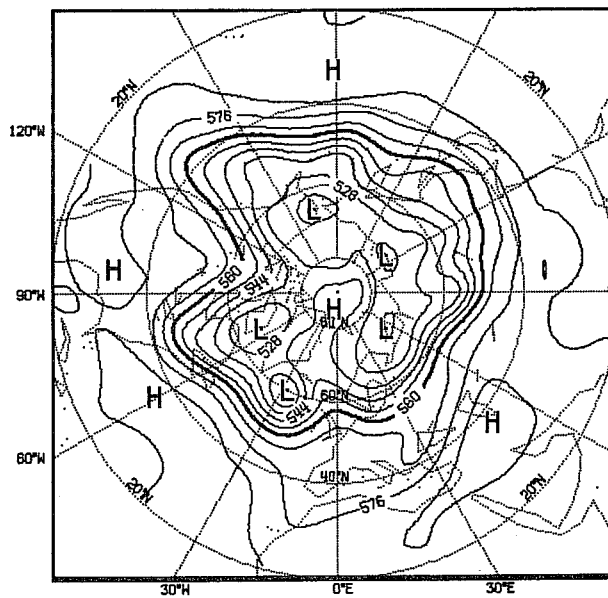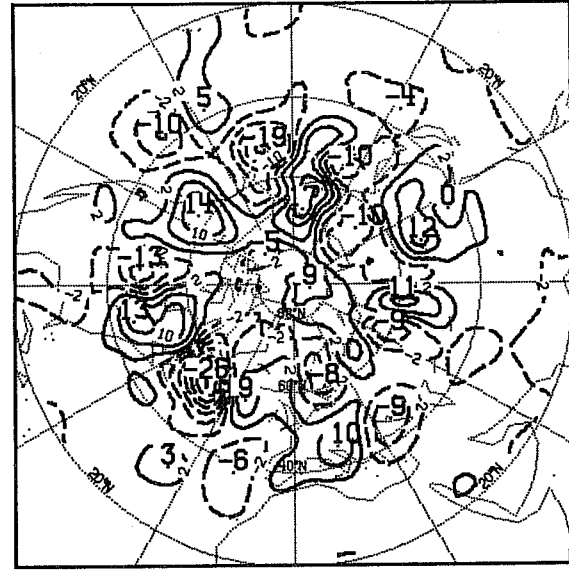
Fig.15b (Cont.)
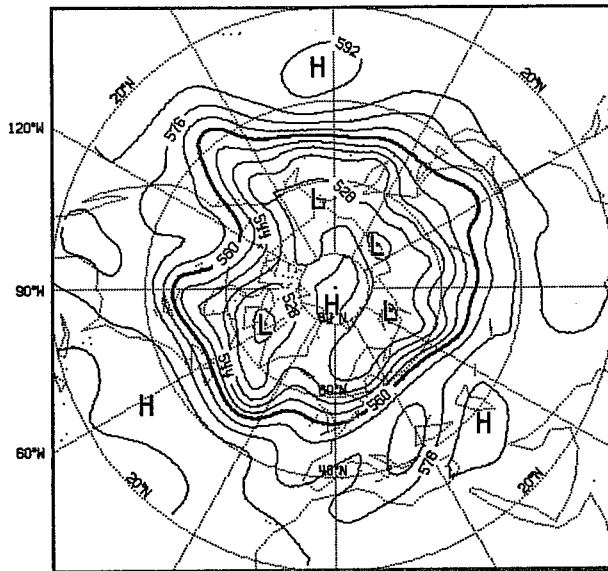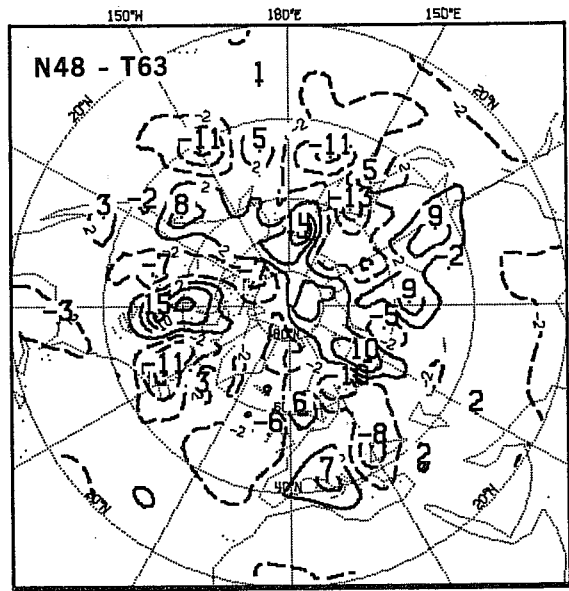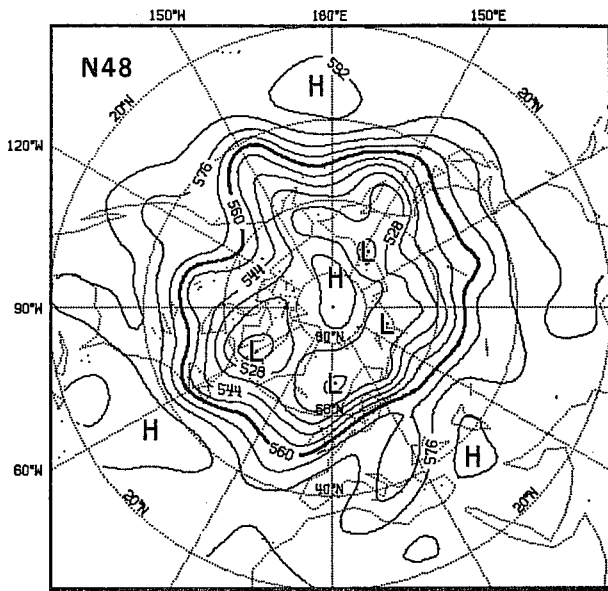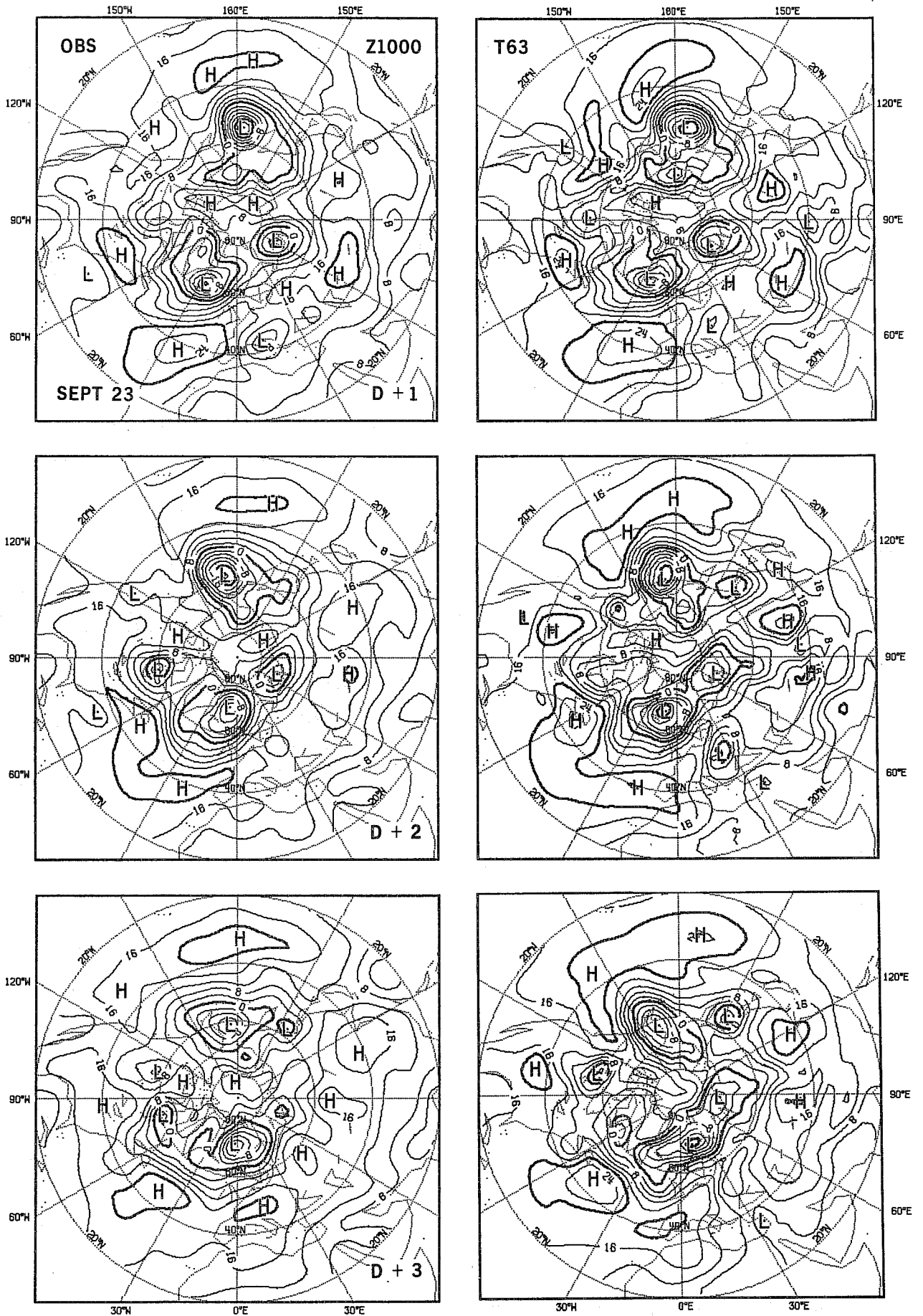
Fig.16a   SEP. 23, Z1000: D+1, D+2, D+3

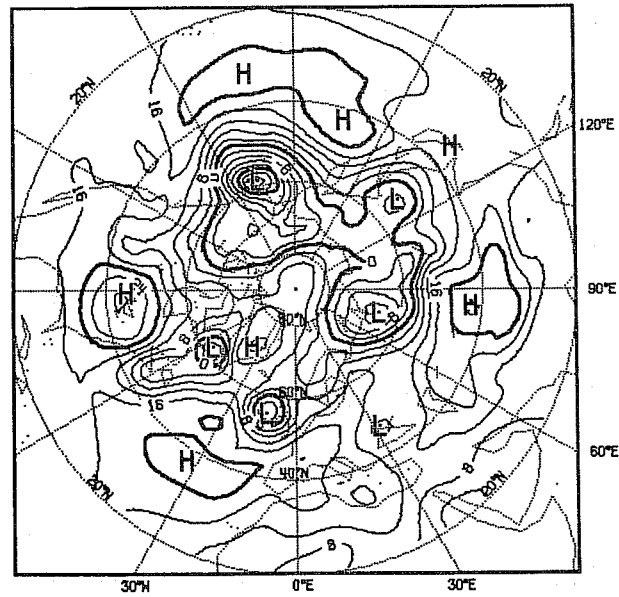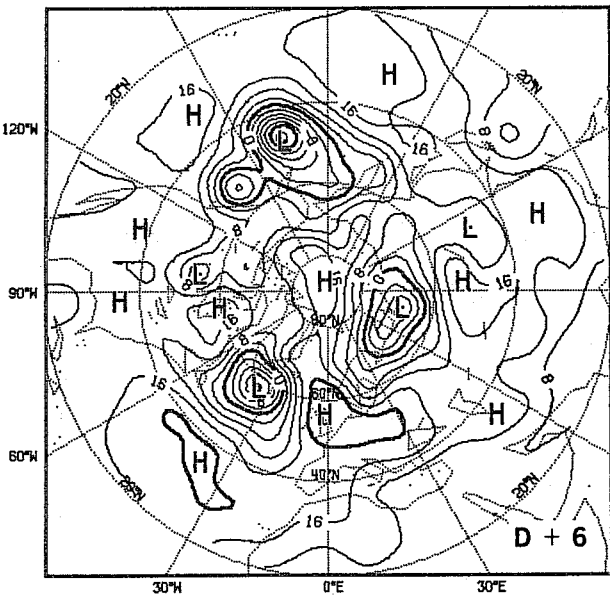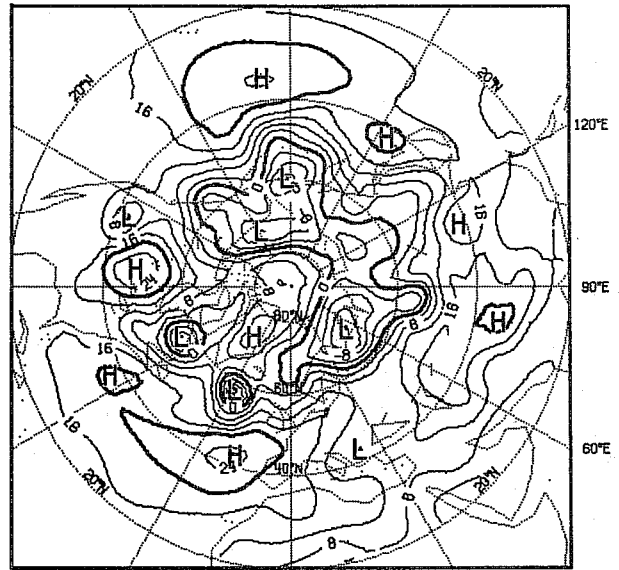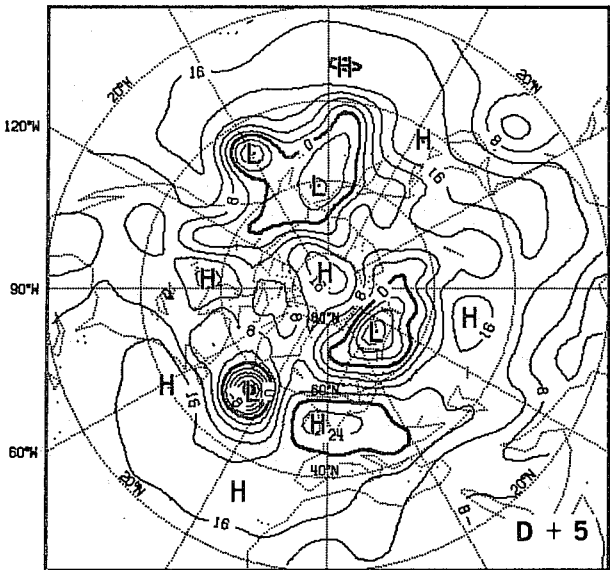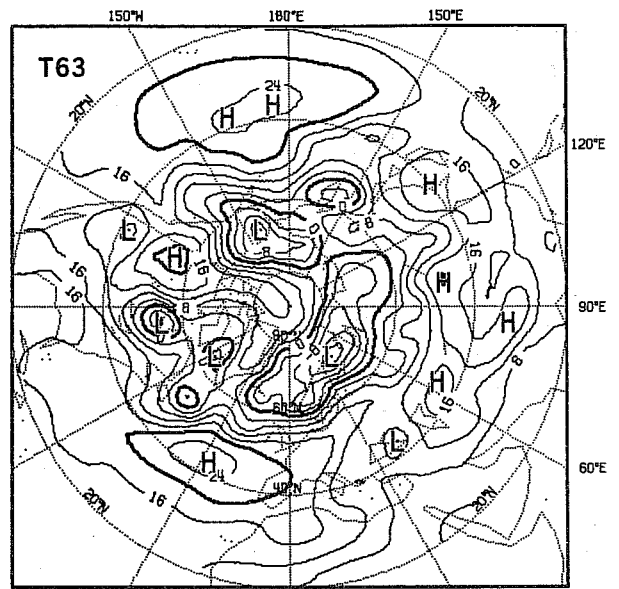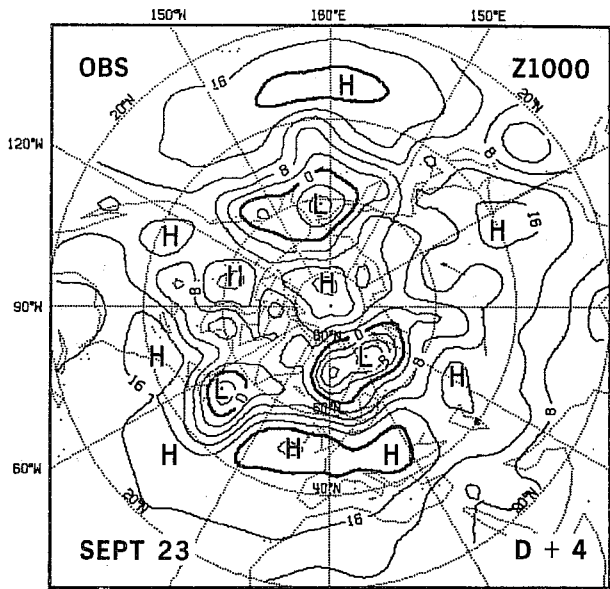Fig.16a (Cont.)

39

Fig.16b SEP. 23, Z1000: D+4, D+5, D+6
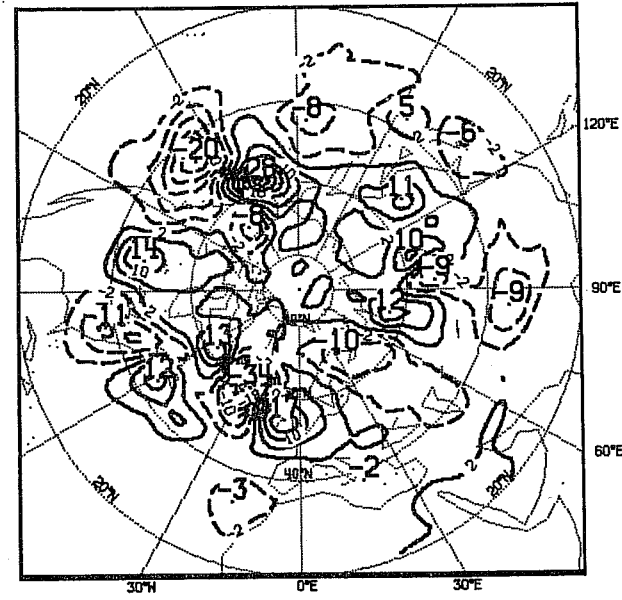
Fig.16b (Cont.)
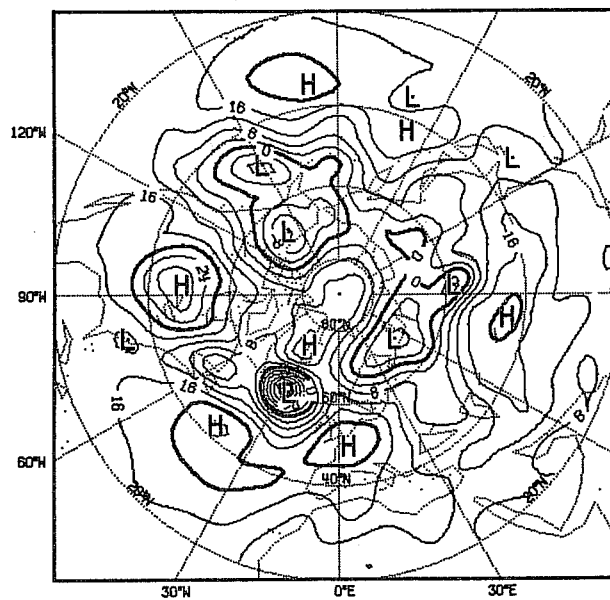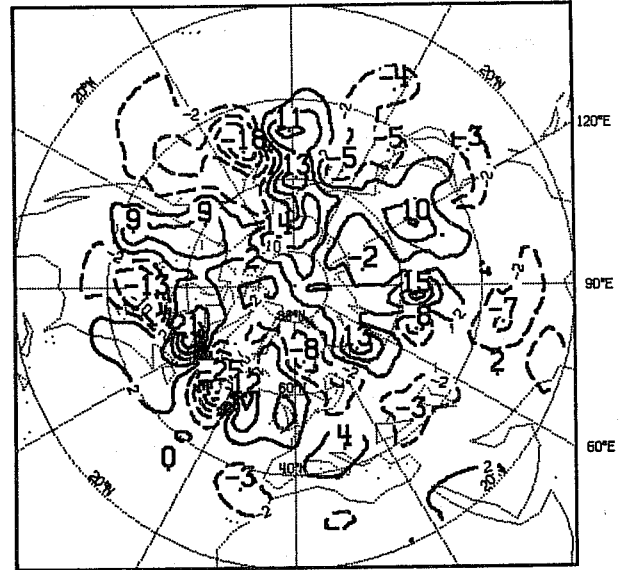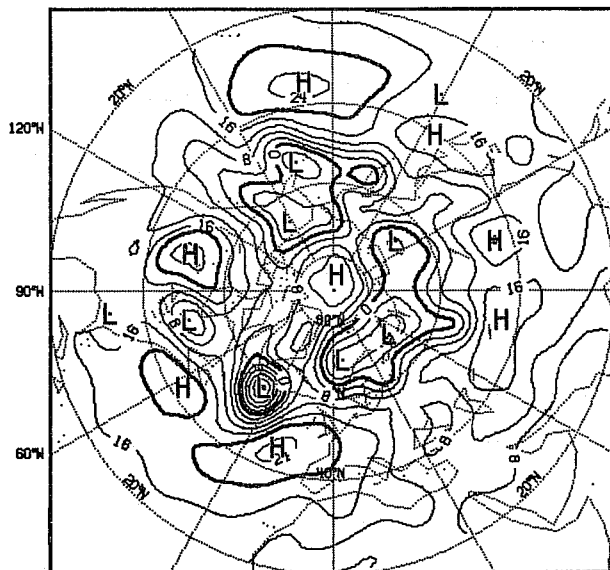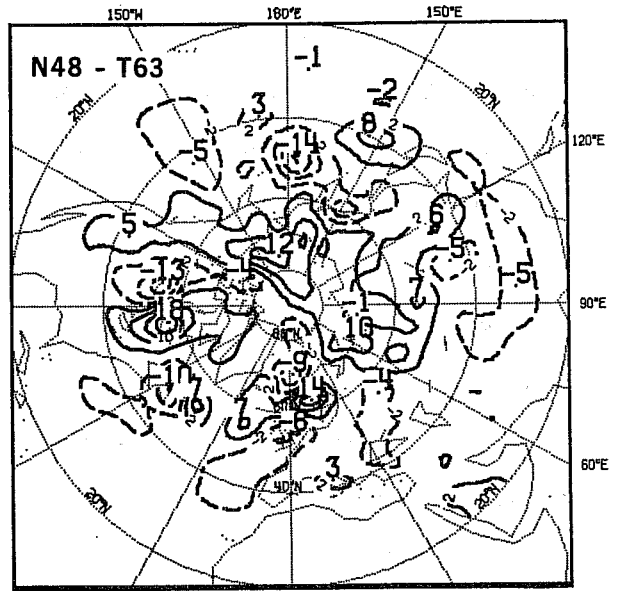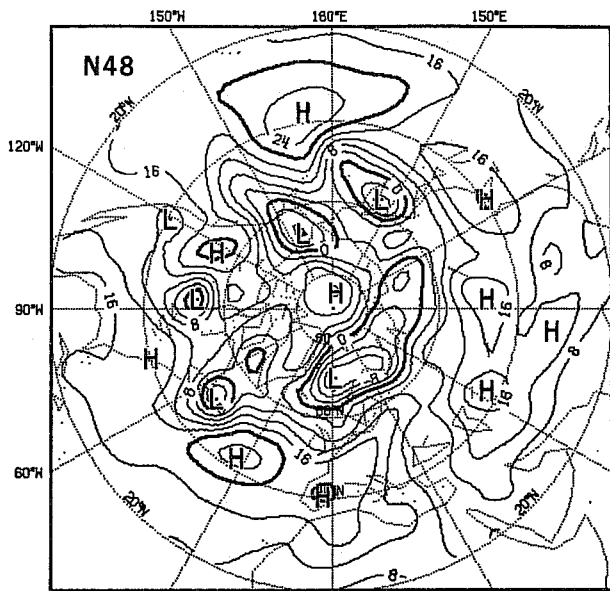
relatively zonal circulation in OBS. However, N48 has better success in predicting the structure over east ATL and EUR. T63 has never developed the trough over ATL and has moved it too near EUR. On the other hand, T63 has produced a better forecast over PAC, despite a phase error on the central trough.

At 1000 mb (Fig.16), the models are unable to predict the decay of the low over ARC north of Russia from D+1 to D+3 and T63 overdevelops the other one north of EUR. On D+5 and D+6, N48 correctly forecasts the development over ATL while T63 predicts a much weaker development there and a too far east position of the low. On the other hand, on D+6, T63 predicts a strong development over PAC which does occur but is further south in OBS while N48 misses it.

At 500 mb, there are fast moving short-wave troughs which we can follow throughout the 6-day forecast period. For example, the small trough on the northern part of the stream, associated with the cut-off low over EUR (10°E) on D+1 is observed moving to 40°E, 65°E, 70°E, 85°E and 105°E respectively from D+2 to D+6. Other troughs upstream of it can also be followed clearly. Similarly a trough observed over EAS (110°E) on D+1 moves to 120°E, 130°E, 165°E and 165°W respectively. Already on D+2, a string of FD centres with alternating signs has developed over east ATL and north of EUR in the N48-T63 maps. Comparing the troughs' position in both forecasts, we may interpret these as slight phase differences. T63 troughs advance with respect to those of N48. Phase differences are, of course, also modulated by amplitude differences. On D+3, for example, a well organized pattern is only apparent for the trough over Scandinavia. But on D+4, the differences show up more clearly. The string of alternating positive and negative centres aligned along the jet extends essentially from ATL to EAS. On D+5 and D+6, both the amplitude and scale of the pattern have increased.

Over PAC and EAS, a similar pattern becomes apparent by D+3 but it is already of much larger scale. By D+5, phase differences are modulated by large amplitude differences.

Moreover, one also observes a characteristic phase difference pattern over the west coast of North America, on D+2. Indeed, a weak trough is observed there with N48 lagging behind T63 which is about in phase with OBS. On D+3, the trough and, with it, the difference pattern have moved well over WAM. From D+3 toD+6, the trough moves across WAM.

The net result is that, by D+6, the pattern seen on the difference maps is mostly composed of this string of alternating difference centres circling the hemisphere and having their extrema near the jet core. The scale and amplitude of the pattern is also well correlated with the scale and amplitude of the travelling perturbations themselves.

At 1000 mb, the features on the difference maps can be matched with the ones on the 500 mb maps, almost one to one, day by day. Their scale and amplitude are very similar.

Before ending our discussion on this case, we would like to attract the attention of the reader to the relatively large area of weak FDs at 500 mb in the subtropics. These are already present on D+1 but do not grow with forecast length. They are entirely due to differences between the PP packages used by T63 and N48. Similarly, at 1000 mb, large area differences are present over regions with high terrain: Himalayas, Rockies and Greenland. These will not disappear completely with the change of post-processing as the mountain fields will remain slightly different, because of the spectral analysis, but they will greatly diminish. The impact of these differences on the objective performance of T63 has been mentioned earlier and will be further analysed in Appendix A.
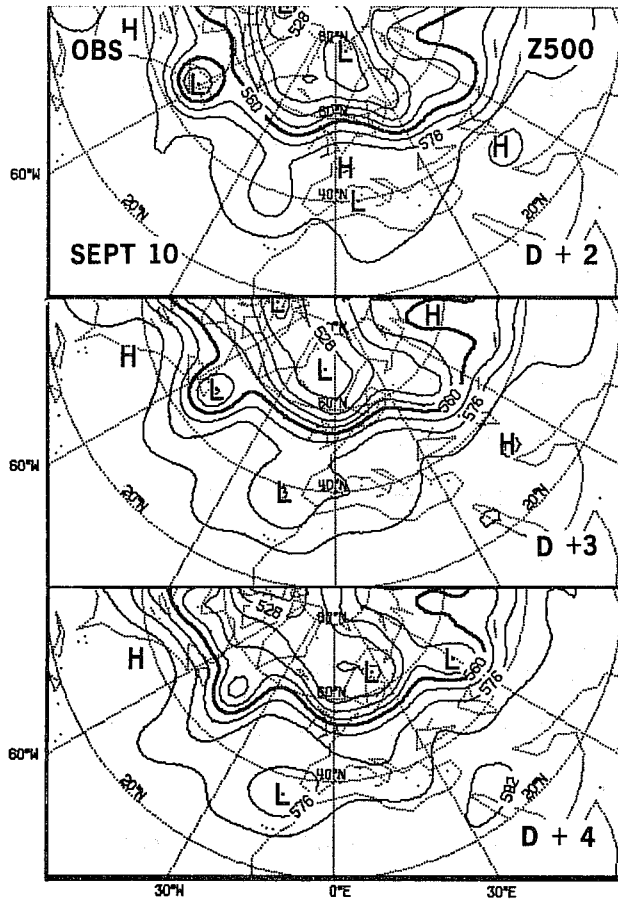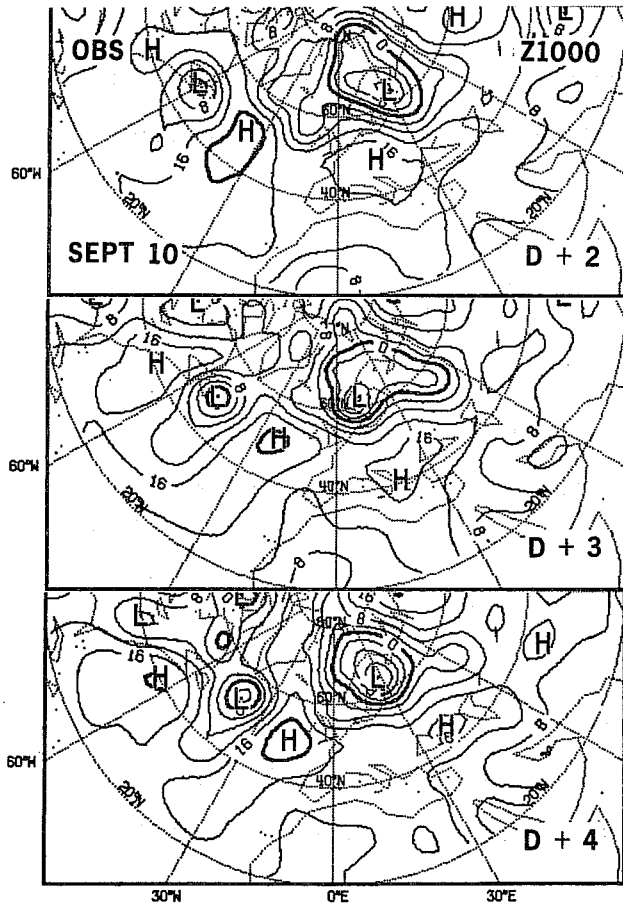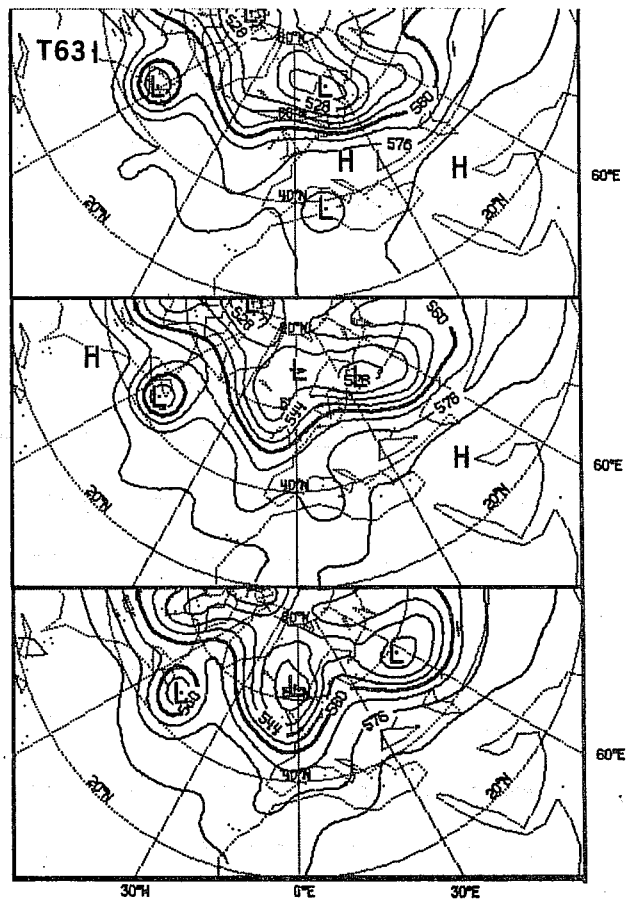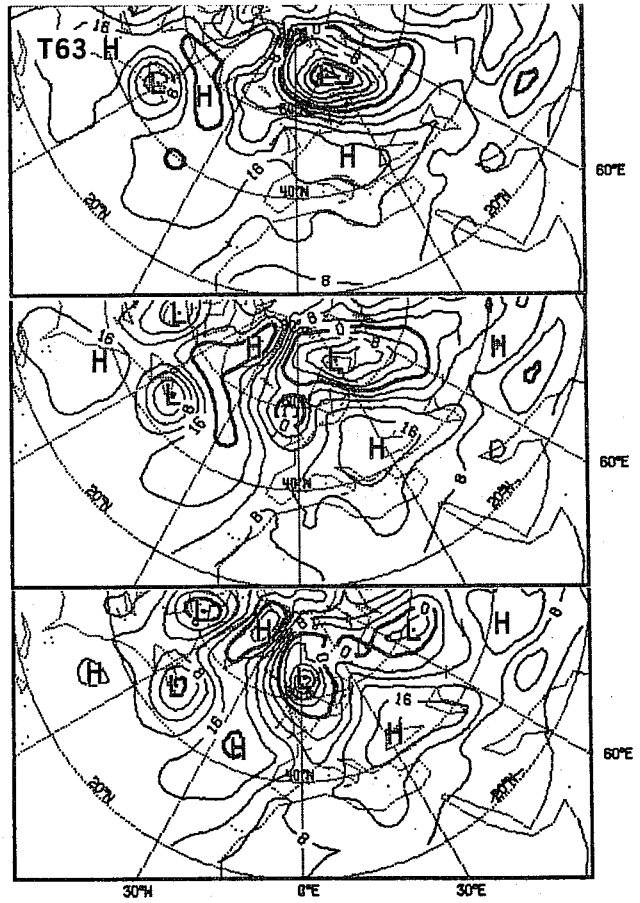
Fig.17   SEP. 10, Z500: D+2, D+3, D+4
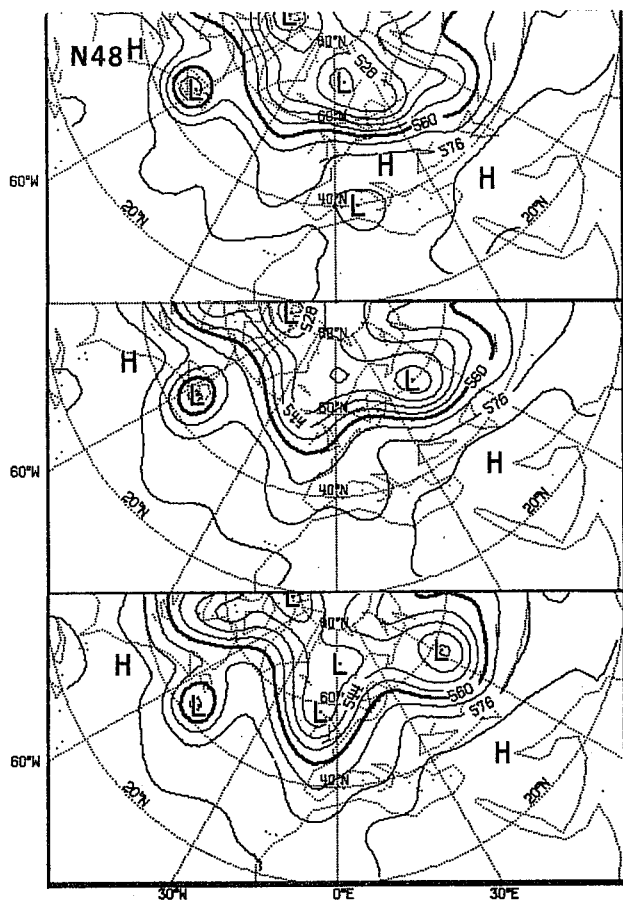

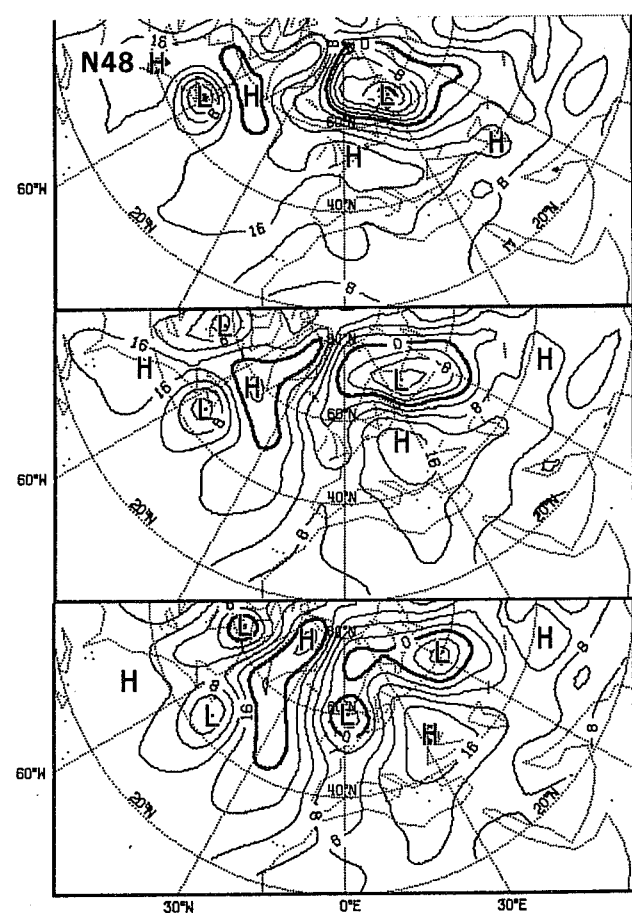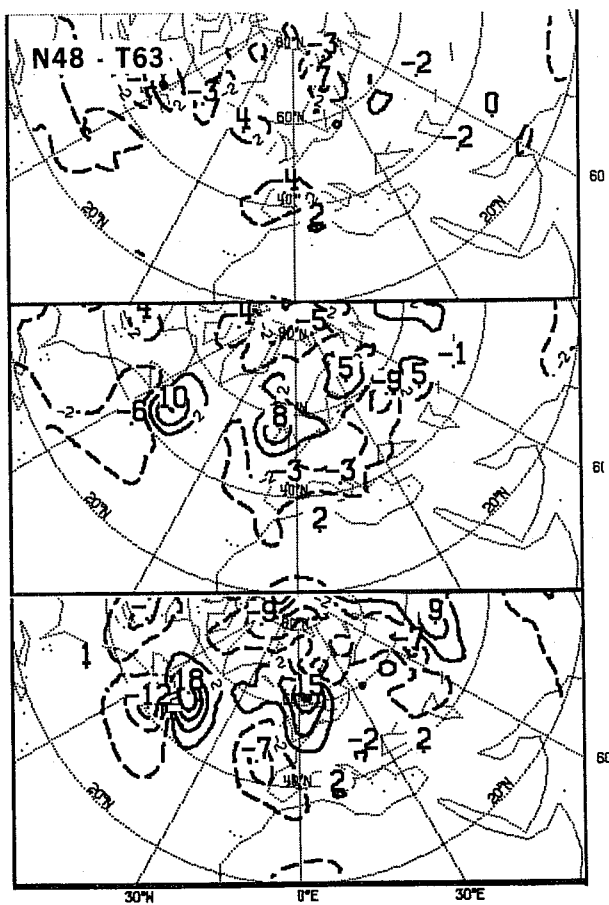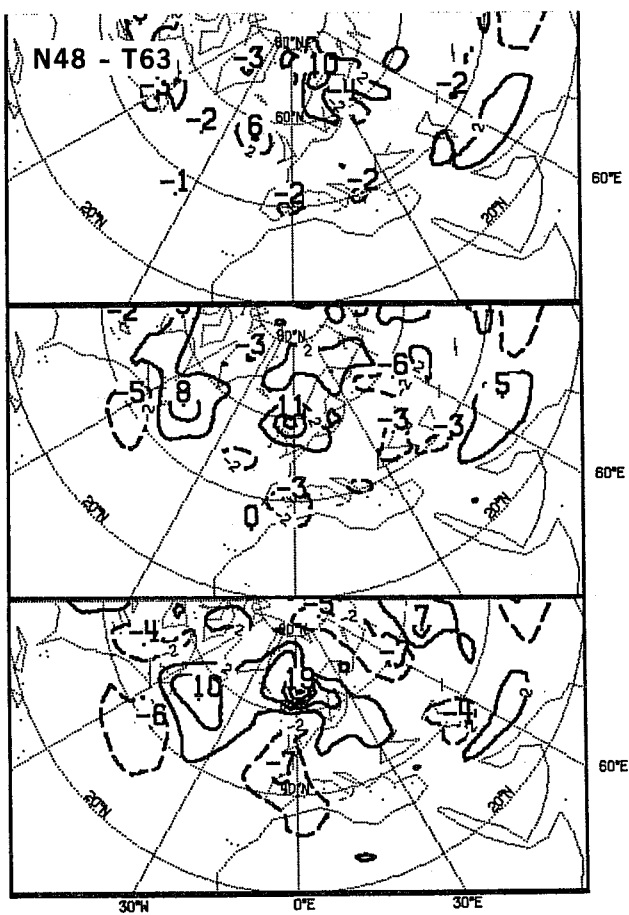Fig.18   SEP. 10, Z1000: D+2, D+3, D+4

Fig. 17 (Cont.)

Fig. 18 (Cont.)

Case of SEP 10

Fig. 17: Z500, D+2, D+3, D+4

Fig. 18: Z1000, D+2, D+3, D+4

At 500 mb, the features of interest are a trough moving over England and the cut-off system upstream. The models forecast an exaggerated intensification of the trough and the ridge upstream and they fail to merge the cut-off with the main stream. As seen from N48-T63 maps, T63 features are deeper than N48's. They follow a more northerly track and are ahead of N48's (in closer agreement with OBS).

At 1000 mb, the features of interest are a trough north-west of England which is absorbed by the large cyclone over Scandinavia and moves quickly north-eastward and the low over the North American east coast which moves offshore and deepens slightly. The trough in T63 overdevelops over Scotland on D+3, moves northeastwards on D+4 but is far behind OBS. The trough in N48 only develops between D+3 and D+4 and is far south of T63 then. An important meteorological consequence is a complete misrepresentation of the high pressure belt over east ATL and southern EUR by N48. Similarly the east coast low has the correct intensity but is too slow in T63. It is less intense and even slower in N48.

The difference maps show that T63 is far deeper than N48 throughout north ATL. The error in T63 is an error of phase rather than of over-development while that in N48 is of both phase and amplitude. However, note that both models have indeed overdeveloped the large cyclone over Scandinavia on D+2, T63 even more than N48.

Such overdevelopments are a frequent feature of both models and, as seen here, they may occur in the early stages of the forecasts. Very often also, in these cases, T63 tends to overdevelop even more than N48.

The maximum difference of 10 dam seen on D+2 on the difference maps is one of the largest differences observed at this range during the course of these experiments (Table 4).

Period from SEP 29 to OCT 8

On SEP 29, the flow becomes much more meridional as we witness the formation of a blocking high over EUR and intensification of the ridge over WAM and troughs over PAC, ATL and WAS (Fig.19 and 20, OBS). The flow remains quasi-stationary until OCT 8. On SEP 30, both forecasts are good, especially at 500 mb, but T63 is better. The suggestion that a spectral model may not do as well as a grid point model in quasi- stationary situations has often been put forward. The following example shows this not to be the case.

Case of SEP 30

Fig. 19: Z500, D+2, D+4, D+6

Fig. 20: Z1000, D+2, D+3 ,D+4

At 500 mb, both models are remarkably successful in maintaining the phase of the quasi-stationary long wave pattern (Note that the maps shown are from D+2 to D+6!). They are less succcessful in maintaining its amplitude: as the forecast range increases, the forecasted jet becomes more zonal, particularly between 90°W and 90°E.

T63 is more able to maintain the amplitude of the troughs (D+2, D+4). On D+2, all the main areas of differences are positive showing that T63 is deeper than N48 over EAS, over WAM and over the ARC north of WAS. On D+4, the amplitude differences over EAS and WAM have increased further. A short-wave trough superimposed on the long-wave ridge positioned over WAM is better indicated by T63 which also better differentiates the structure of the troughs over PAC and ATL. On D+6, the low observed over ARC north of the Hudson Bay is much better represented by T63.
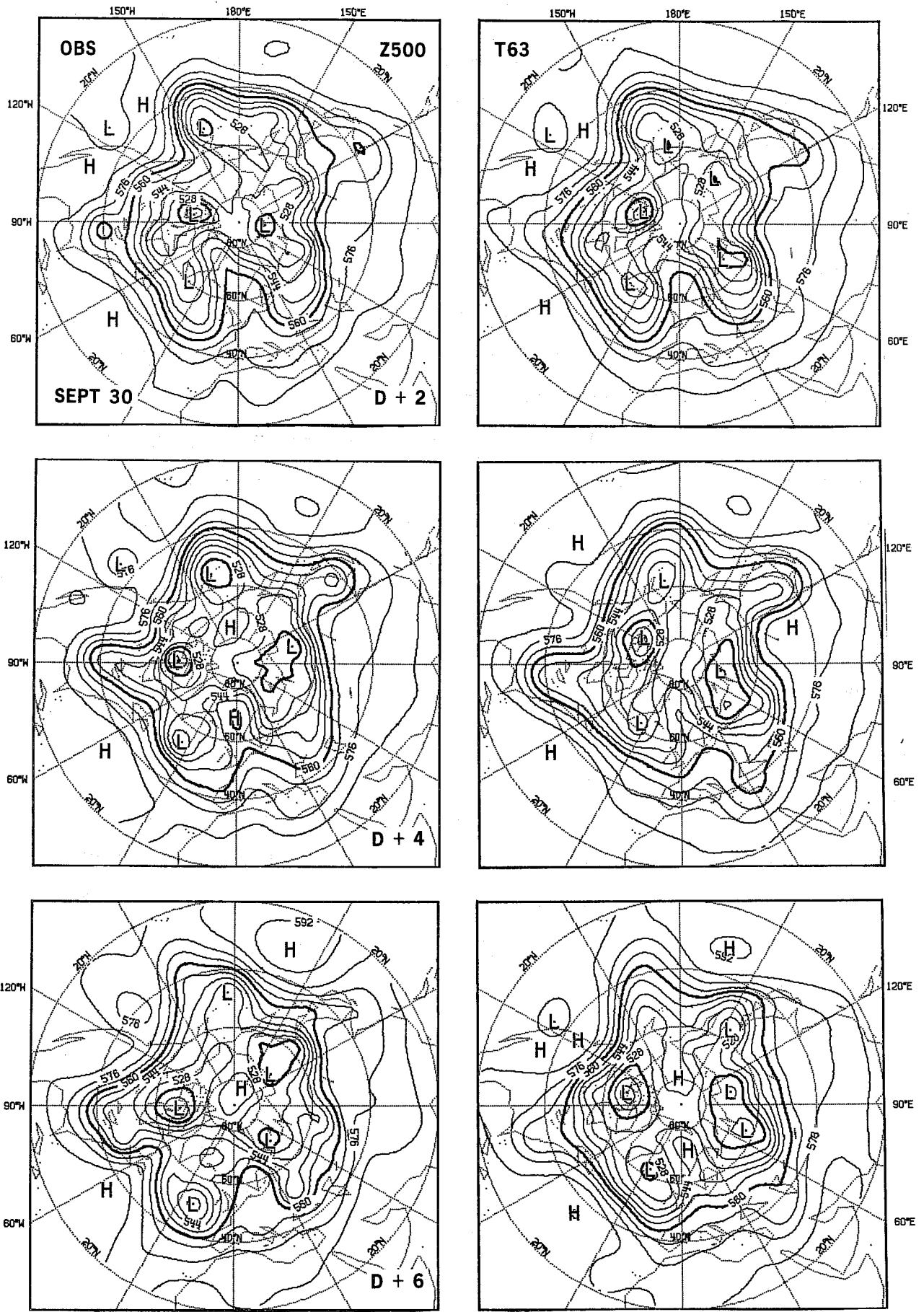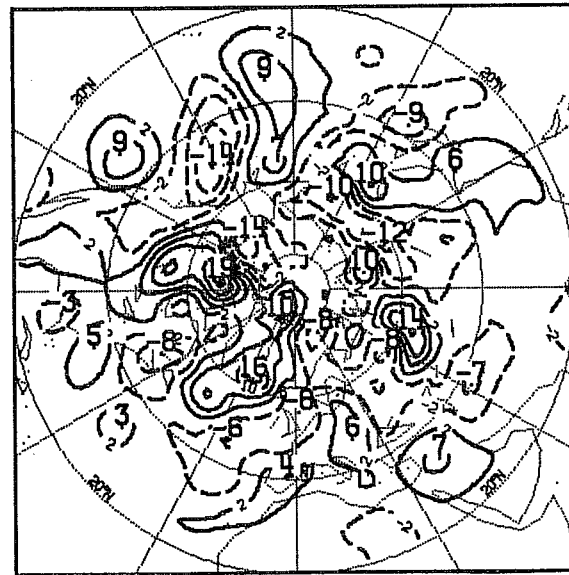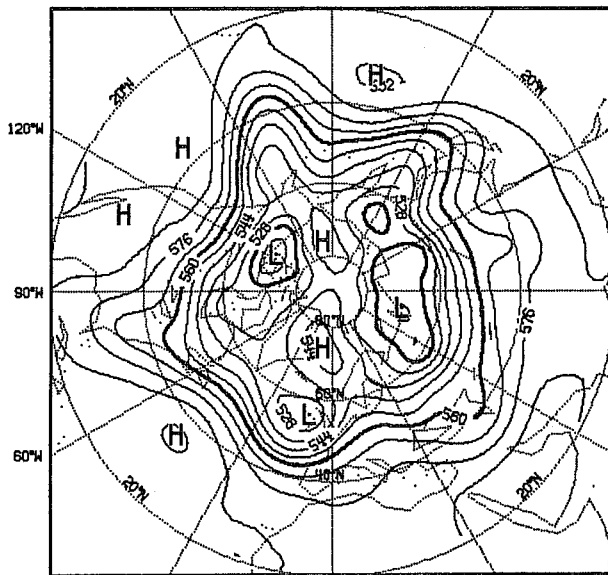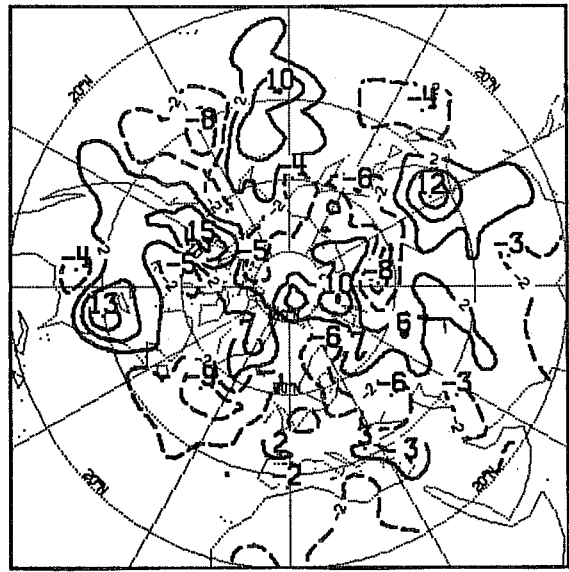
47

Fig.19 SEP. 30, Z500: D+2, D+4, D+6
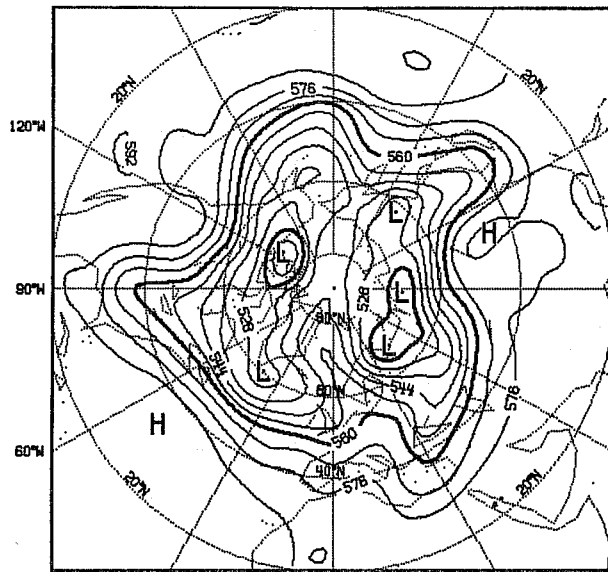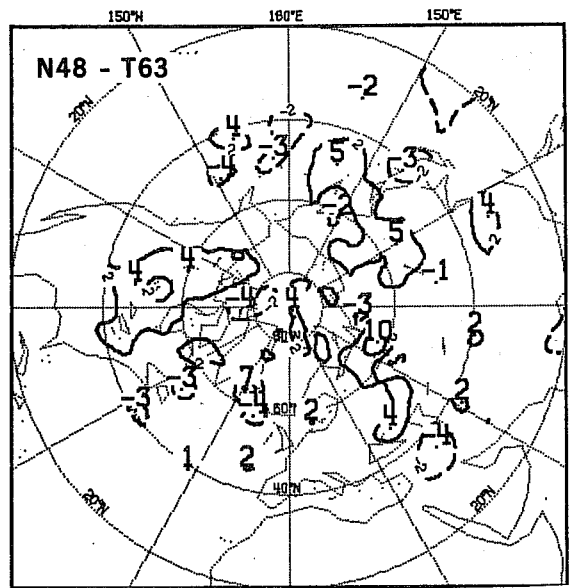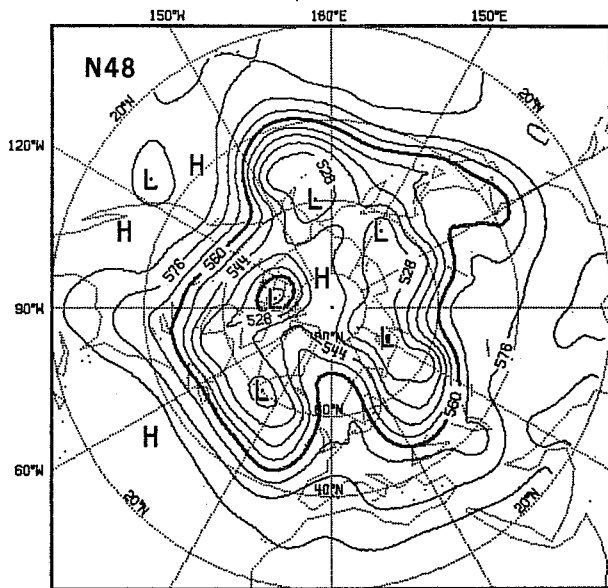
Fig.19 (cont.)
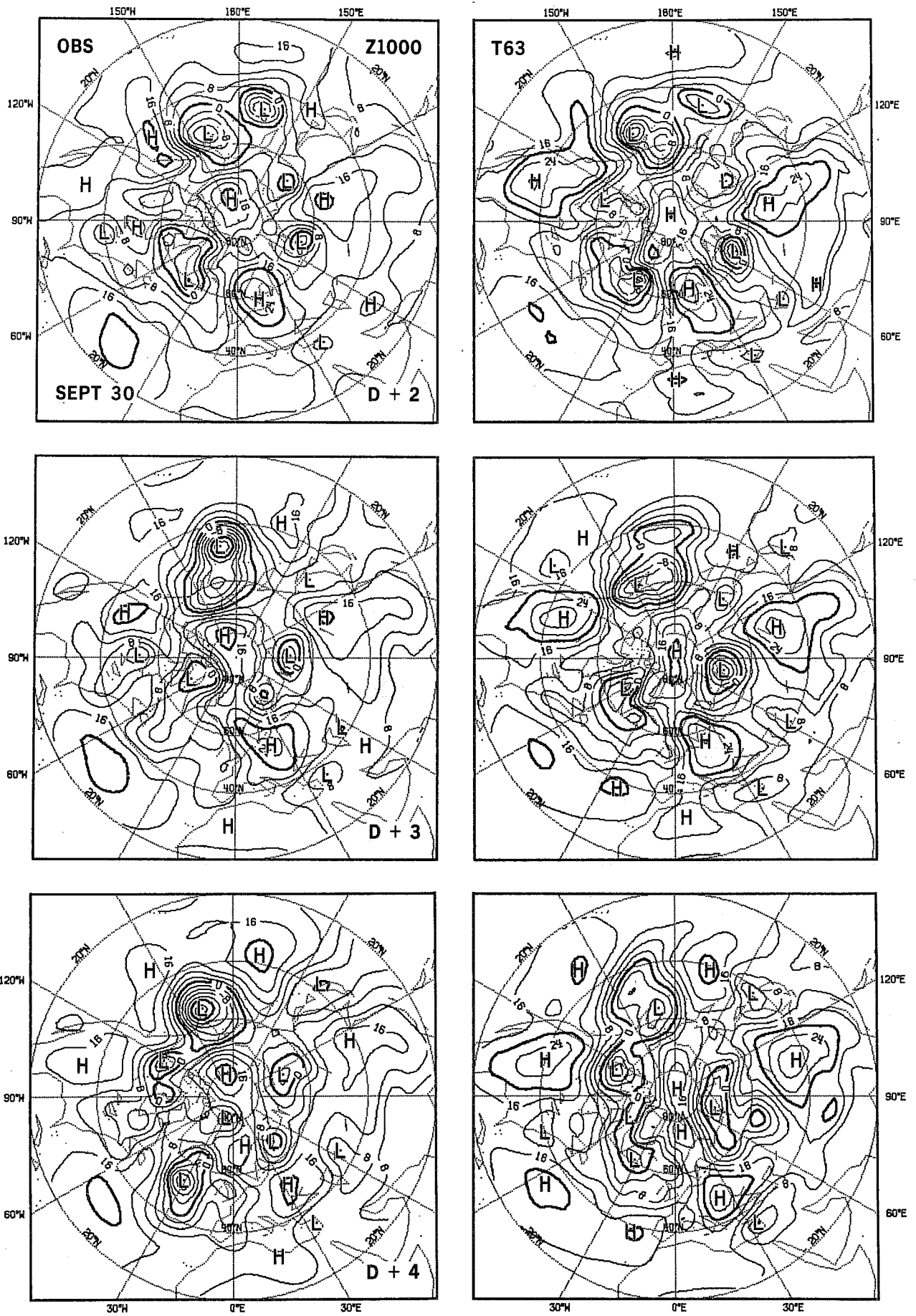
Fig.20   SEP. 30, Z1000: D+2, D+3, D+4

Fig.20 (cont.)

51
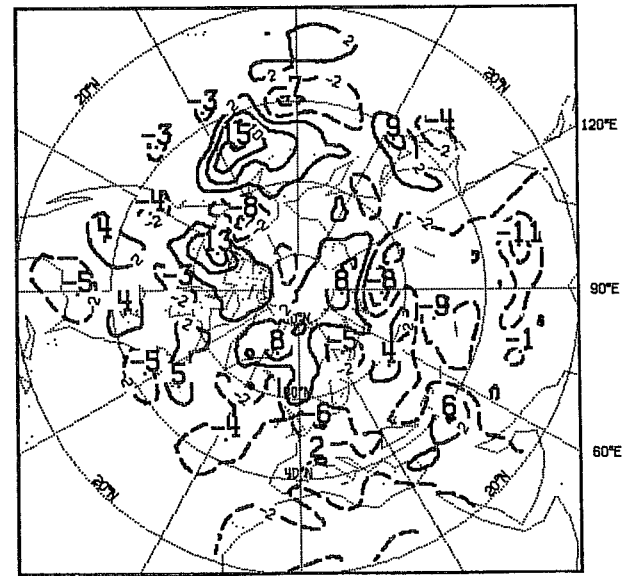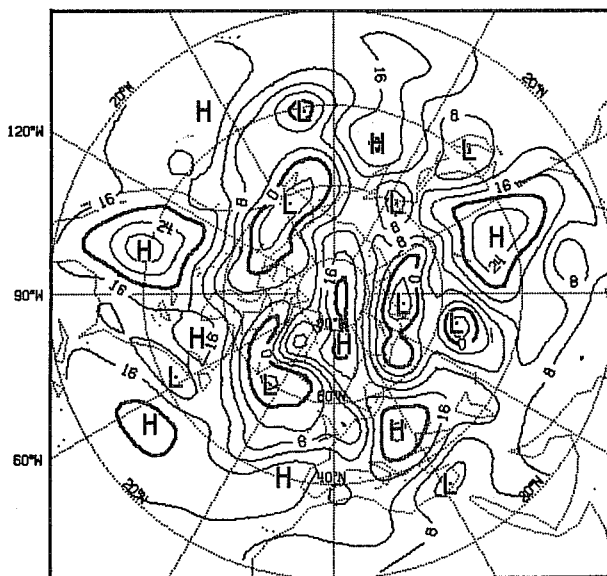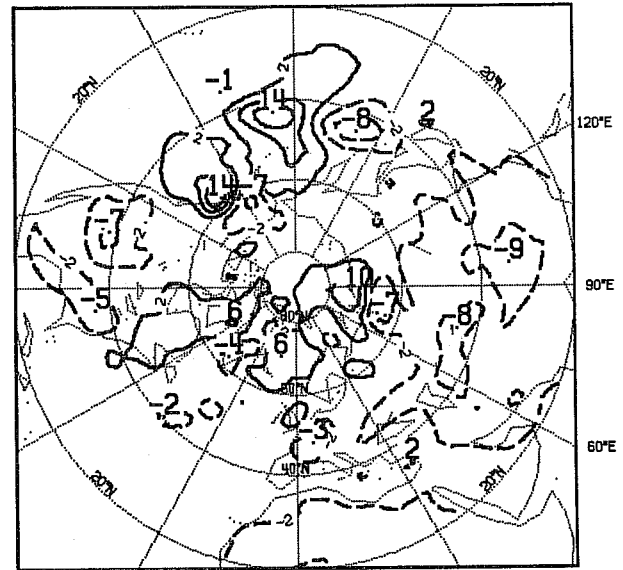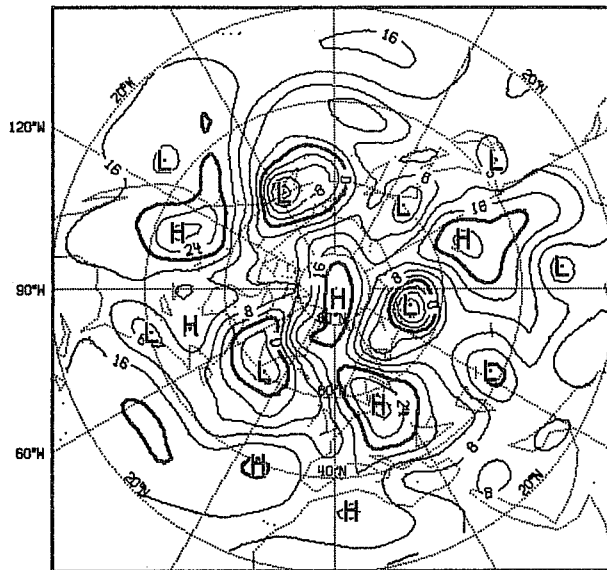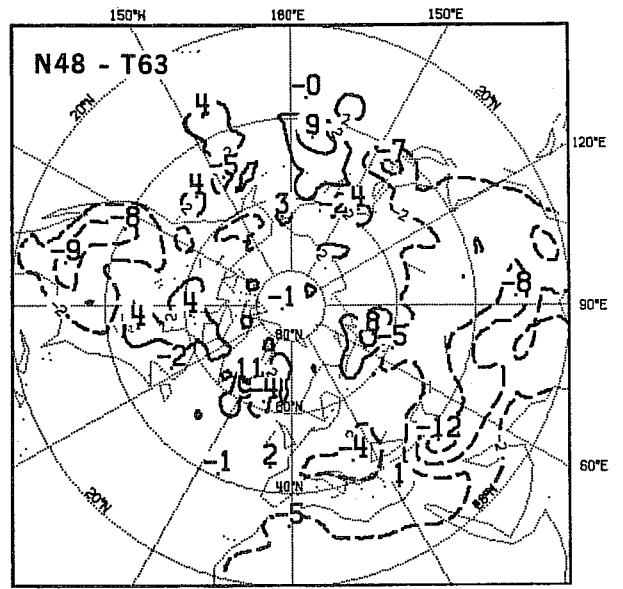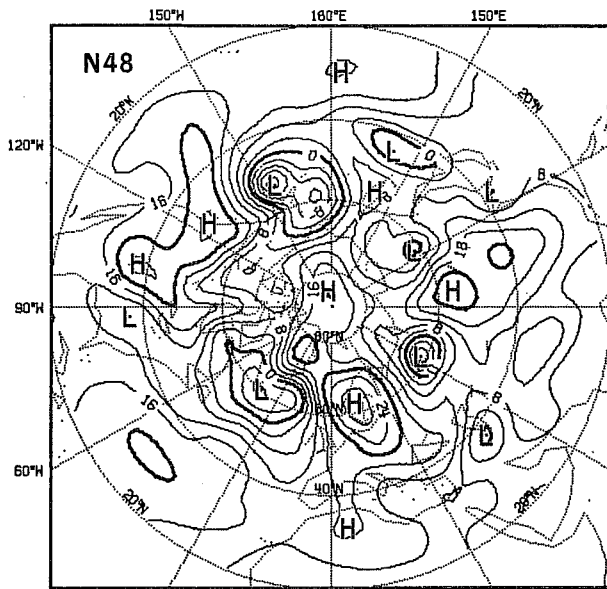
At 1000 mb, note once again the tendency of the models to overdevelop the main lows in their mature stage, over PAC and over north ATL (T63 only) on D+2 and over WAS on D+3, while they fail to capture the new developments, over PAC and over ATL on D+2 and D+3. However, T63 offers a more credible evolution of a new low over PAC. It forecasts more correctly a more northerly track for the low over WAS and better captures on D+4 the development of a new low over north west Canada (WAM).
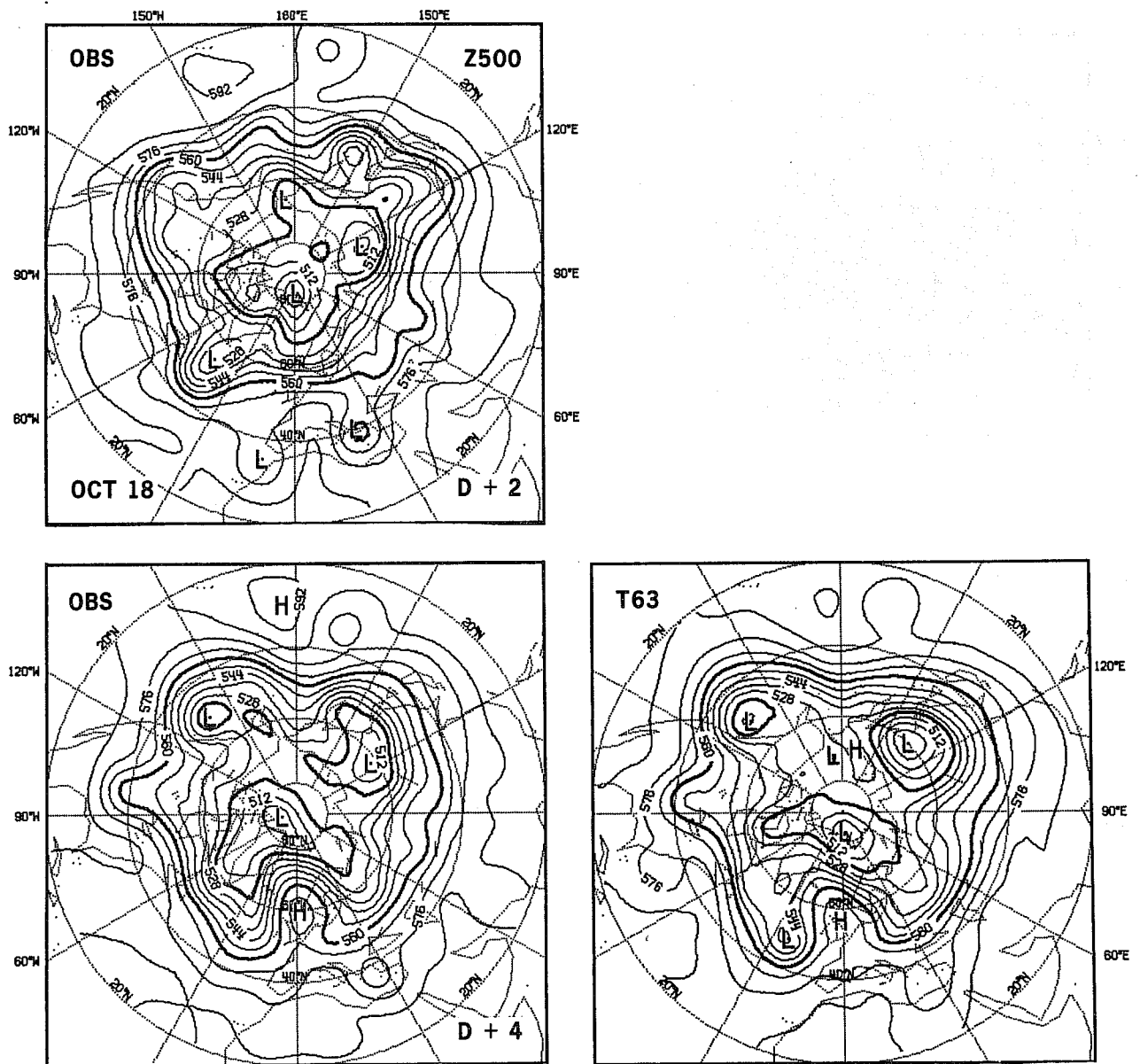
Fig. 21 OCT. 18, Z500 D+2 OBS (top); D+4 (bottom).

Period from OCT 9 to OCT 30

Following the decay of the blocking situation over EUR, the 500 mb long-wave pattern becomes more transient. The models do not score well then. But around OCT 20, heights start rising again over EUR and by OCT 22, a sharp ridge has developed (Fig.21, OBS). In the following days, the blocking high strengthens further and moves slowly north-eastward over Scandinavia towards the ARC (Fig.22, OBS). The onset of the block and its evolution are well caught up by the models on the forecasts from OCT 18 and OCT 25. But the FDs are quite large on the OCT 25 case. In both cases T63 performs better than N48.

Case of OCT 18

Fig. 21: Z500, D+4

The figure displays OBS on D+2 in addition to D+4 showing the formation of the ridge over EUR in the interval. The largest FDs are in the same region (over north EUR and ATL). The pattern observed is again that of a string of centres with alternating signs aligned, not along the D+4 position of the jet, but rather along that seen on D+2. The pattern is well correlated with

Fig. 21 (Cont.)

53
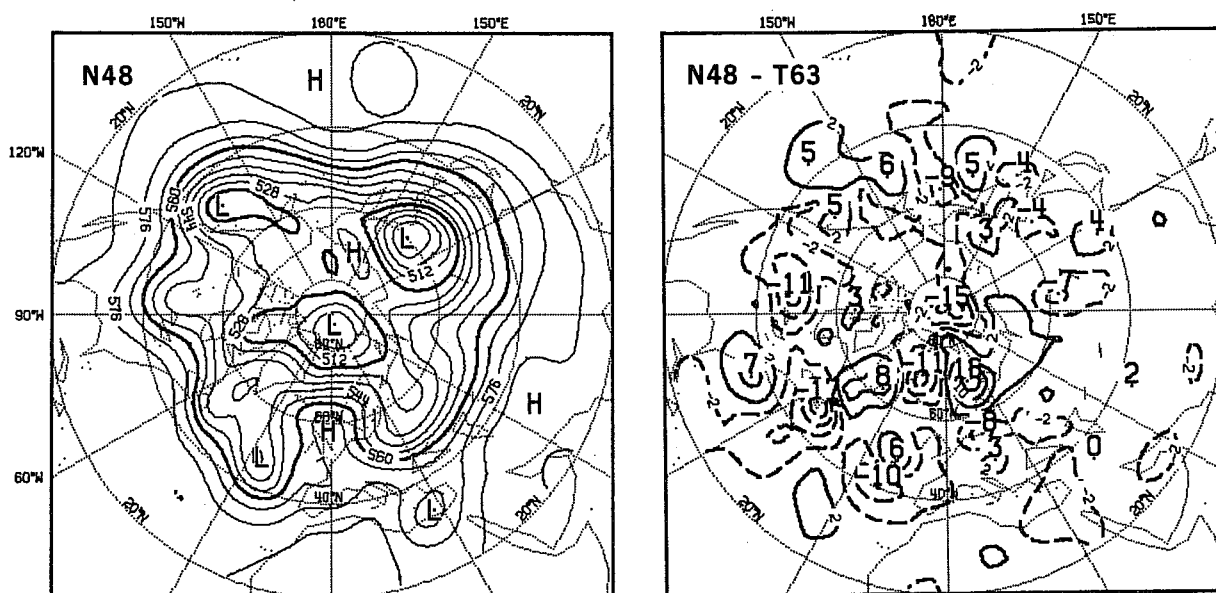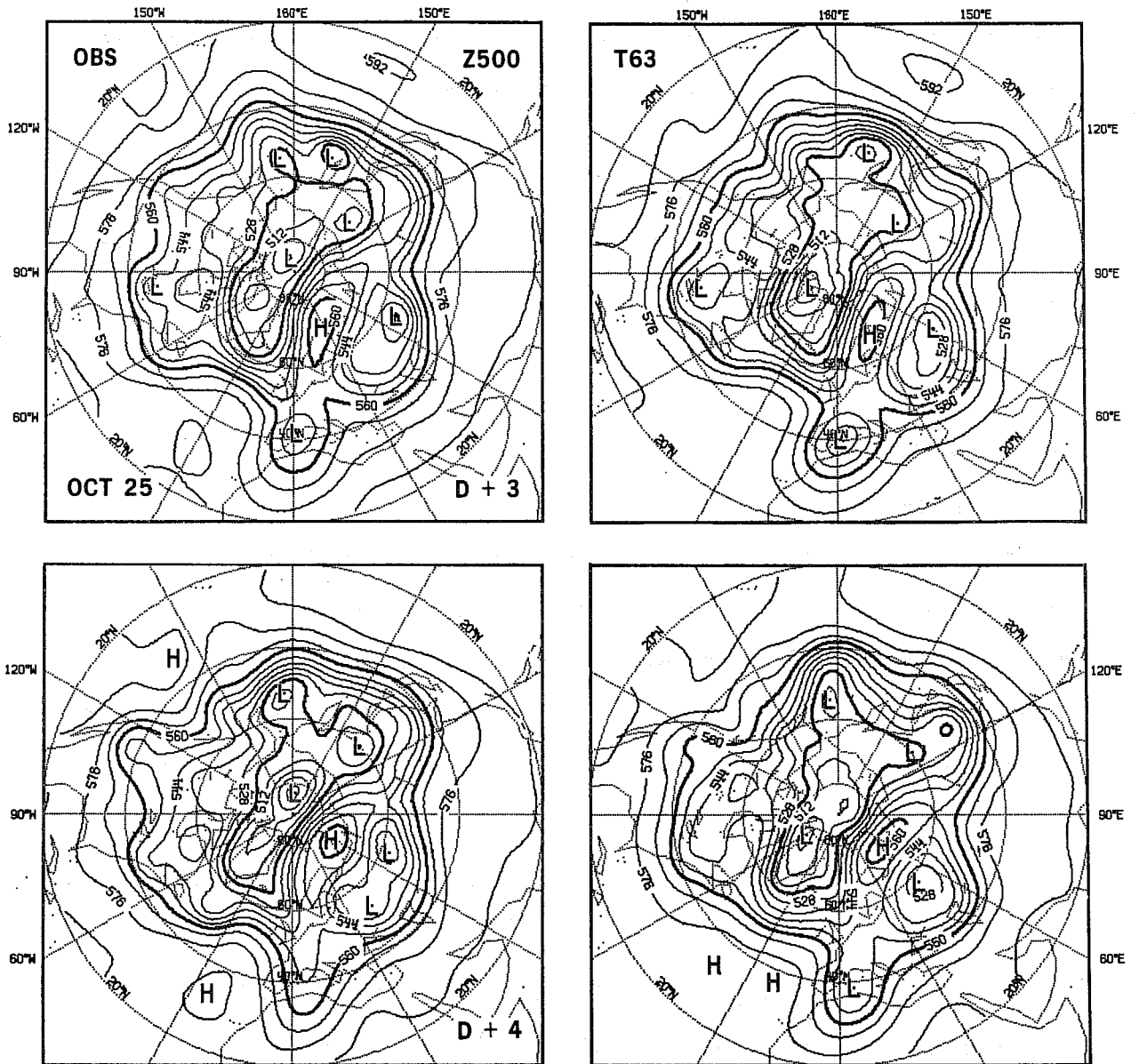
Fig.22 OCT. 25, Z500: D+3, D+4
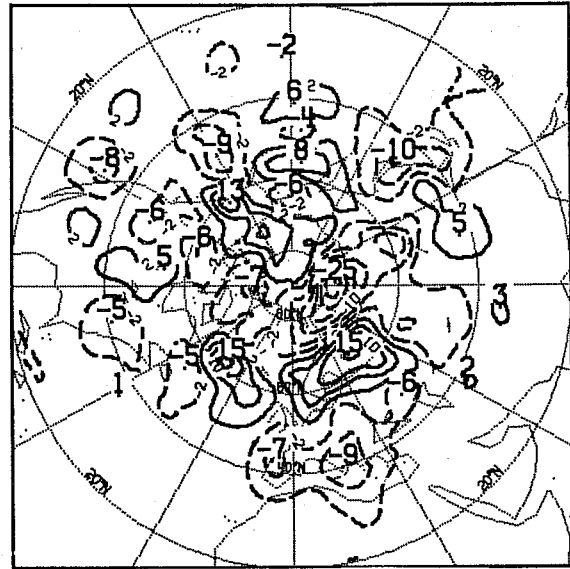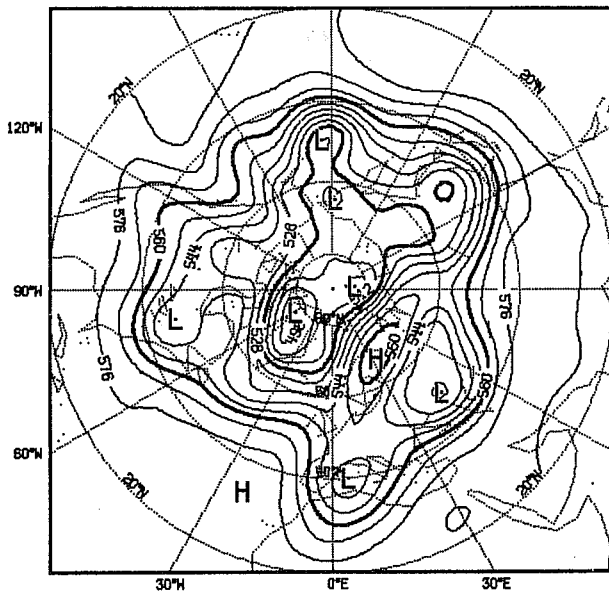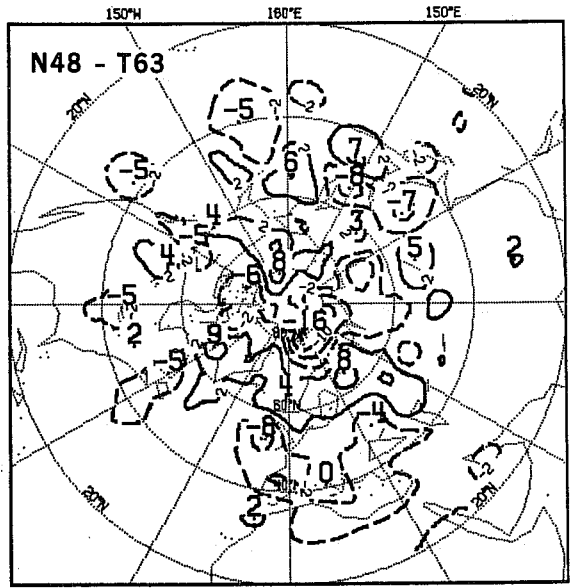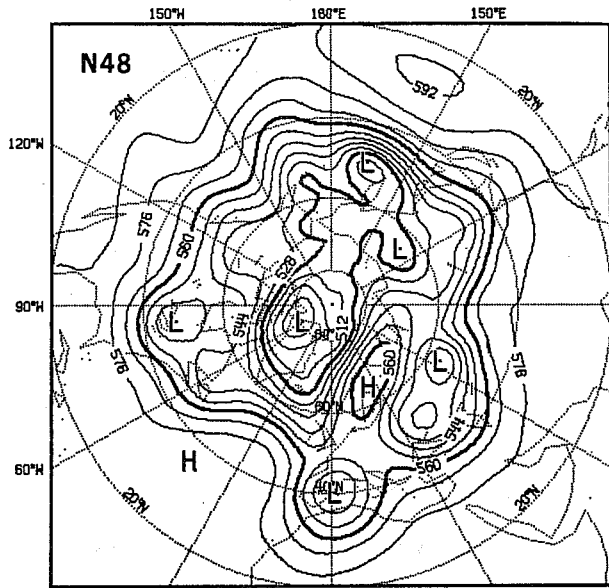
Fig.22 (Cont.)

a large phase difference of the weak trough seen over north Scandinavia in T63 and east of Greenland in N48 for the first dipole and of the small low south-east of Greenland in T63 and south-west of it in N48 for the second dipole. It can also be verified that the same systems in T63 nevertheless still lag behind OBS. The T63 structure of the troughs over ATL and WAM is also more realistic.

Case of OCT 25

Fig. 22: Z500, D+3, D+4

The main FD is the high over ARC which is more in phase with OBS in T63 and also further north than in N48. Note the scale and amplitude of the difference. Amongst the other differences on D+4, we can mention a phase difference in the minor trough over east PAC in OBS, an amplitude and phase difference in the trough south of Greenland. Finally the large trough over Western Europe is broader and deeper in T63 than in N48.

—

Period from NOV 1 to DEC 5

From NOV 1 onwards, the 500 mb long-wave pattern becomes and stays more transient. The models perform very well throughout NOV with a mean P(50%) of more than 6 days at 500 mb and more than 7 days at 1000 mb. Note incidentally that mean predictability for NOV is higher at 1000 mb than at 500 mb. T63 achieves nearly systematically better scores up to D+5, with very large differences in the large scales (wavenumber group 1-3). On these cases it will be interesting to relate these differences with the structure of the N48-T63 maps. For this period we shall concentrate our attention on the largest FDs (Table 4), showing the D+5 forecasts from NOV 1 and the D+3 forecasts from NOV 15 and NOV 29.

Case of NOV 1

Fig. 23a: Z500, D+5

Fig. 23b: Z1000, D+5

Although part of the FDs can be attributed to phase differences another part must also be attributed to amplitude differences. For example, over EAS and PAC, T63 has its systems not only better in phase with OBS but equally much more intense. In particular, over PAC at 500 mb, the intense trough-ridge-trough pattern, well captured by T63, is almost inexistant in N48. The 47 dam difference seen over EAS at 1000 mb is the largest one for this series of experiments on D+5. There, as often, T63 is even deeper than OBS. Over EUR, the structure of the low is apparently better in N48, but T63 has the low centre in a more correct position although typically too deep.

Case of NOV 15

Fig. 24: Z1000, D+3

On D+3, the two main FDs form two dipoles both with a SW-NE orientation. Here are typical examples of N48-T63 differences. N48 moves developing systems more slowly, following a track further south than does T63. This commonly occurs over the oceans: ATL and PAC and also EAS (where one can observe a smaller structure of the same type associated with a smaller third low). All differences are in favour of T63.

Case of NOV 29

Fig. 25: Z1000, D+3

There is only one area of large FDs on D+3 in this case. The type of pattern is not substantially different from the ones in NOV 15. The leading edge of the huge low in T63 is further east and slightly further north and has a better orientation. The cyclonic circulation in both models is characteristically too strong. On **this case, it is easy to understand why** the improvement brought by T63 benefits mostly to the long waves.

57
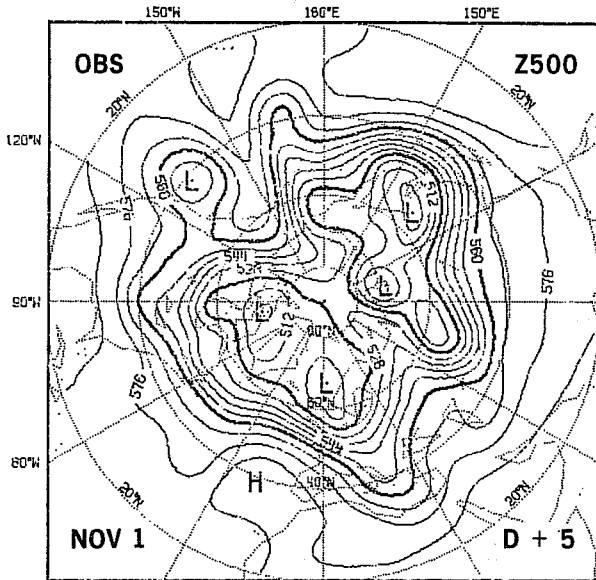
Fig.23a NOV. 1, Z500: D+5
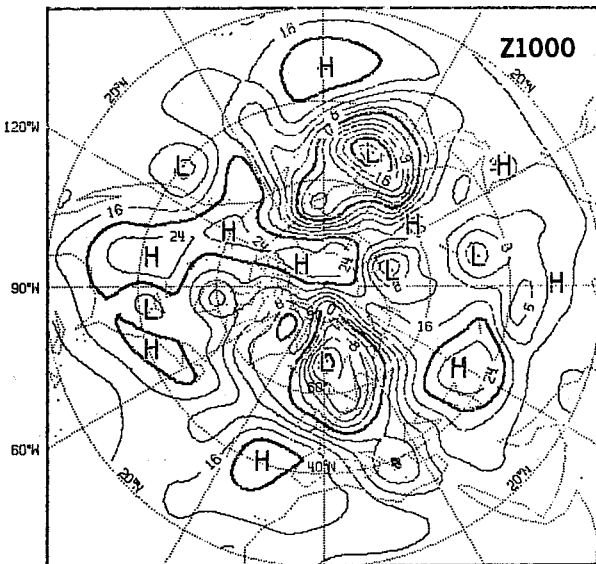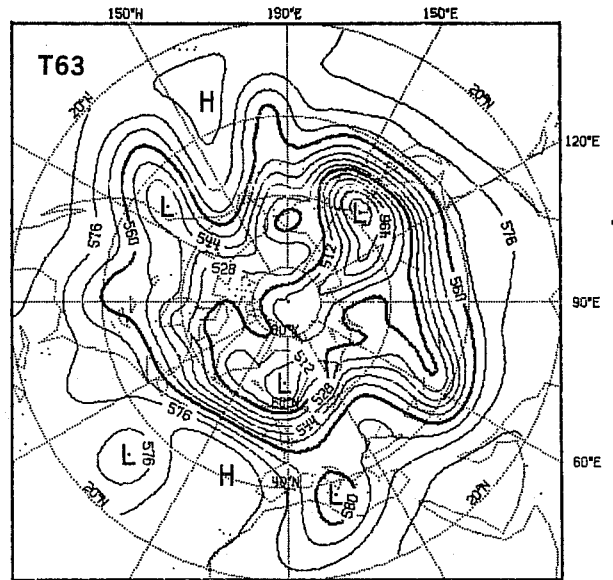
Fig.23b NOV. 1, Z1000: D+5
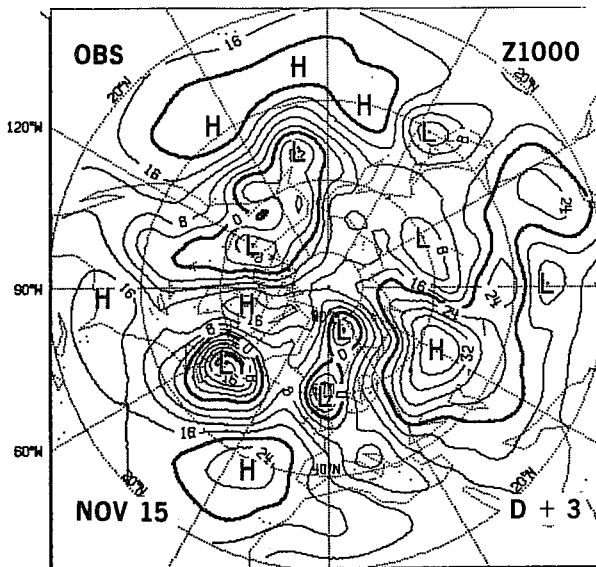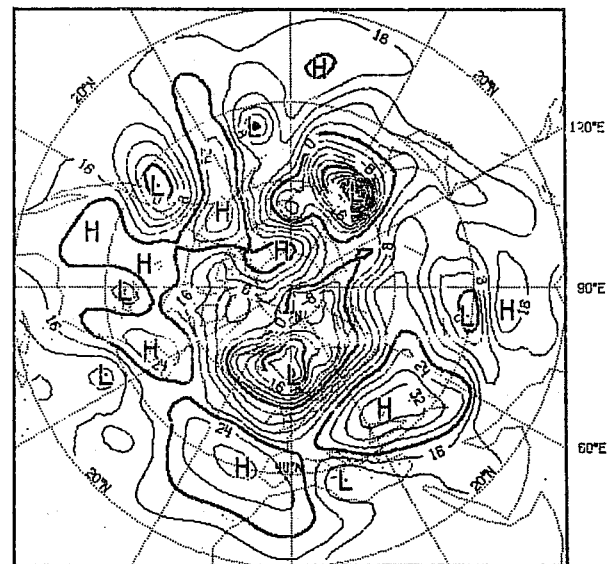
Fig.24 NOV. 15, Z1000: D+3
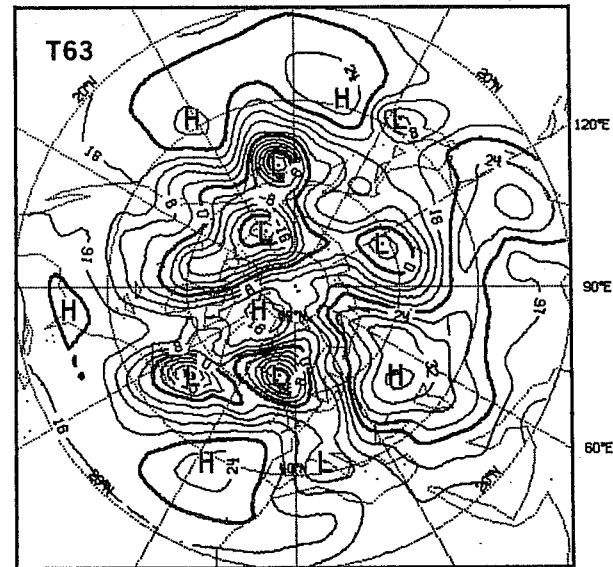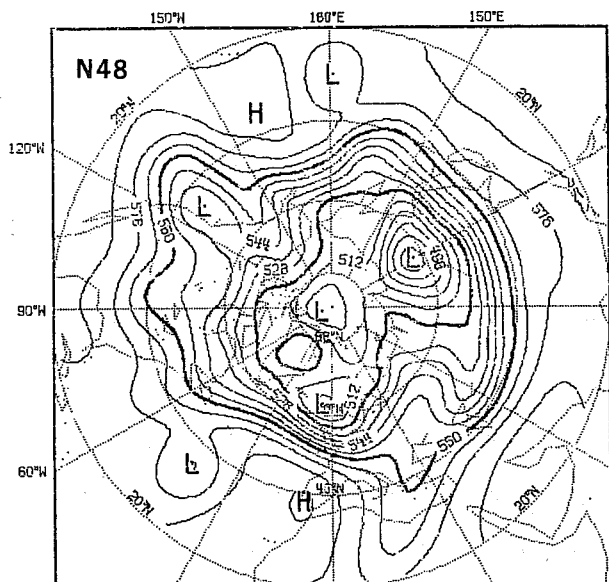
Fig. 23a (Cont.)
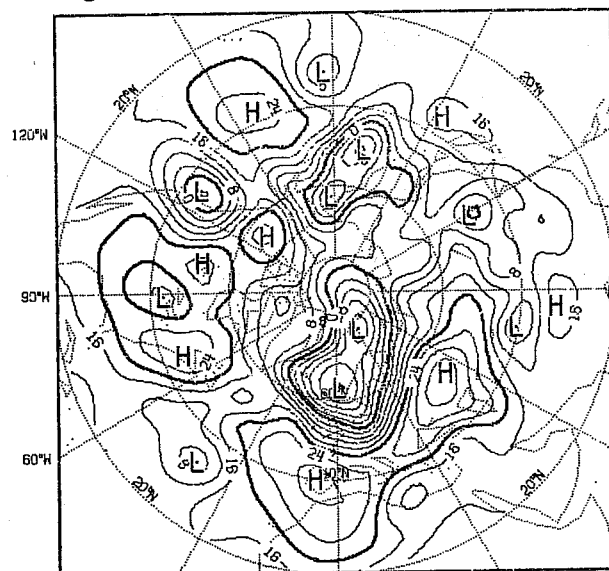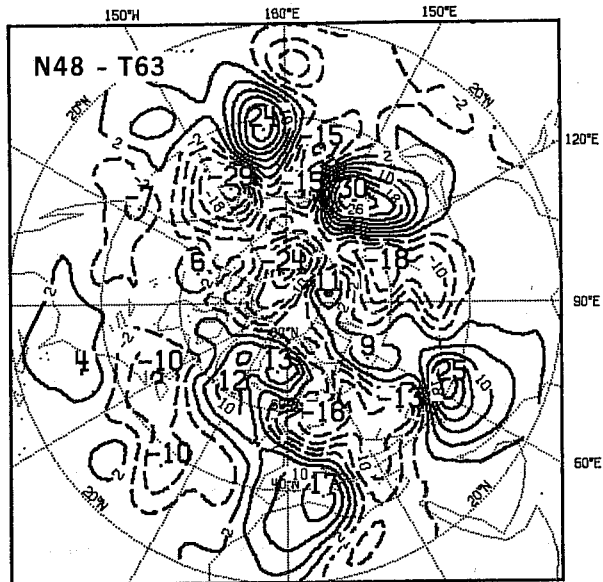
Fig. 23b (Cont.)
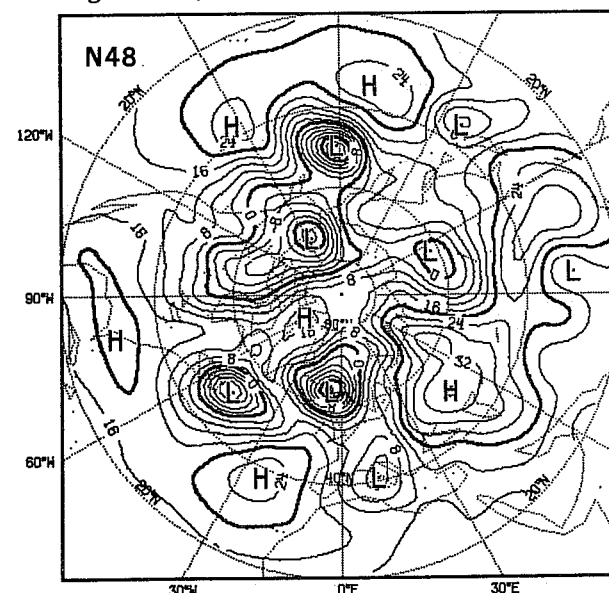
Fig. 24 (Cont.)
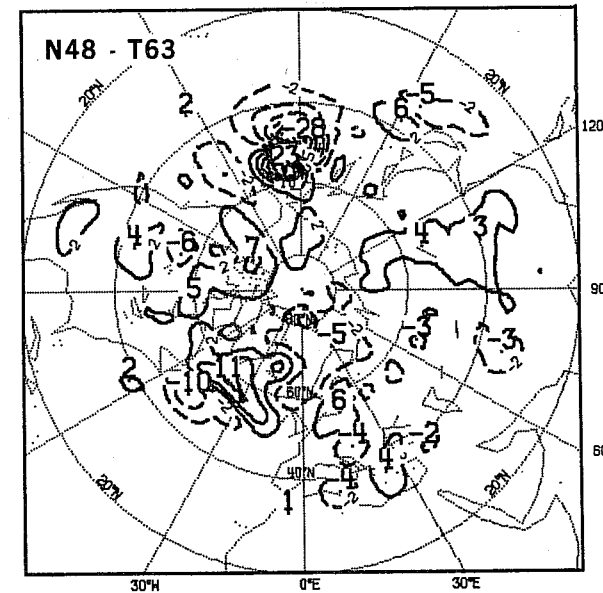
3.3 Winter cases

Period from DEC 6 to FEB 6

The forecasts made during DEC 79 and JAN 80 are, on average, less successful than during NOV 79. Actual FDs as well as differences in model performance are also in general smaller, at least in the first days of the forecasts. The atmosphere itself demonstrates considerable variability throughout the period. Sometimes, as during DEC 6 to DEC 12 period (Fig. 26) the flow is fairly zonal and the jet very strong all around the hemisphere. Other times as during JAN 10 to JAN 15 (Figs. 27 and 28), the flow is more meridional and the jet is fairly weak, except in one or two places. Because the differences are smaller and more variable - the atmospheric flow being fairly unsettled, the FDs also lack time continuity - it is more difficult to analyse and interpret them and to find representative cases. Our choice of cases for this period, DEC 6, JAN 10 and JAN 31, will, in particular, illustrate some of the difficulties and short-comings of single case comparison.
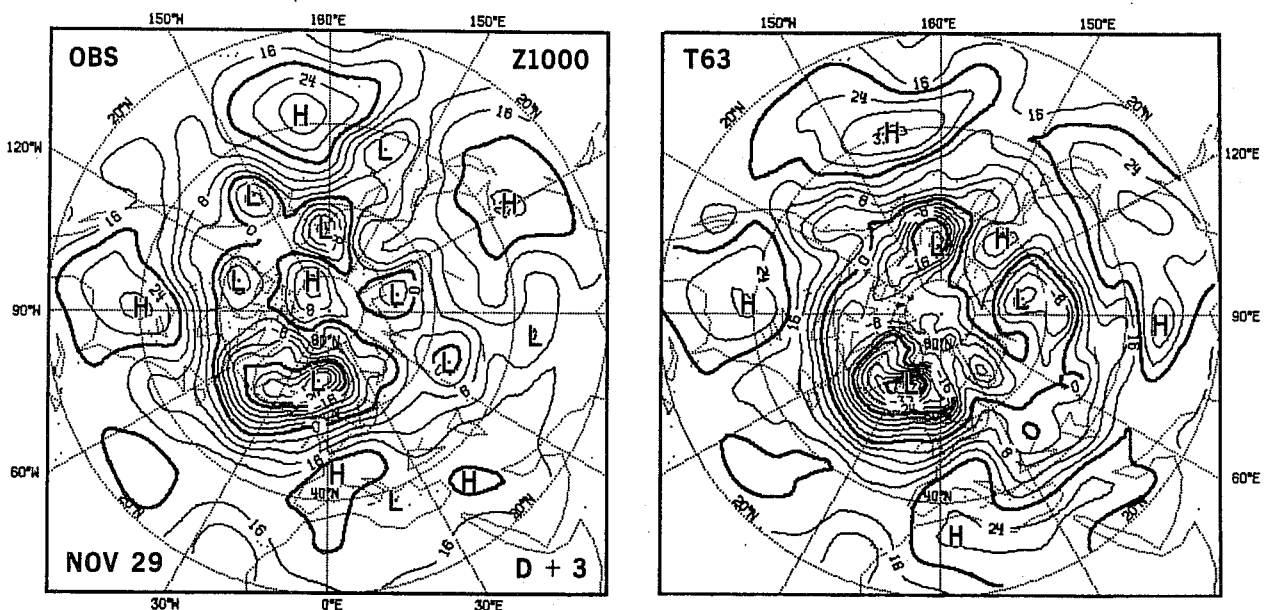
Fig.25 NOV. 29, Z1000: D+3

Case of DEC 6

Fig. 26: Z500, D+4. Also forecast by T40 (spectral model with triangular truncation up to wavenumber 40) and difference between T40 and T63 are shown in the same figure.

This case was chosen mainly beause a T40 integration had also been made. The OBS 500 mb flow is fairly strong and zonal, except over EAS. All forecast circulations are more meridional, especially over WAM and ATL. Comparing first T63 to N48, one notes four main areas of differences. Over PAC, the trough at 170°W and ridge at 145°W are more pronounced in T63 and this is
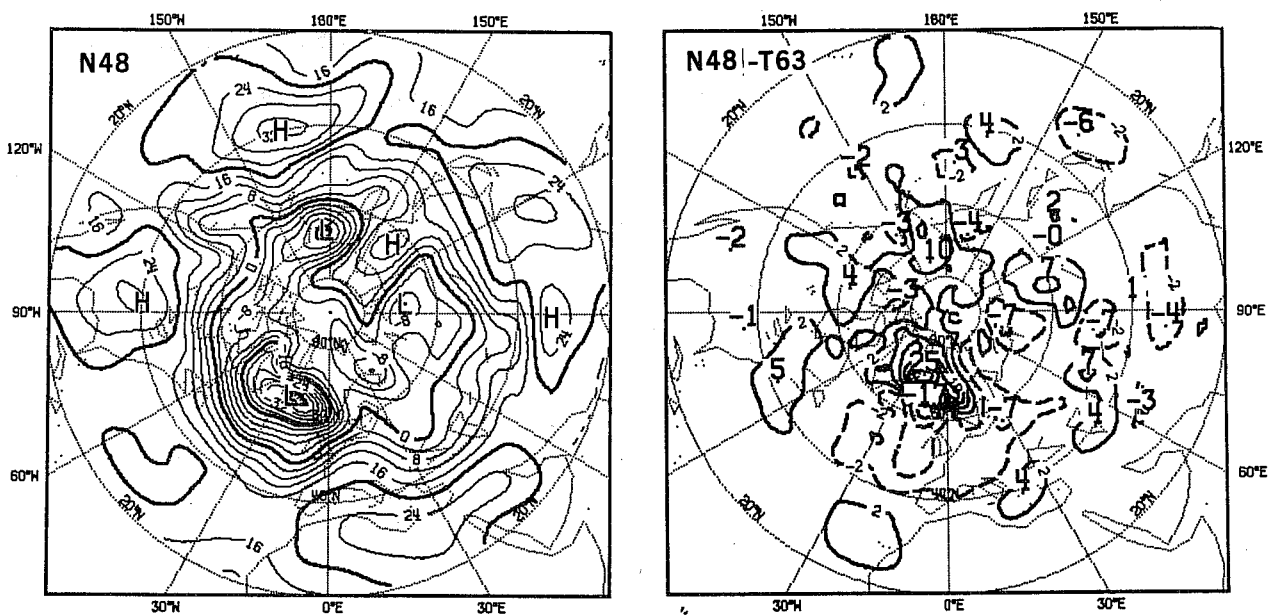
Fig.25 (Cont.)
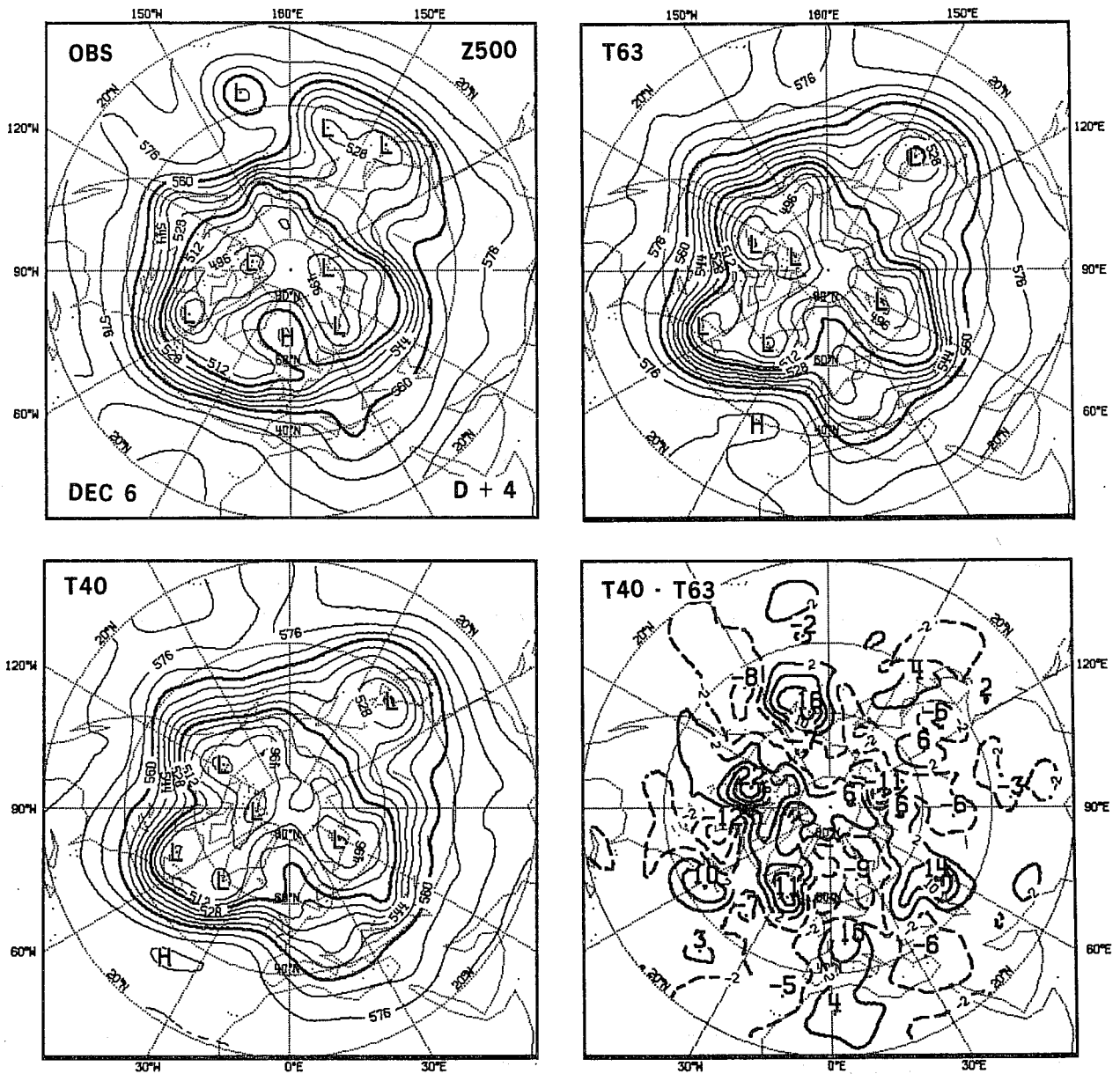
Fig.26 DEC. 6, Z500: D+4   OBS, T63, N48, N48-T63 (top); T40, T40-T63 (bottom).

slightly more correct. The other differences can hardly be interpreted with respect to OBS as the forecasts are too far from it. Over WAM, for example, the trough-ridge system in both forecasts is too intense. The positive difference centre indicates however that part of the double low system is deeper and further east in T63. The low centre at the southern tip of Greenland and the large trough over Russia are also further east in T63.
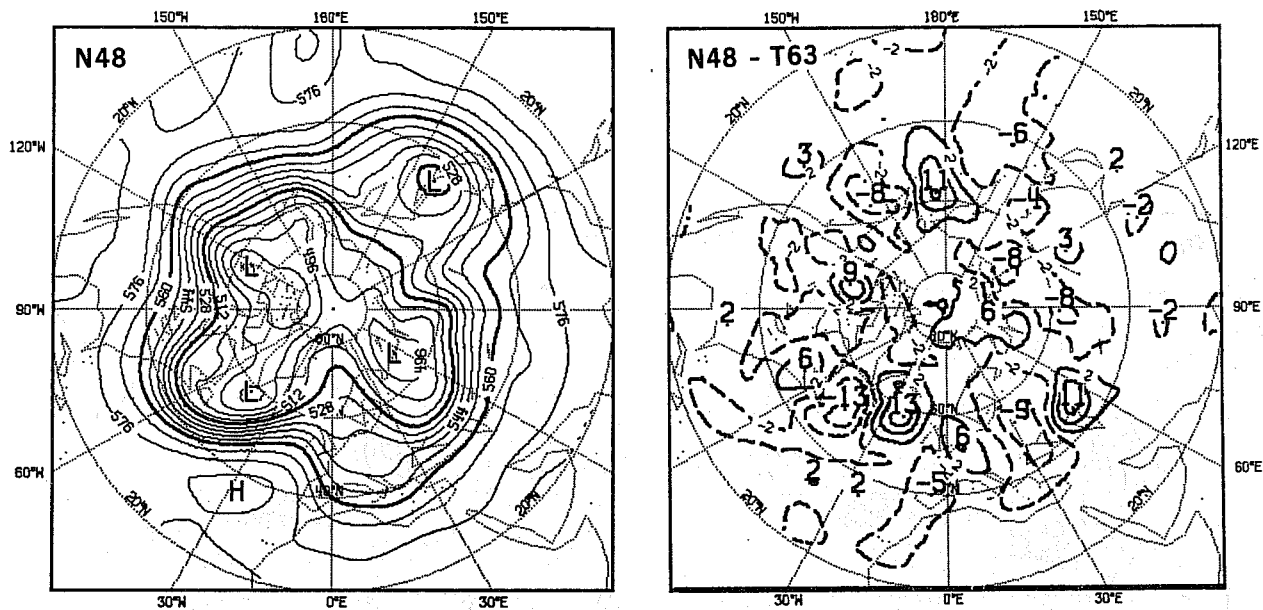
Fig.26 (Cont.)

Second, comparing T63 and T40, one notes that the largest FDs are roughly in the same areas without being truly similar. For example, the intensity of all four above mentioned features is less intense in T40 than in the other forecasts. In three of them (the exception being over PAC), T40 is more correct. It is thus not surprising that objective scores are in its favour (for Z500, on D+4, it scores 70% against 67% AC for both T63 and N48).

It is worth noting that T40-T63 differences are larger than those between N48 and T63, particularly over PAC and WAM. Thus, on this case, the numerical solution, while appearing to converge as resolution is increased (assuming that N48 effective resolution is between T40 and T63:see TR23), does not converge towards the true solution, the deviations from OBS being larger at higher resolution.
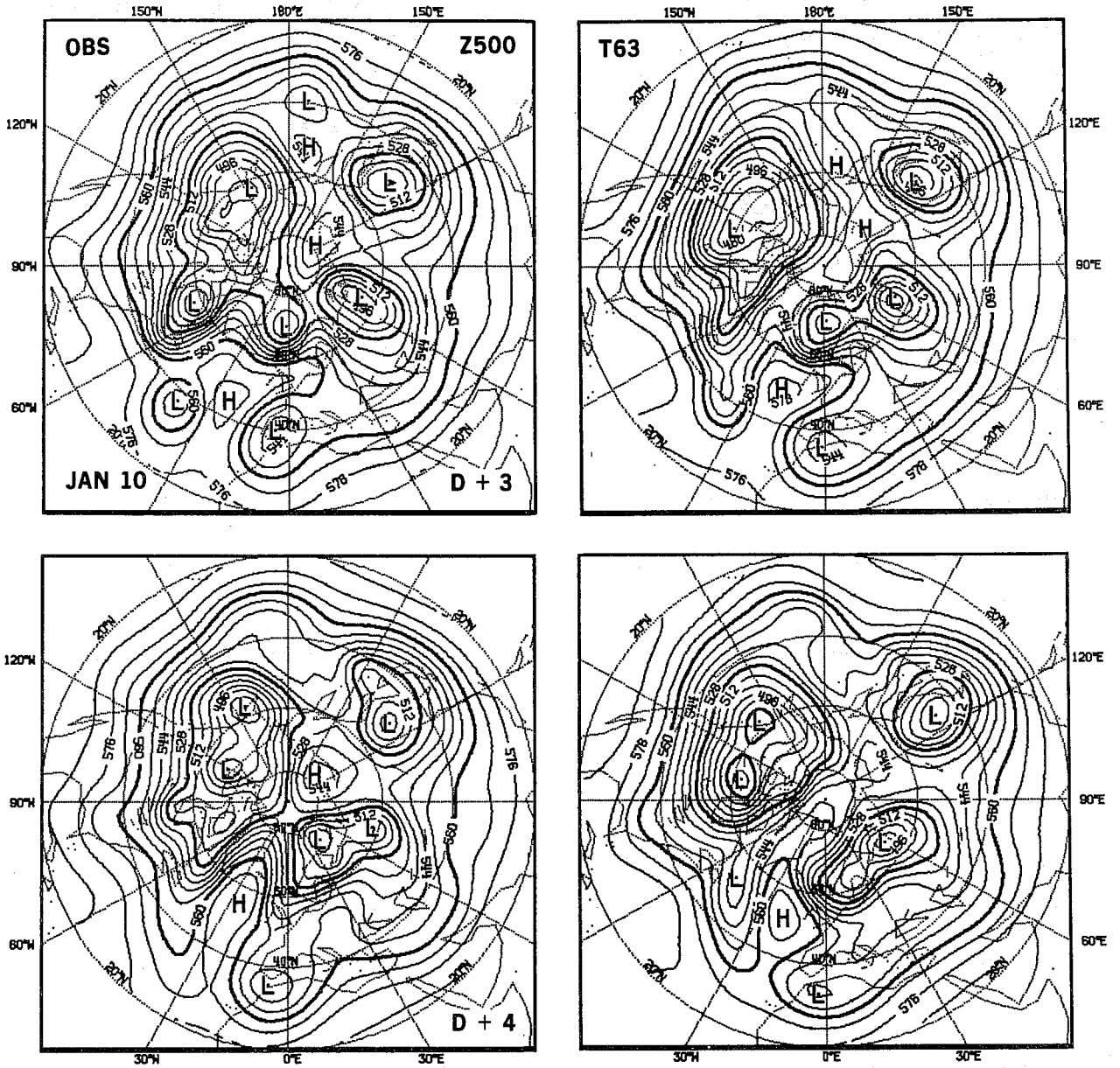
Fig.27   JAN. 10, Z500: D+3, D+4
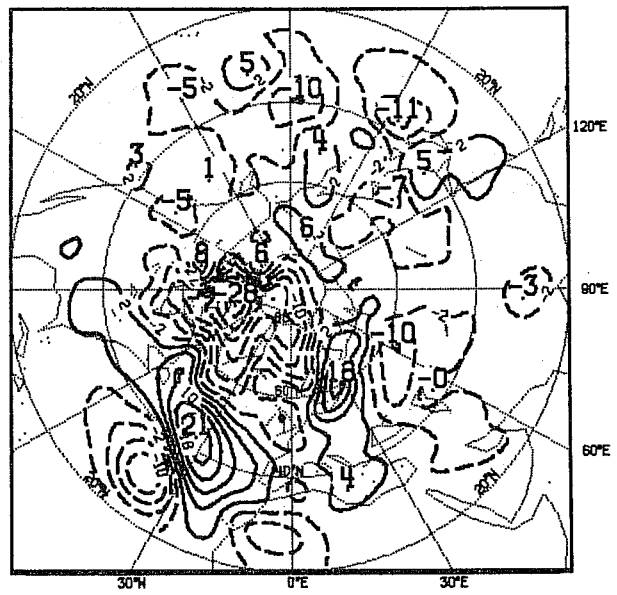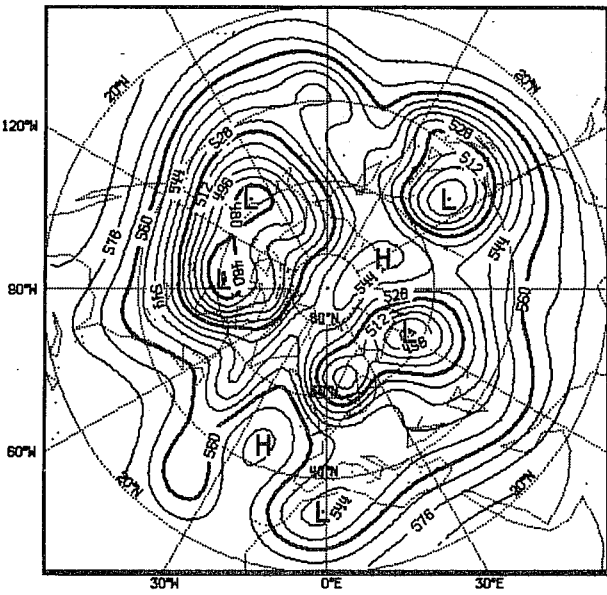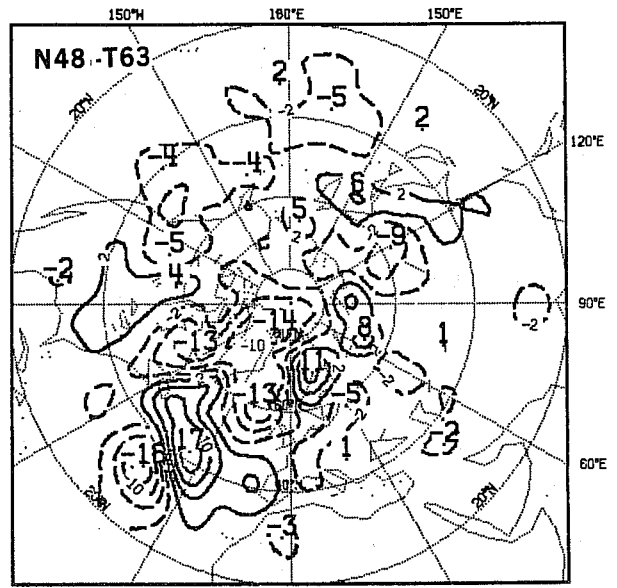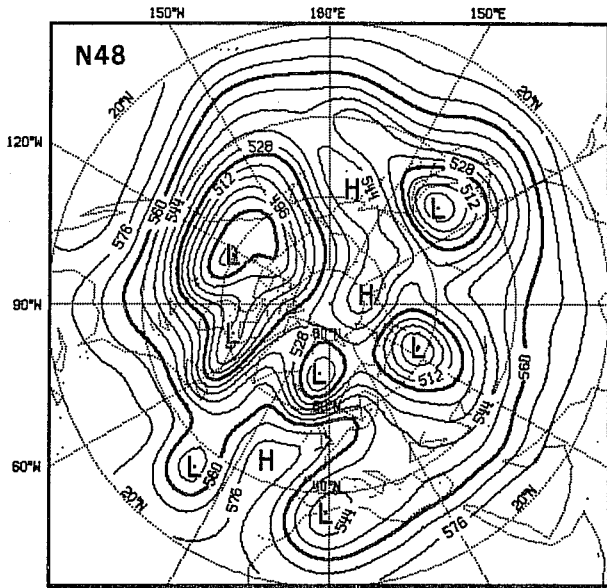
Fig.27   (Cont.)
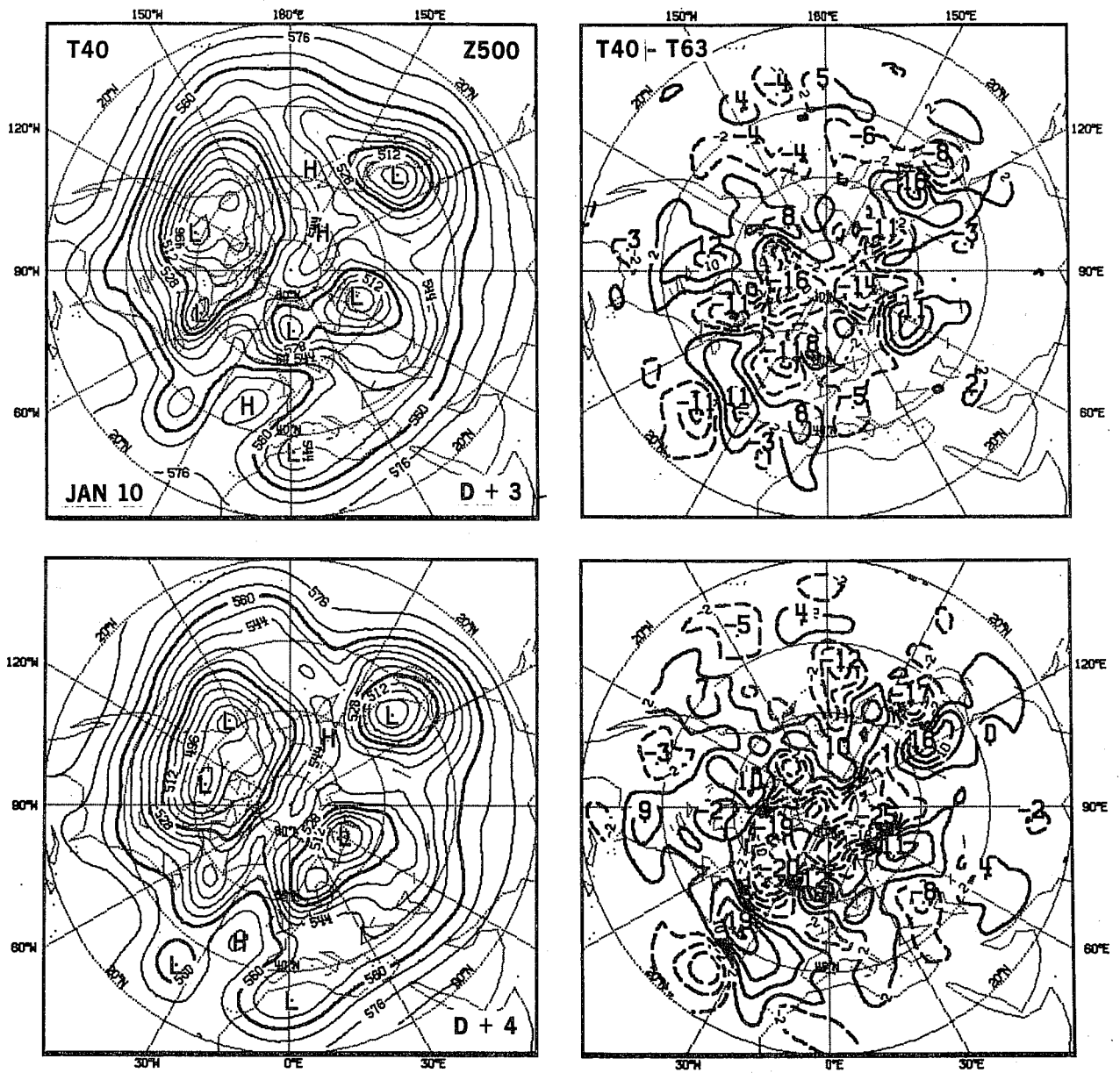
Fig.28 JAN. 10, Z500: D+3, D+4 (T40, T40-T63)

Case of JAN 10

Fig. 27: Z500, D+3, D+4

Fig. 28: Z500, forecast by T40 and T40-T63 differences, D+3, D+4

Fig. 29: Z1000, D+4

Fig. 30: Z1000, forecast by T40 and T40-T63 differences, D+4.

This case was chosen because it showed large FDs (Table 4). At 500 mb (Fig. 27), they are mainly confined to the ATL and EUR sectors, extending into the ARC. On D+3, one sees in OBS a trough-ridge-trough system over north ATL with to the south of it two cut-off lows separated by a blocking high. On D+4, the west cut-off has merged with a trough, the ridge downstream has greatly intensified and the east cut-off has become in phase with the other trough.

On D+3, T63 appears to have begun the merge of the west cut-off with the trough to the north too early and the circulation over Greenland is much too weak. N48 is more successful in these two aspects. However, the position of the low over 0°W is good in T63 while in N48 it lags noticeably behind. On D+4, the heights over and to the west of Greenland are much too high in T63. The ATL trough is too far to the east. Again N48 is much better and although T63 continues to handle slightly better the low system over Northern Europe, N48 overall forecast must be better. This is indeed confirmed by the mean hemispheric scores (Fig. 12 and 13).
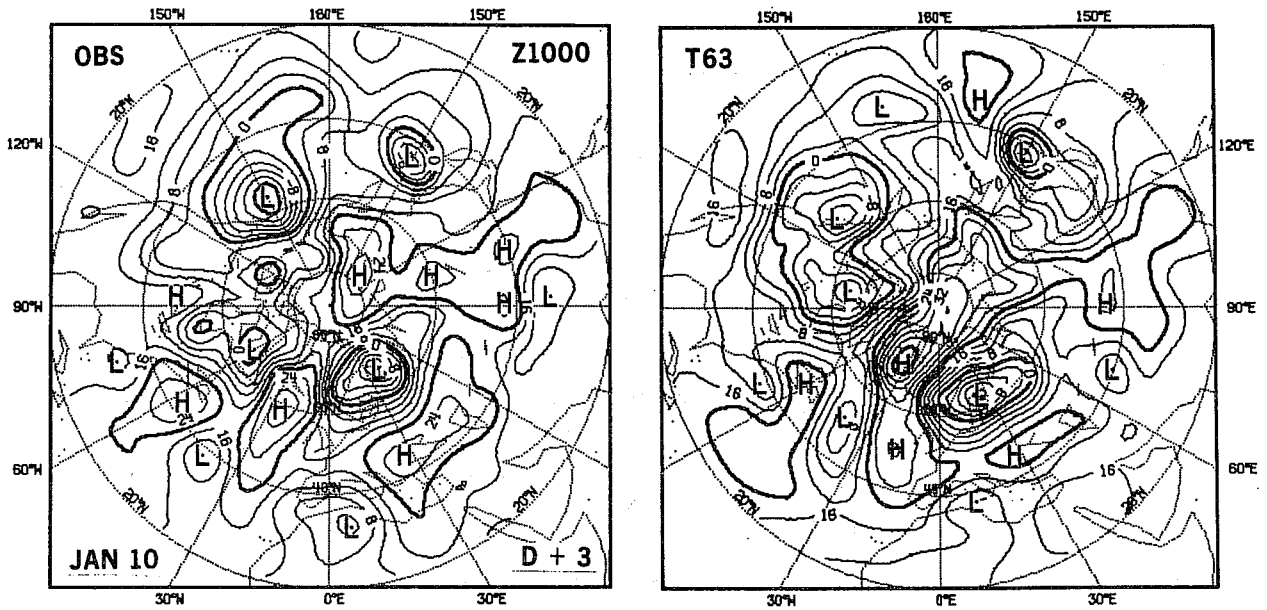
Fig.29   JAN. 10, Z1000: D+3

The phase difference in the low around 0°W (D+3) is typical but the other

differences are more complex and the fact that N48 performs better is fairly

unusual.   In the hope to gain more understanding of this case, we also made a

5-day T40 integration.   Interestingly, the T40 Z500 forecast (Fig. 28), over

the region where N48 performed much better than T63, is very similar to N48

(as seen also from the difference maps) on both D+3 and D+4.   On the other

hand, for the system moving over Northern Europe, it is more similar to T63.

For the rest of the hemisphere however T40-T63 differences are larger than

N48-T63 differences.   Objectively, T40 performs well: it scores in fact

better than T63 and almost as good as N48.

At 1000 mb (Fig. 29), FDs have a similar structure over both ATL and EUR but

their magnitude is smaller than at 500 mb.   One can note in particular the

much too intense high build-up over Greenland by T63.   T40 again resembles

more N48 in this area (Fig. 30).   Over EAS, however, a new development is

correctly captured by T63 but over-developed and out-of-phase in N48.   T40

offers a third solution.
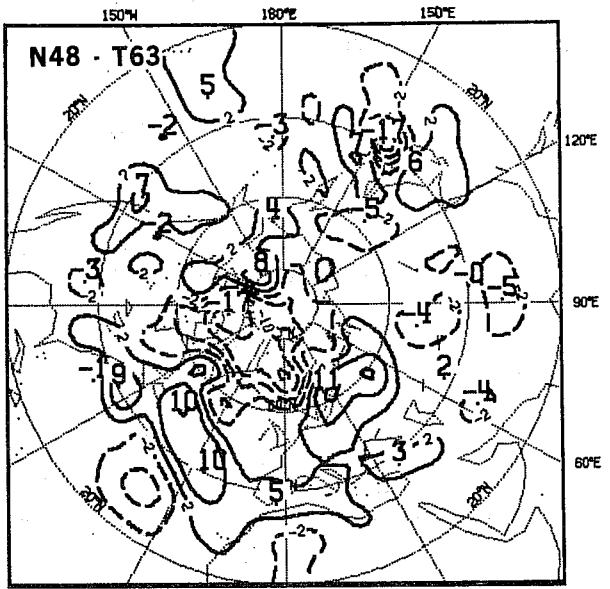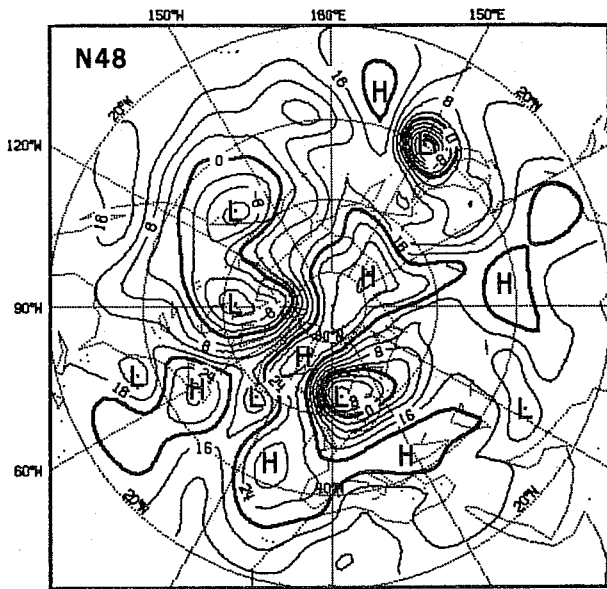
Fig. 29 (Cont.)
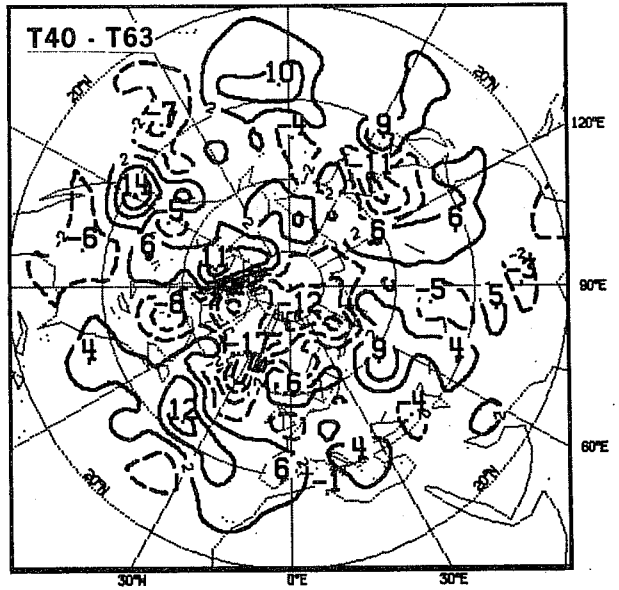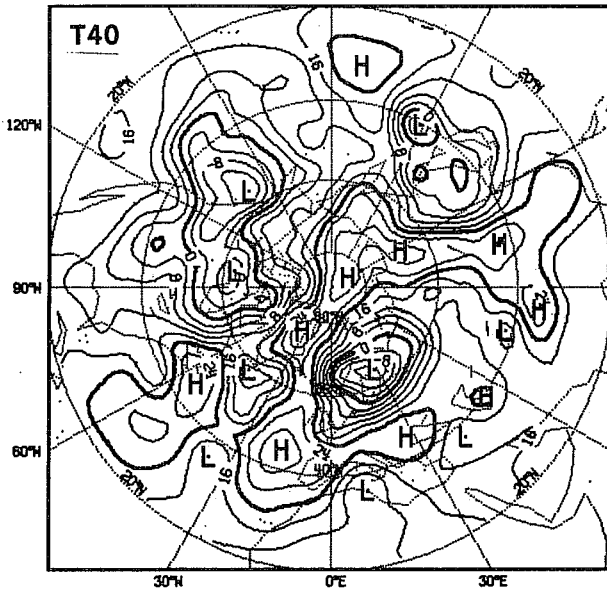
Fig.30   JAN. 10, Z1000: D+4 (T40, T40-T63)

We shall come back to this case in Section 5 but we would already like to emphasize the fact that **in the areas where T63 and N48 greatly differed, N48 being better than T63**, a coarser resolution spectral model **(T40) produced a forecast remarkably similar to that of N48.**
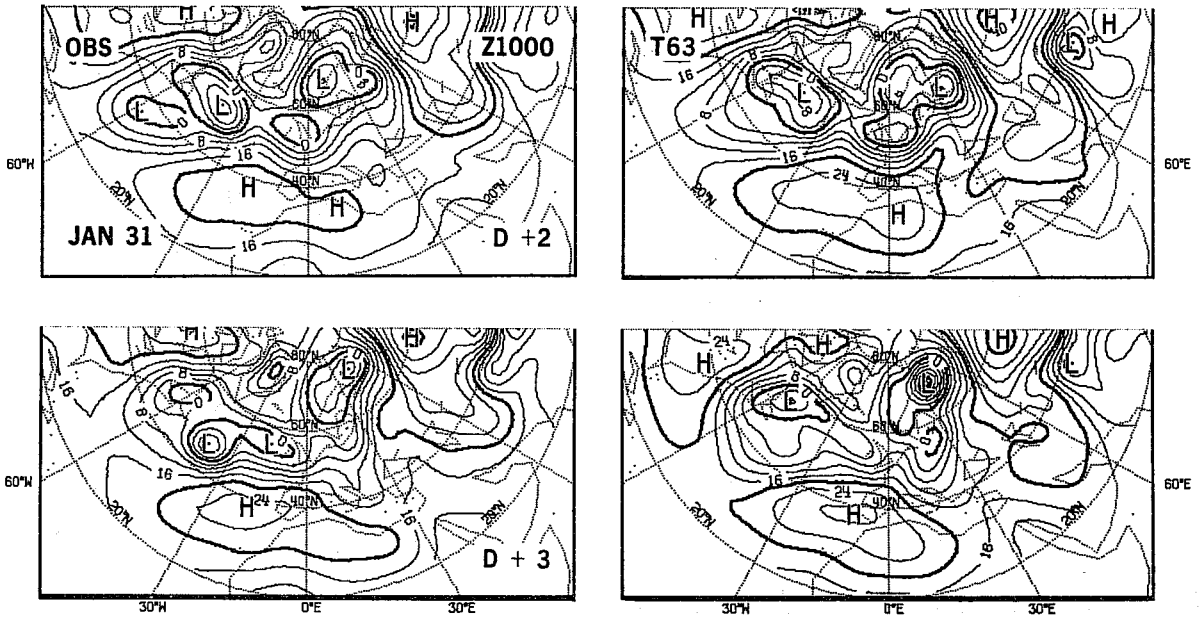
Fig.31 JAN. 31, Z1000: D+2, D+3

Case of JAN 31

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

The performance of the models is here very nearly identical but FDs are

relatively large in magnitude: on D+2, the maximum difference is of 10 dam.

The region of interest is ATL, where T63 better describes the fine detail

structure of the flow. But over Scandinavia, the low in T63 is deeper than

in N48 and is in fact much too deep. Elsewhere, the differences are not very

relevant from a synoptic point of view. Thus, although the hemispheric mean

scores are the same in both models the meteorologically significant

differences are in favour of T63. This case also brings further evidence

that overdevelopment is more of a problem in T63.
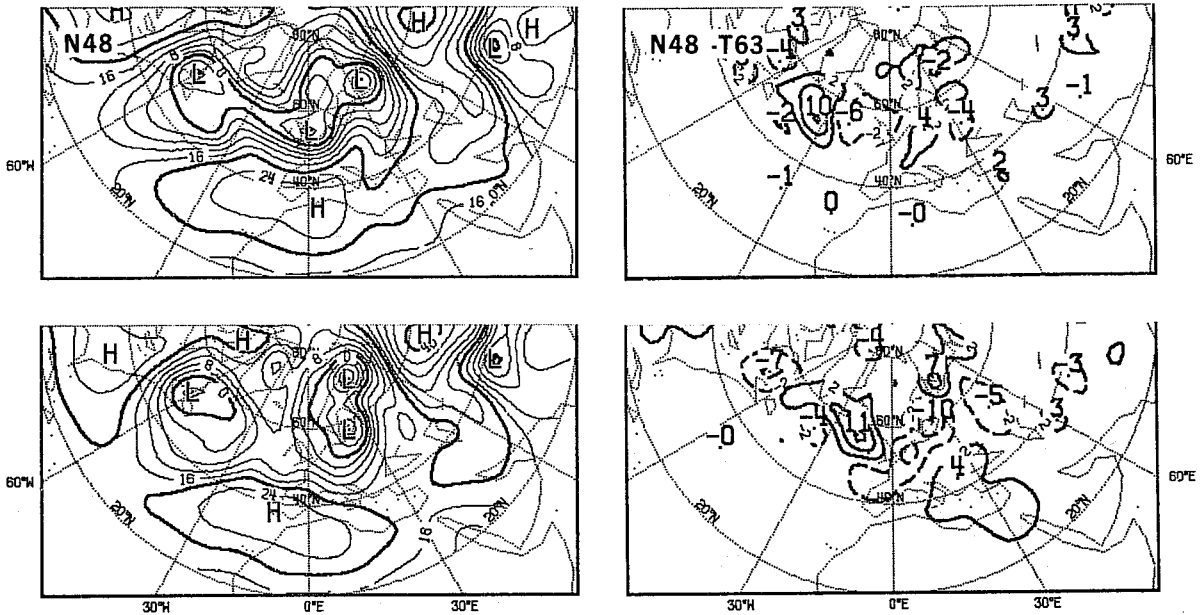
Fig. 31 (Cont.)

Period from FEB 7 to MAR 5

For this period, we would like to emphasize one aspect only of the

atmospheric circulation: namely the presence of a stationary long-wave trough

over west ATL and on which are superimposed travelling short-waves. At 1000

mb these are accompanied by lows forming near Florida, developing rapidly and

moving northeastward along the North American east coast and into ATL. Figs.

32 to 35 only display the lower left quarter of the 1000 mb maps for D+2 of

some, and D+3 and D+4 of all 4 FEB forecasts. As may be verified from Table

4, the forecast differences in this region are in general the largest for the

hemisphere. The difference patterns observed are symptomatic of the
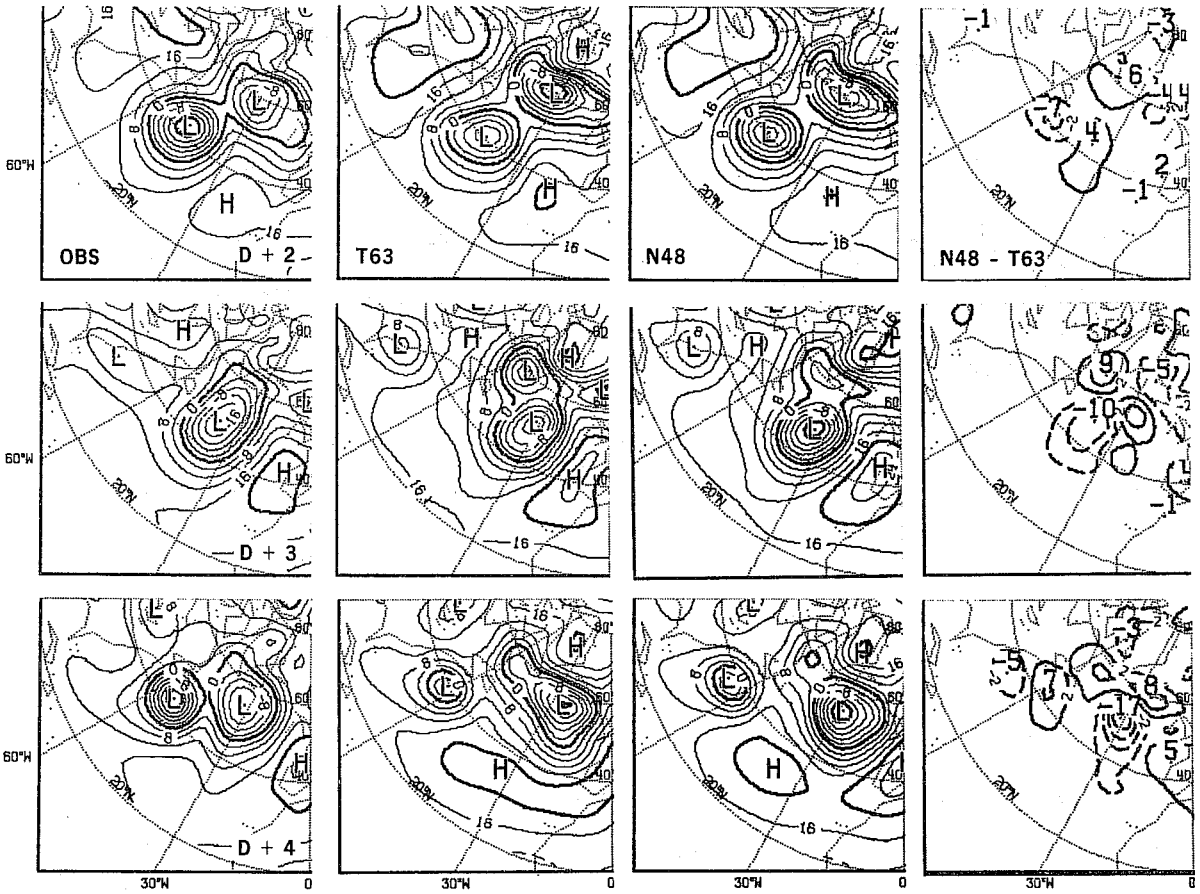
Fig.32 FEB. 7, Z1000: D+2, D+3, D+4
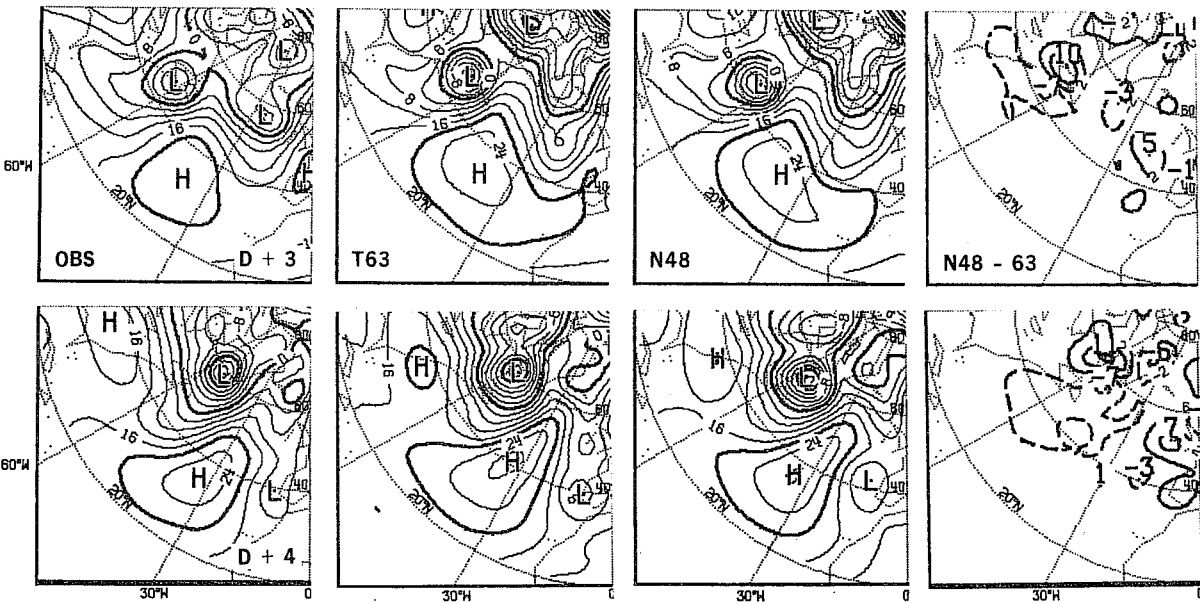
Fig.33 FEB. 14, Z1000: D+3, D+4
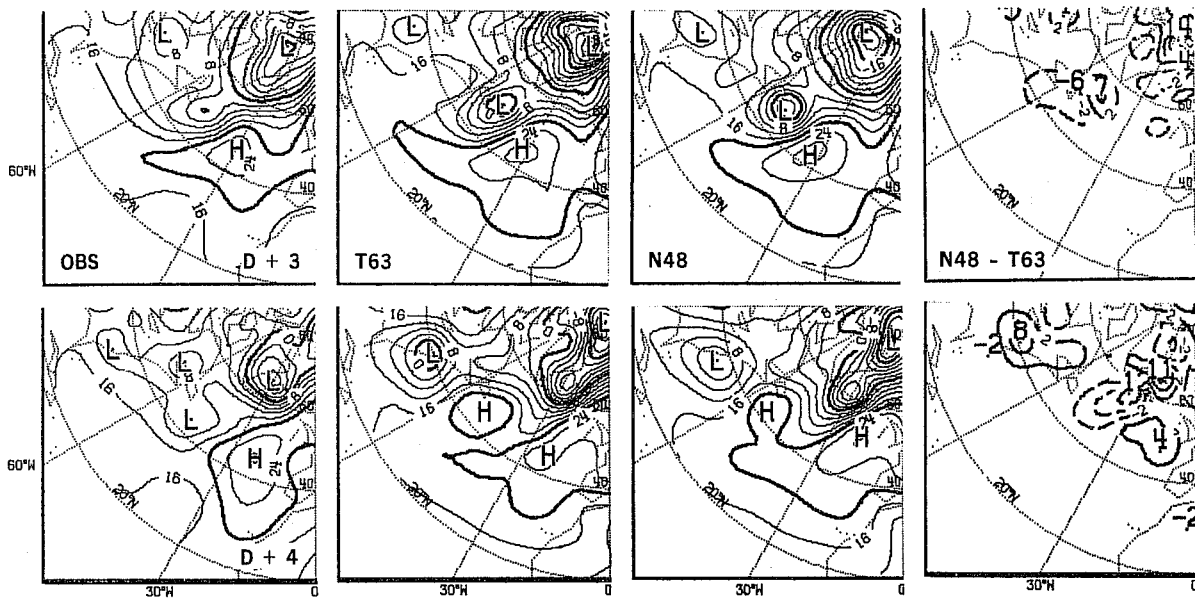
Fig.34   FEB. 21, Z1000: D+3, D+4
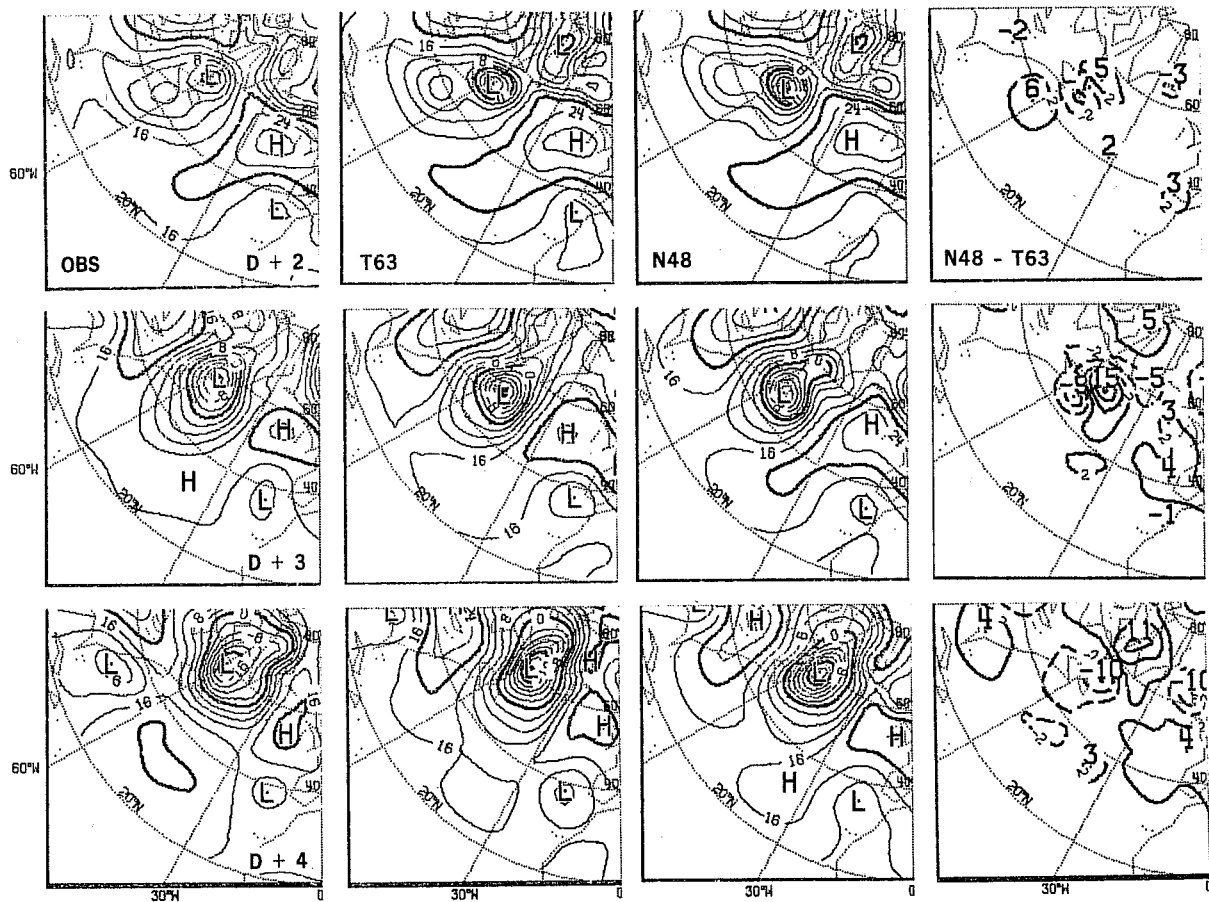
Fig.35   FEB. 27, Z1000: D+2, D+3, D+4

73

phase-like differences already seen in many of the previously described cases. What is more characteristic of this period is the apparent regularity with which situations leading to such differences recurred. It must however be admitted that the impact of these differences on forecast skill cannot have been large since both models score similarly on average during FEB.

Case of FEB 7

Fig. 32: Z1000, D+2, D+3, D+4.

A system of two lows is seen to occupy the OBS map on D+2. Both lows are seen to have travelled eastward on D+3. The leading low has nearly moved off the map and a new low has formed in the wake of the second, near the coast. On D+4, the system of lows has continued to move northeastward and the newly formed low on D+3 has been quickly developing.

The model forecasts differ from OBS in many respects. They nevertheless succeed in bringing out a new low on D+3. Between D+3 and D+4, they move the low correctly northeastward, although too slowly. They predict its development, with underestimation. Differences between the forecasts show in particular a phase speed difference between the lows. Always T63 lows are ahead of N48 lows in their direction of motion.

Case of FEB 14

Fig. 33: Z1000, D+3, D+4.

Case of FEB 21

Fig. 34: Z1000, D+3, D+4.

Case of FEB 27

Fig. 35: Z1000, D+2, D+3,D+4.

These figures are presented to further emphasize the quasi systematic development of similar difference patterns during this period.

3.4 Spring cases

Period from MAR 6 to APR 2

The circulation pattern remains of the wintry type during this period with major troughs in the west ATL and west PAC sectors. The jet is strong over a large part of the hemisphere except over WAS where a blocking high (Figs. 36 and 37) separates the flow into weak southern and northern jets. It slowly moves over north EUR. Models perform very well there. FDs do not grow to become very large but they are well structured. At the end of the period the jet weakens except over PAC where the largest FDs can then be observed

Case of MAR 6

Fig. 36: Z500, D+3, D+4

Fig. 37: Z1000, D+3, D+4

Along the tight stretches of the jet, models develop well organised FDs with a phase-like structure. At 500 mb, on D+3, the huge low over ARC is east of OBS in T63 and slightly west of it in N48. A double trough extending from it over north ATL is more pronounced and tilts more southward in the models. Differences can largely be explained by differences in phase speed. On D+4, these differences have further increased in the same direction. Similar phase-like differences have been developing over EAS and PAC in relation to the two troughs there.
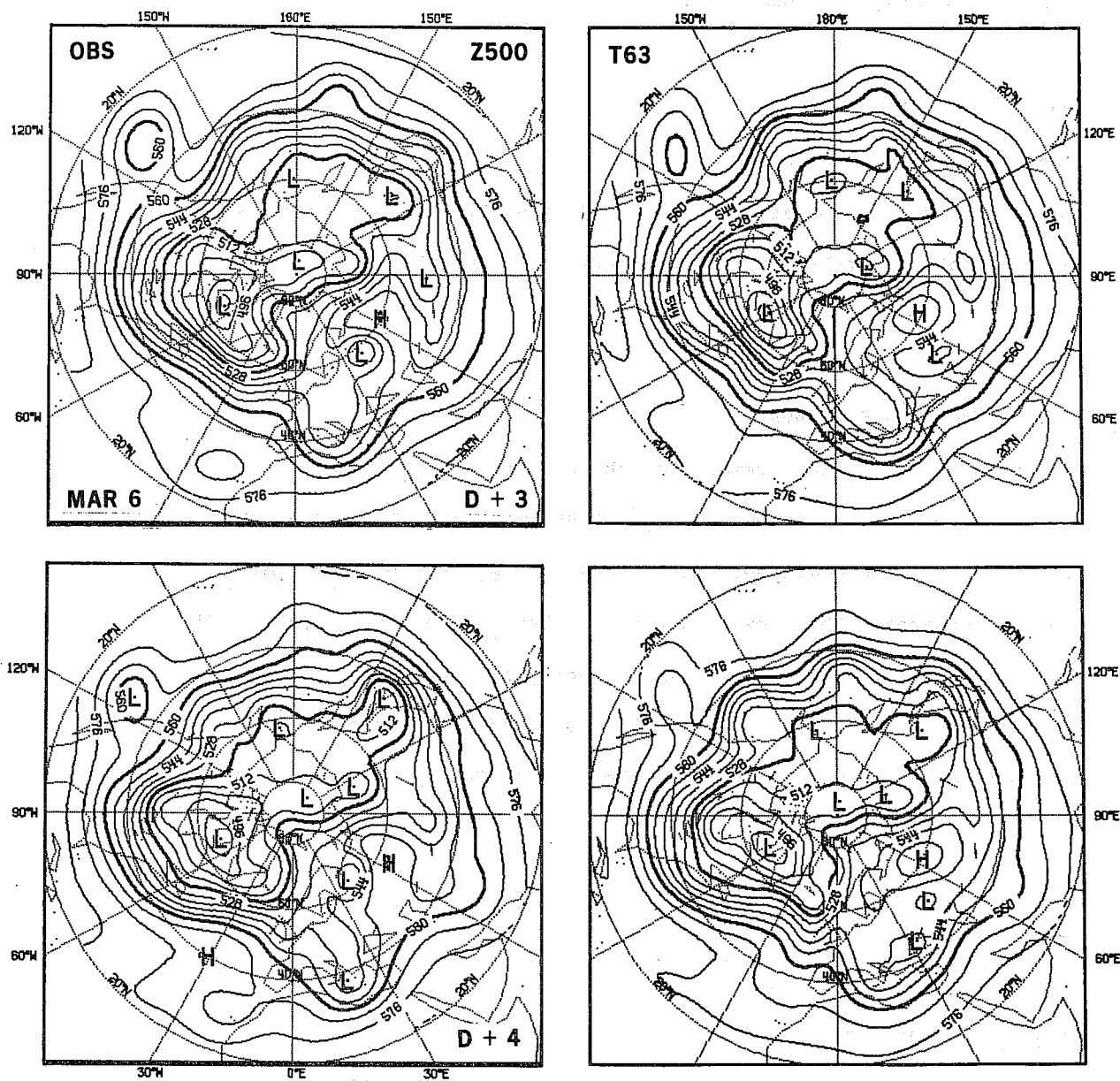
Fig.36   MAR. 6, Z500: D+3, D+4
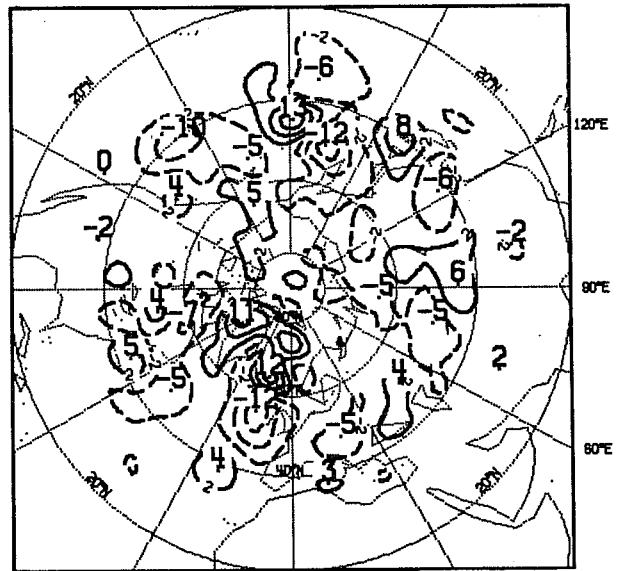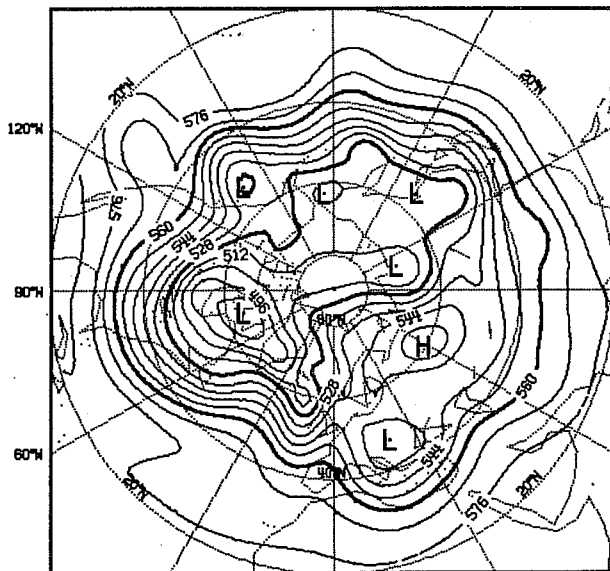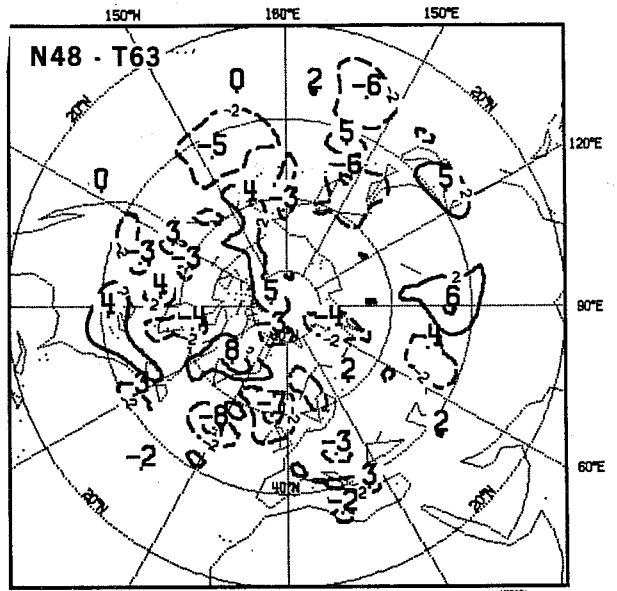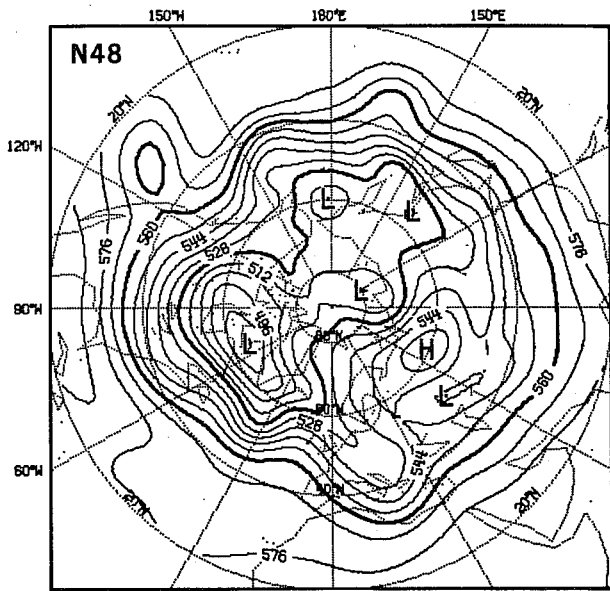
Fig.36   (Cont.)
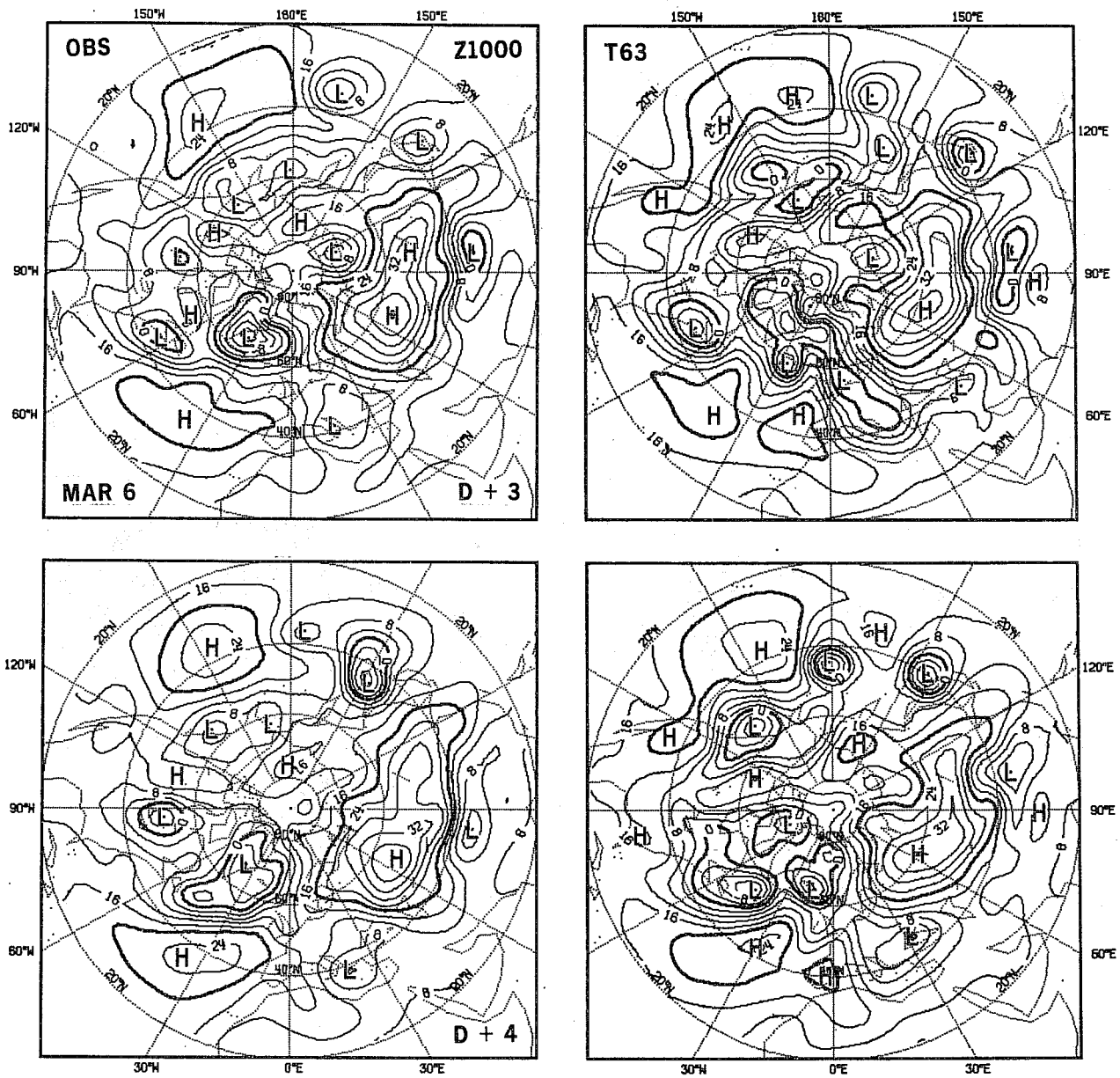
Fig.37　MAR. 6, Z1000: D+3, D+4
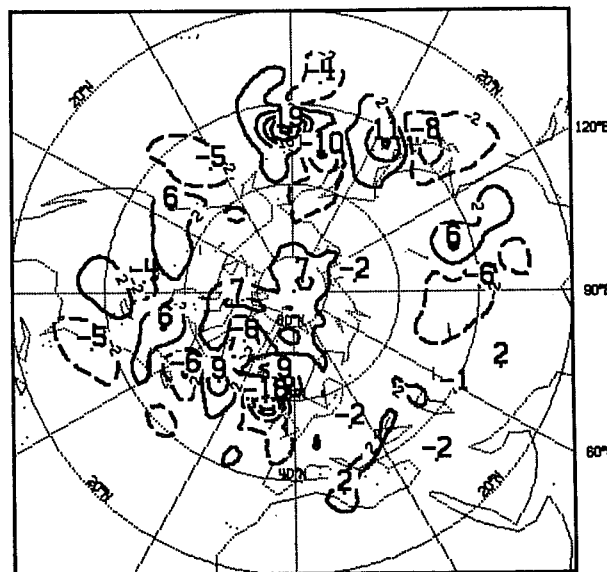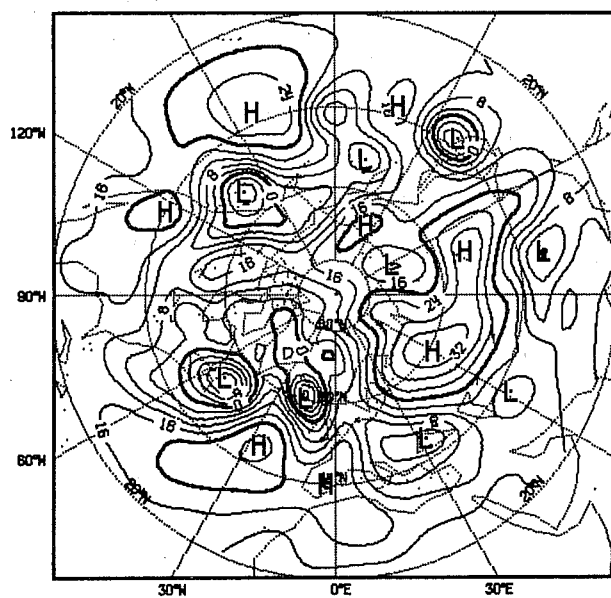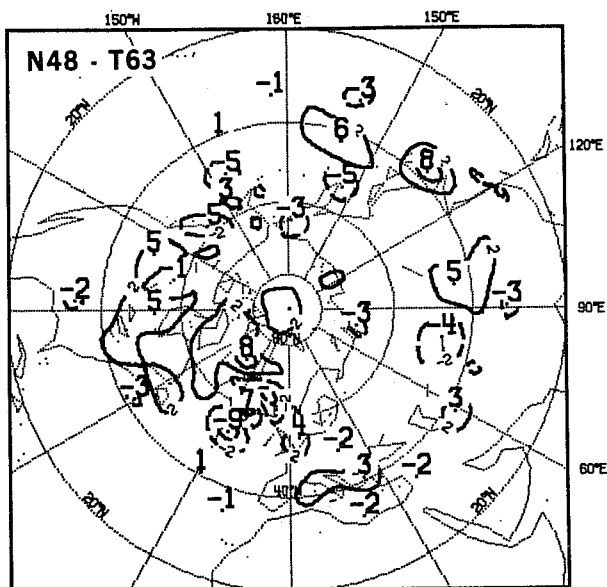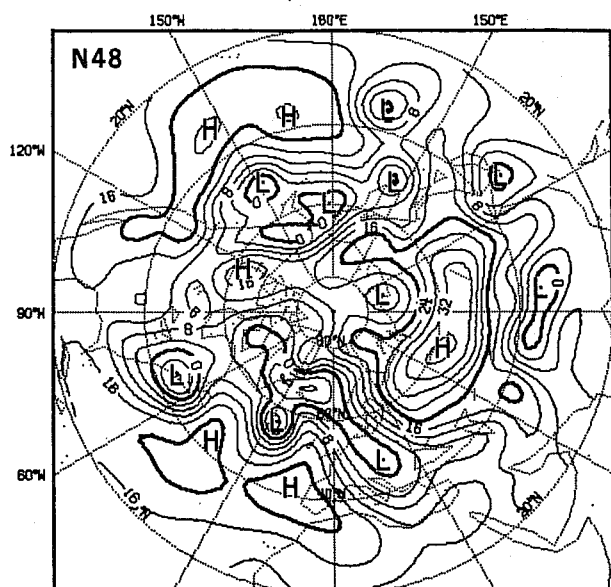
Fig.37 (Cont.)
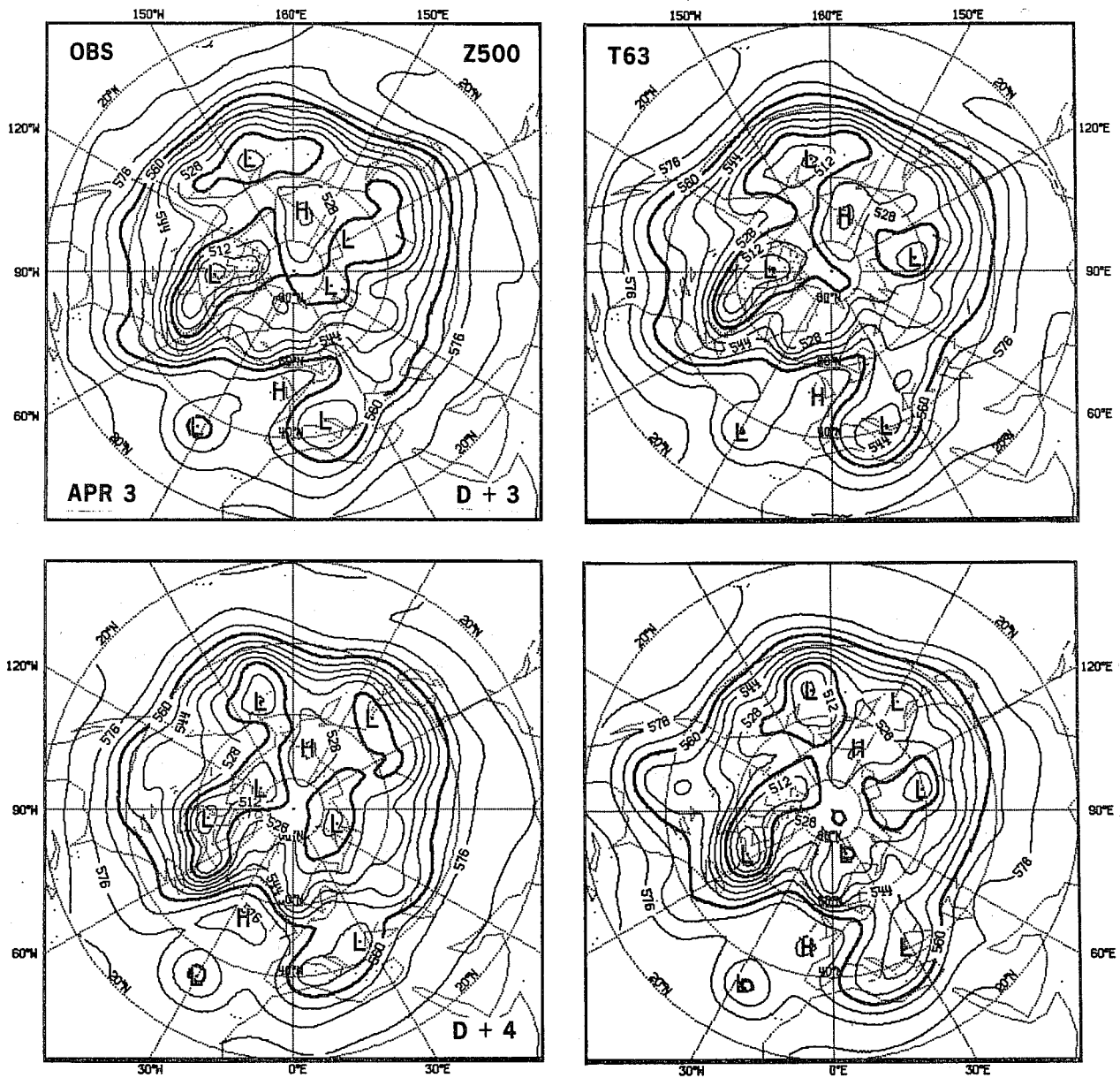
Fig.38 APR. 3, Z500: D+3, D+4
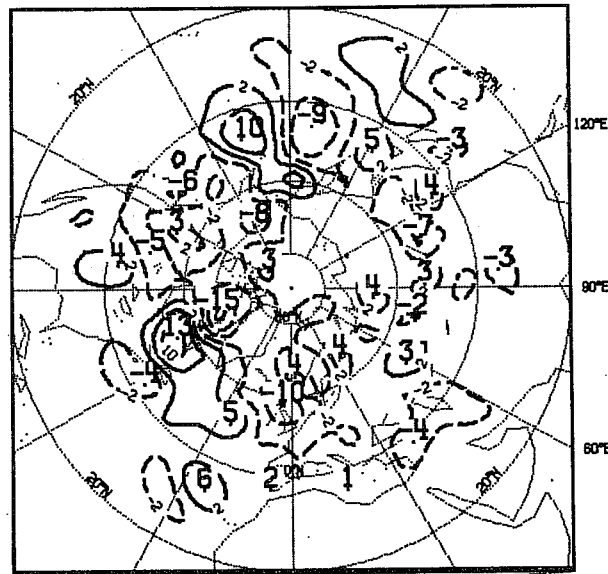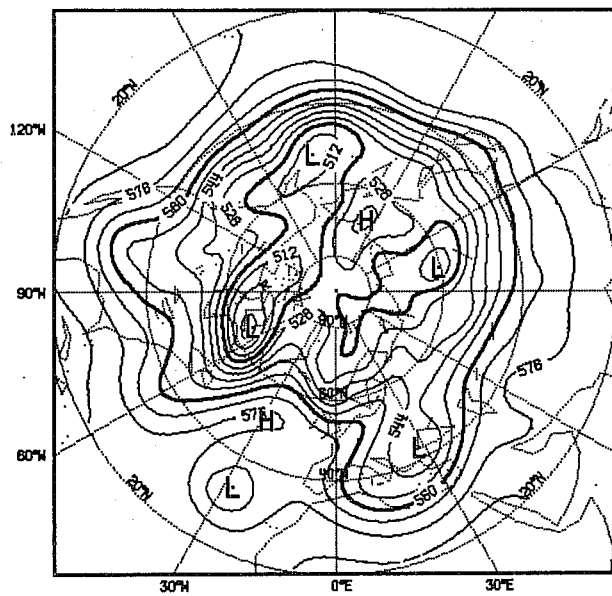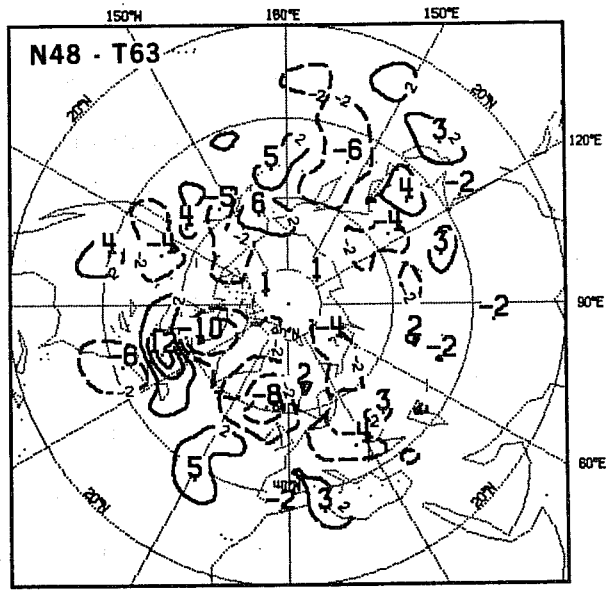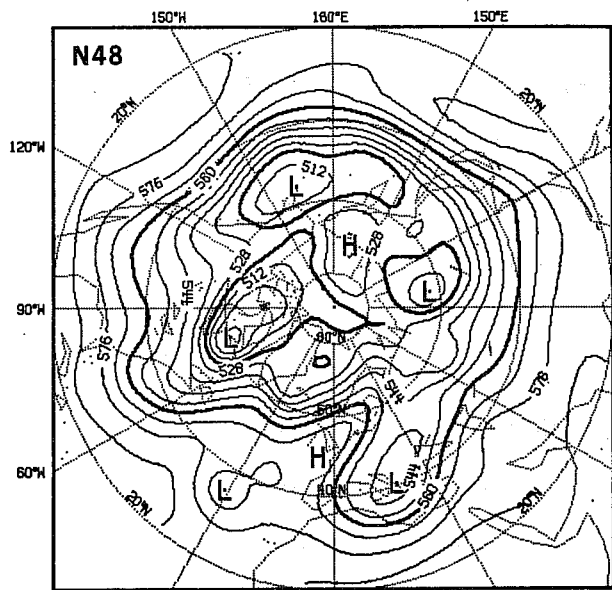
Fig.38   (Cont.)

At 1000 mb, the difference patterns over the same regions are well correlated

with the 500 mb patterns, in relation to the position of the lows in these

regions. Note, however, that differences in amplitude (larger than at 500

mb) are also noticeable, particularly in relation to the development of the

leading lows in both regions, which the models have differently emphasized.

This is apparent on D+4 when T63 has a more intense system over EAS, and N48

over ATL. More details of this case will be given in Section 5.2 along with

an additional comparison with a T53 spectral forecast.

Case of MAR 20

No maps will be shown, but we feel worth mentioning the FDs observed on D+3,

D+4 at both 500 and 1000 mb, aligned in the direction of the 500 mb jet and

related to three troughs. Again T63 puts them slightly ahead of N48, but

both are lagging behind OBS.

Period from APR 3 to APR 30

By then, the block has decayed completely. The 500 mb flow has strengthened

over north ATL and became more transient everywhere (Figs. 38 and 39). APR

forecasts are reasonably good with the exception of APR 10 case. T63 is

better at short range on all cases. Largest score differences occur on APR 3

and APR 24.

Case of APR 3

Fig. 38: Z500, D+3,D+4

The features of interest are the trough over PAC and the two short waves over

north ATL. Over PAC, on D+4, T63 more successfully forecasts the amplitude

and position of the trough and upstream ridge. Over north ATL, on both D+3

and D+4 the double trough system is better handled by T63, mostly in terms of

amplitude. In particular, the ridge from Greenland to North Canada is much

too weak in N48.

Case of APR 24

Fig. 39: Z500, D+4, D+5

Again phase differences appear over PAC. But the main FDs are near the coast of WAM, where T63 manages to keep a trough almost correctly in phase with OBS on D+4 and D+5, although the structure of the trough is not satisfactory. N48 has developed the trough excessively and does not bring it over land. Along the east coast of North America, the trough and cut off are slightly more correct in T63 as are the structures of the trough over north ATL and ridge over ARC. Over WAS, T63 is too fast however and N48 forecast is to be preferred.

Period from MAY 1 to JUN 4

At the beginning of MAY, we observe a general weakening of the circulation, but the 500 mb jet remains well defined over PAC. There is a ridge over WAM and a high over north ATL with weak westerlies to the south (Fig. 40). The structure of the flow is not substantially altered during the month. Often though, it is perturbed by large amplitude travelling waves . Although lower, the predictability is good for the time of the year. T63 scores consistently better. Case of MAY 1 is a typical example of the differences in models behaviour during the month.

Case of MAY 1

Fig. 40: Z500, D+4

Differences of phase are apparent over PAC at 500 mb. T63 is perhaps slightly too fast but N48 is clearly too slow. Over ATL, and east North America, the complex system of highs and lows is better simulated by T63 and to some extent also that over WAS (120°E).
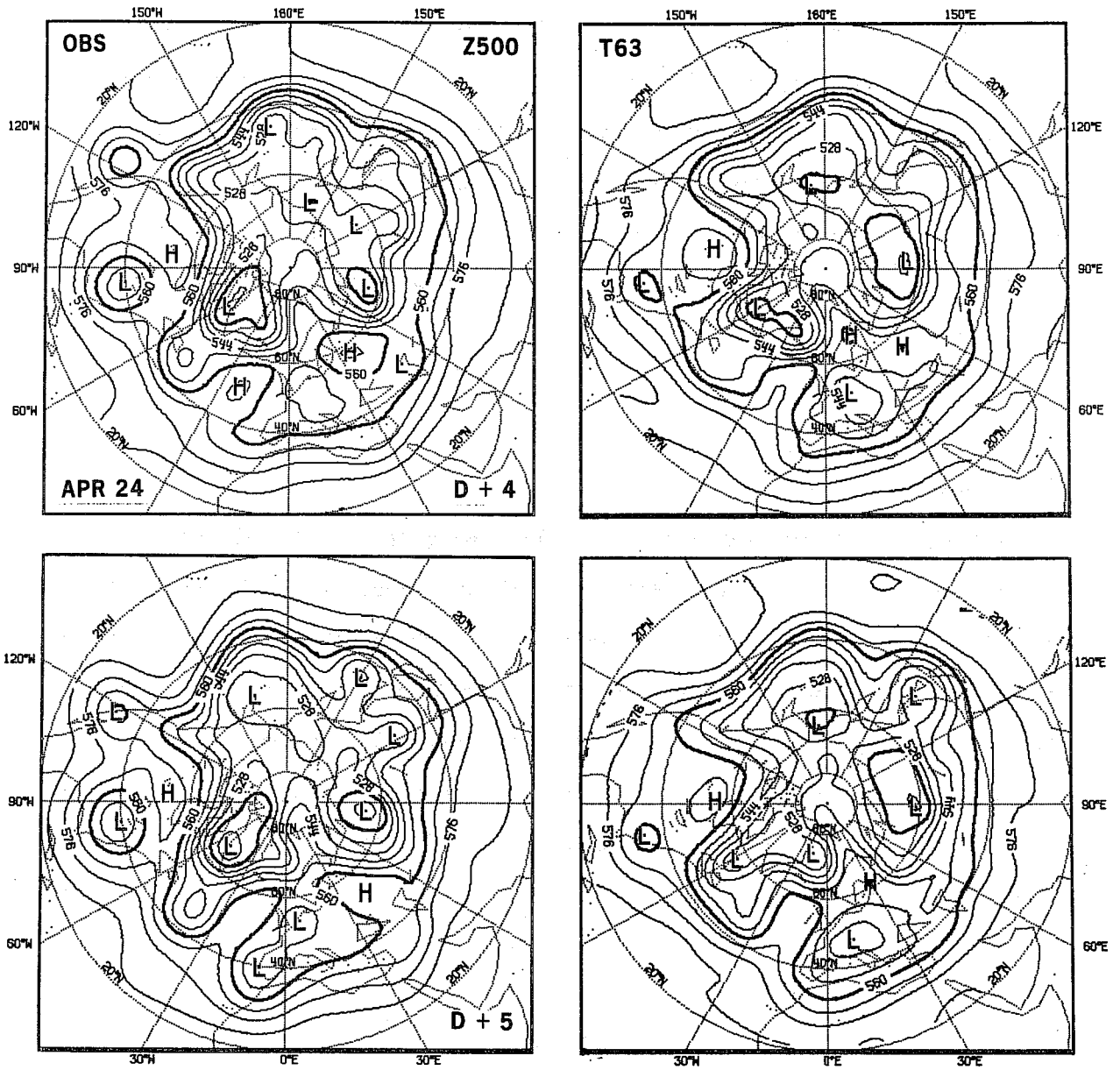
Fig.39   APR. 24, Z500: D+4, D+5

Fig.39 (Cont.)
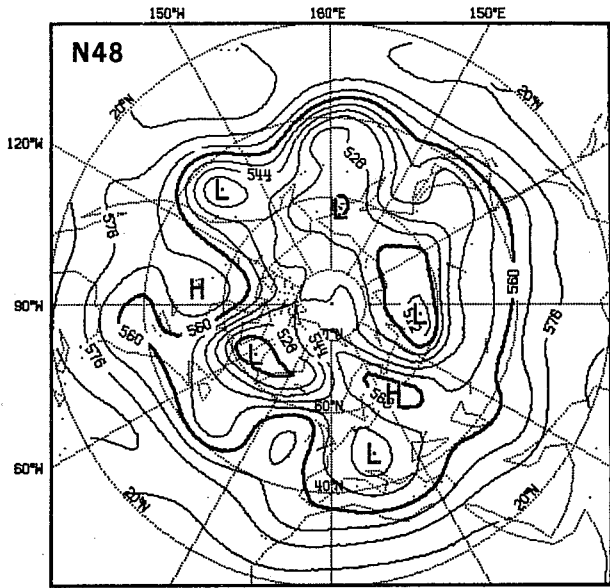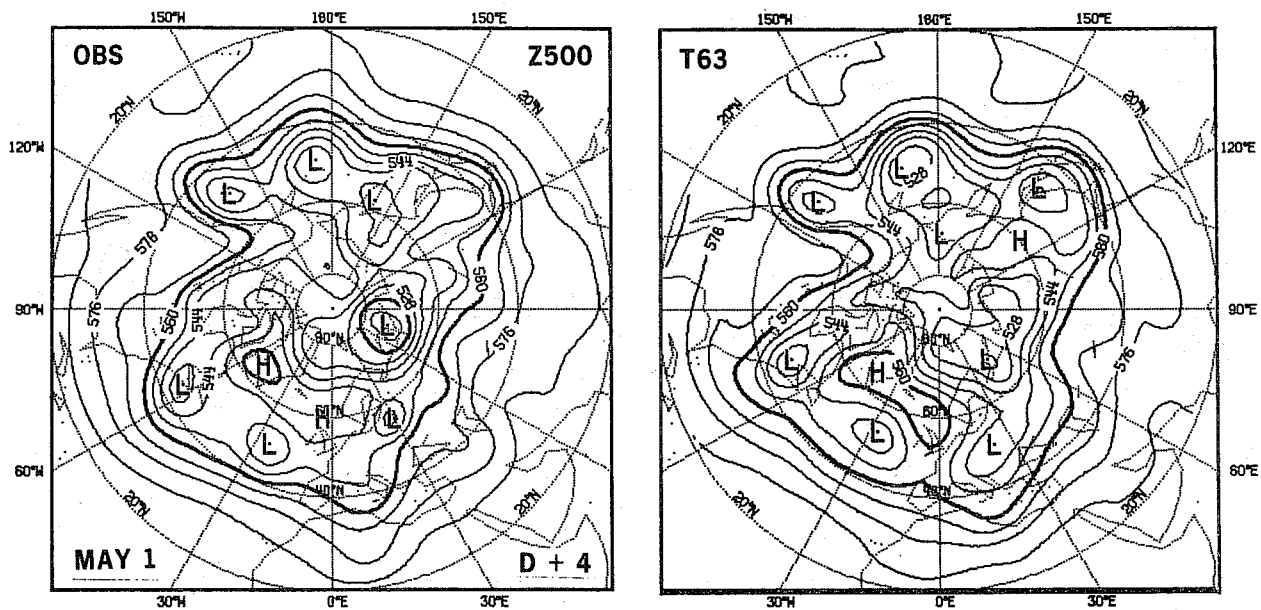
Fig.40   MAY 1, Z500: D+4
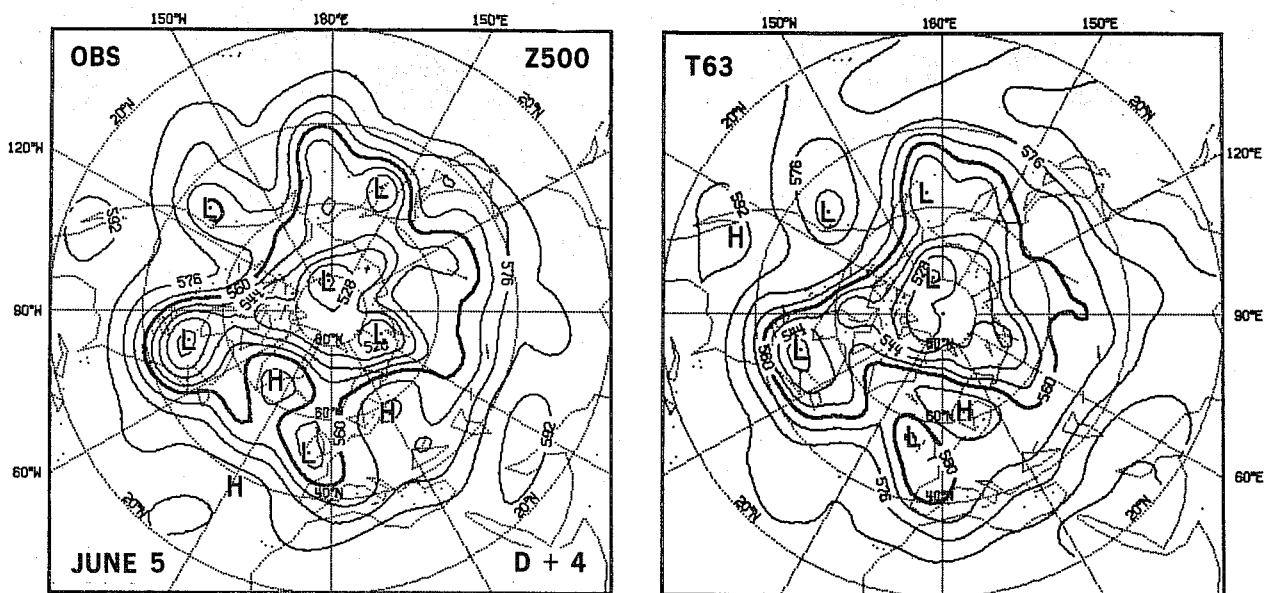
Fig.41   JUN. 5, Z500: D+4
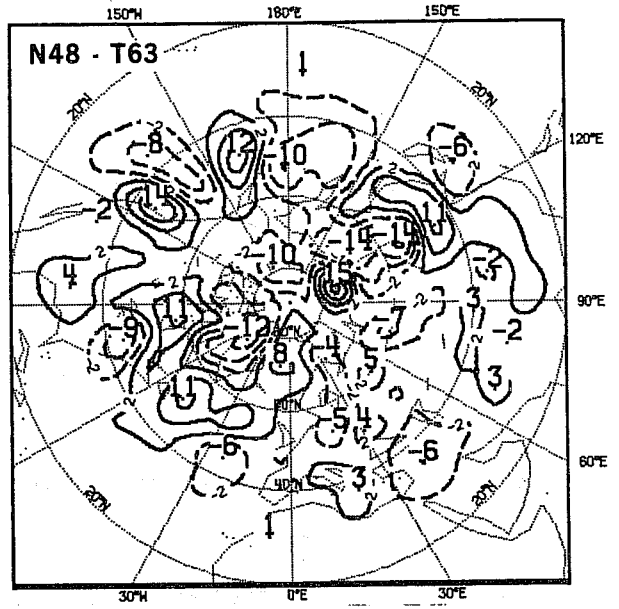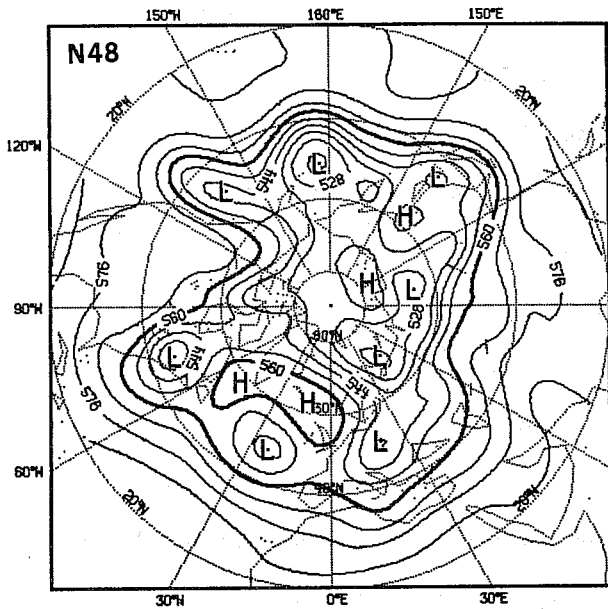
Fig.40 (Cont.)
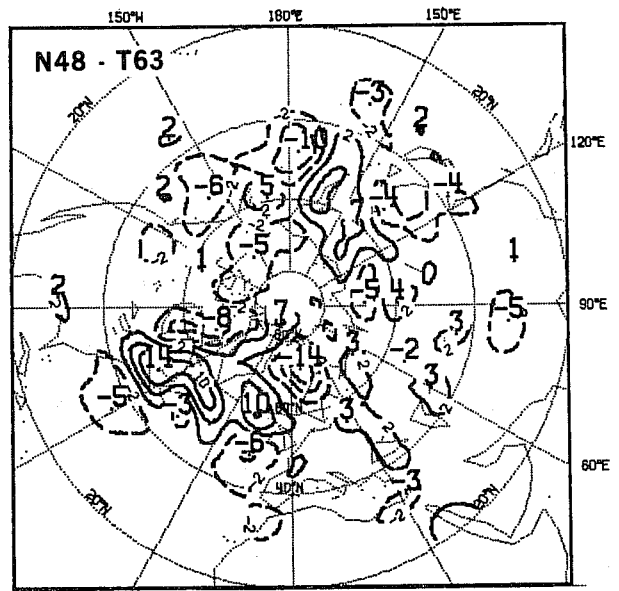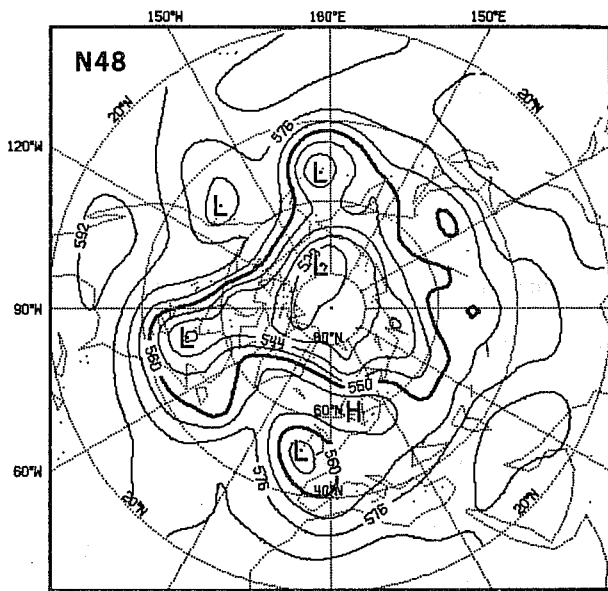
Fig.41 (Cont.)

3.5 Summer cases

Period from JUN 5 to JUL 2

During JUN, the 500 mb flow remains weak. A stronger jet circles the pole at latitude 60°N. A weaker jet runs at about 40°N. Perturbations travel along both jets which sometimes merge locally to form larger perturbations (Fig. 41). At other times, the northern circulation intensifies forming a fairly deep low near the pole. At 1000 mb, over the PAC and ATL oceans and over WAS, the circulation is largely anticyclonic. Over the ARC, it is largely cyclonic. The major cyclone tracks are from southern US to north ATL and from south EAS to north PAC (Fig. 42). Accordingly, largest FDs are found over ARC, EAS and ATL. Although these are much smaller than in WIN for example, they produce relatively large differences in AC scores. Because the circulation is weaker, the anomalies are smaller and the AC become presumably more sensitive. At 1000 mb the scores are very much in favour of T63, as for JUN 5 which shows the largest differences.

Case of JUN 5

Fig. 41: Z500 D+4

Fig. 42: Z1000 D+3, D+4

At 500 mb, on D+4, large FDs are seen in connection with the two main troughs. Over Hudson Bay, T63 better captures the cut-off structure of the low and its position (N48 is lagging behind). Over PAC, T63 is also closer to reality, but this time, by not generating a cut-off low over 180°E and by better simulating another trough over 160°E (N48 has a ridge in its place). There is another large FD near Scandinavia in association with a fast moving spurious trough in N48.

At 1000 mb, the main FDs are similar in structure, but in order to show their evolution as well we display both D+3 and D+4. Over North America, on D+3, the differences are related to the position and shape of a deep depression on

the east coast. The low is elongated in the NS direction and has many minor troughs extending from it. The overall pattern of T63 is more accurate. On D+4, heights are on average lower and the low is further inland in T63.

Over north PAC (180°E), on D+3, OBS has a depression and to the west of it a secondary trough. As the former feature moves northeastward with little change in amplitude, the latter develops to become a well organized low on D+4. T63 indicates all of these features fairly well on D+3 and predicts the correct trend from D+3 to D+4. N48 has missed the secondary trough on D+3, showing instead a broader ridge west of the main low. On D+4, it has incorrectly intensified both the low and ridge.

Further details (D+5) of this case will be given in Section 5.2.

Period from JUL 3 to SEP 3

During JUL, the 500 mb height field shows a weak and broad stream at middle latitudes with however more pronounced gradients from WAM anticlockwise to EUR (Fig. 43a). In AUG, the circulation is characterized by the presence of a deep depression over ARC with troughs extending south (Figs. 45 and 46). At 1000 mb anticyclones still dominate east PAC and south ATL regions, but with gradually diminishing strength. Cyclones are more frequent over ARC, EAS and north ATL and their mean depth increases from JUL to the end of AUG (Figs. 43b and 47).

FDs have their minimum in JUL but grow again in AUG. They are largest on average near the pole but also occur whenever strong gradients and deep depressions are found. Differences in AC scores remain relatively large, in favour of T63.
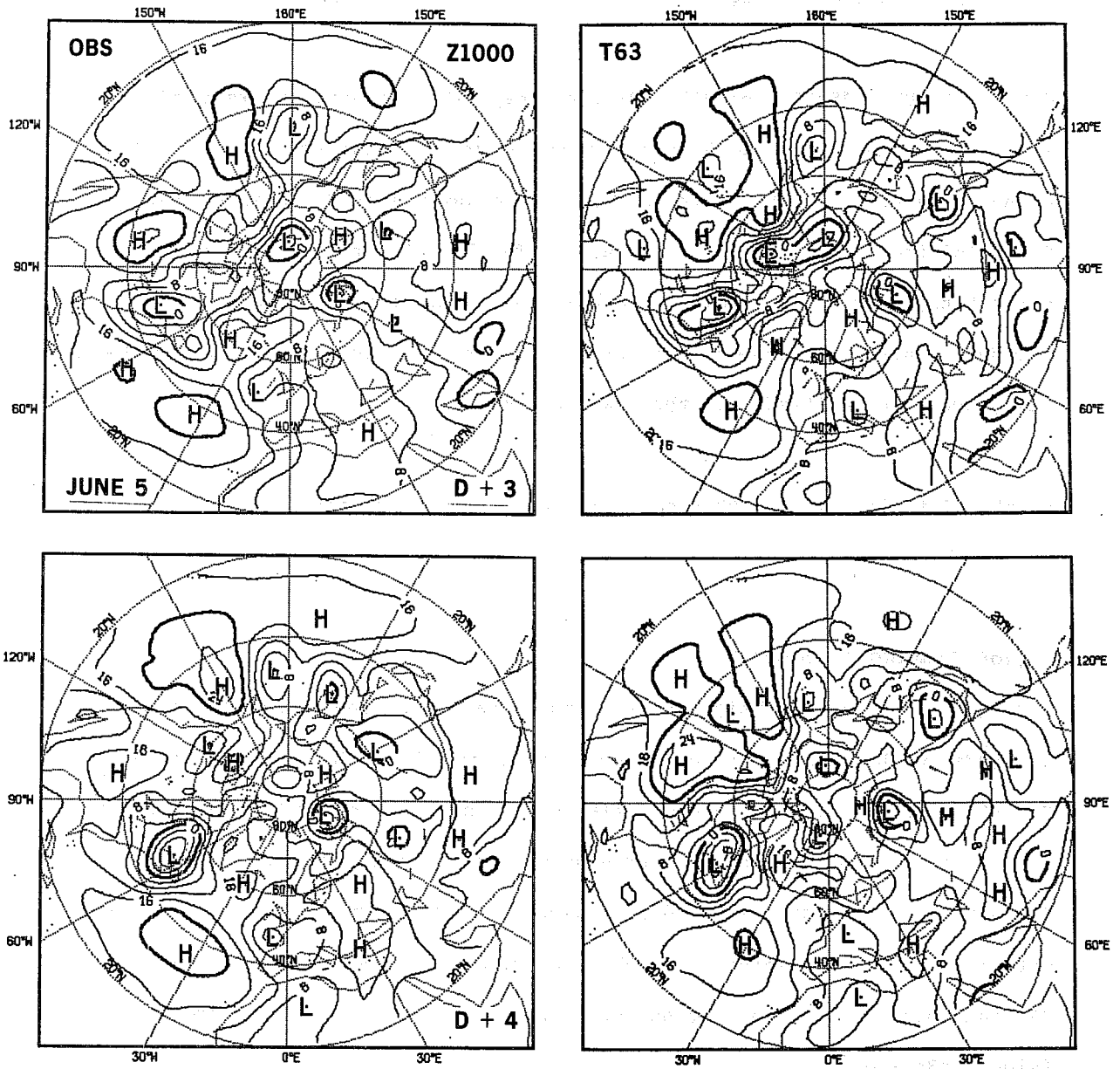
Fig.42   JUN. 5, Z1000: D+3, D+4
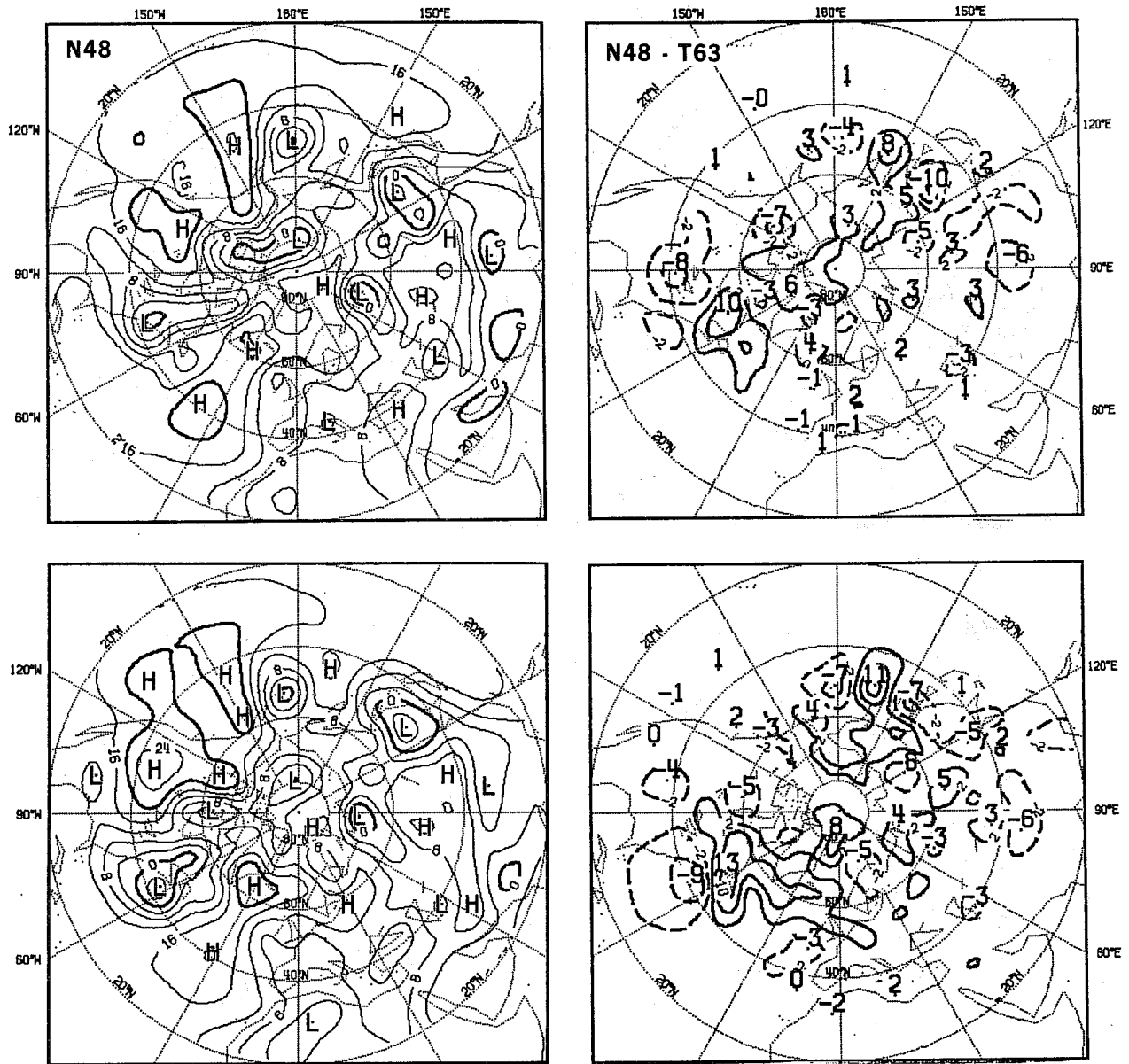
Fig.42   (Cont.)
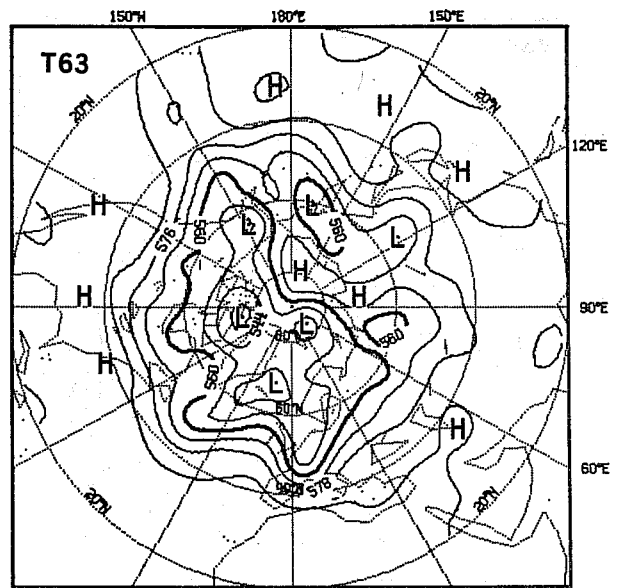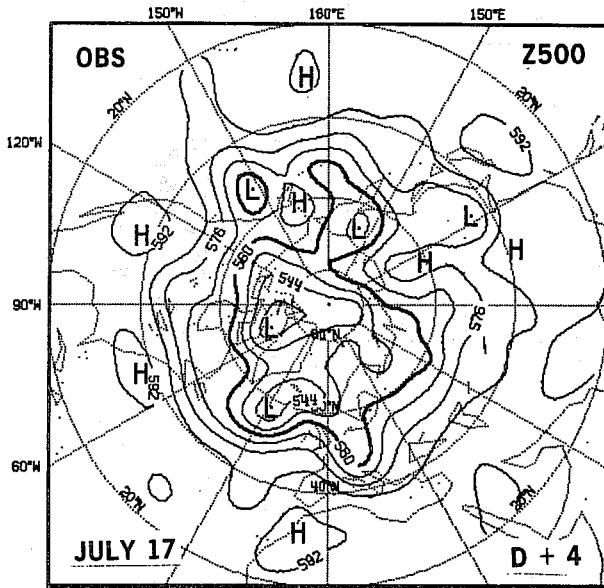
Fig.43a   JUL. 17, Z500: D+4
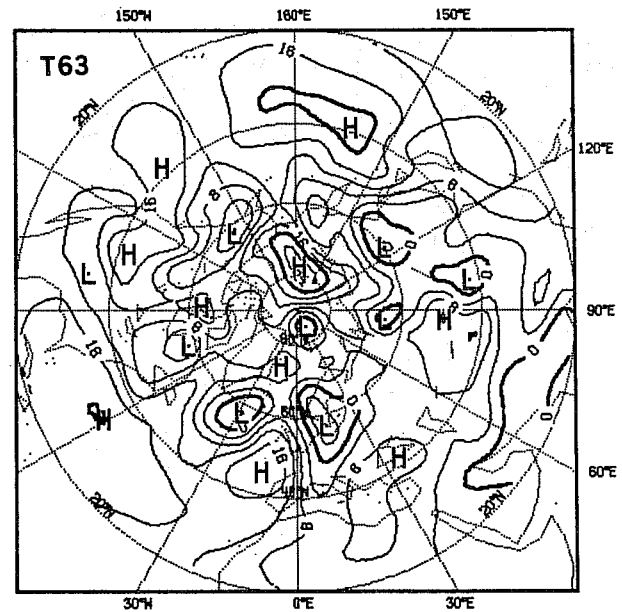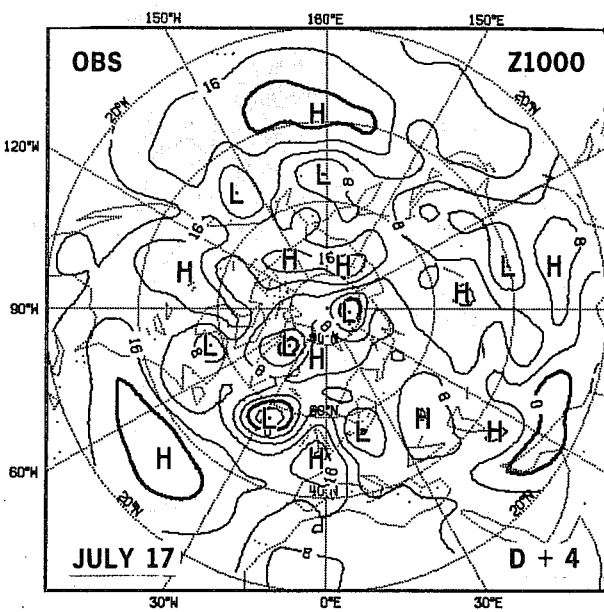
Fig.43b   JUL. 17, Z1000: D+4
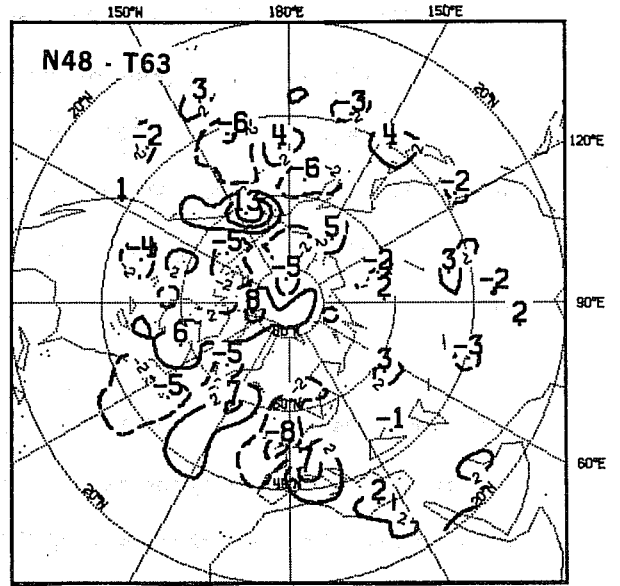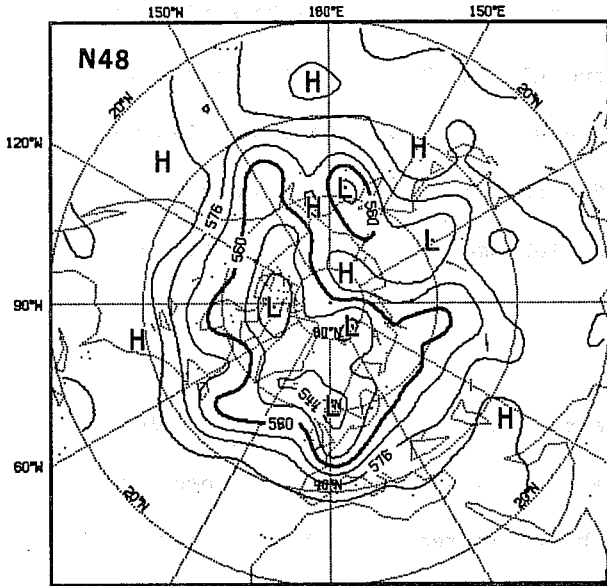
Fig.43a (Cont.)
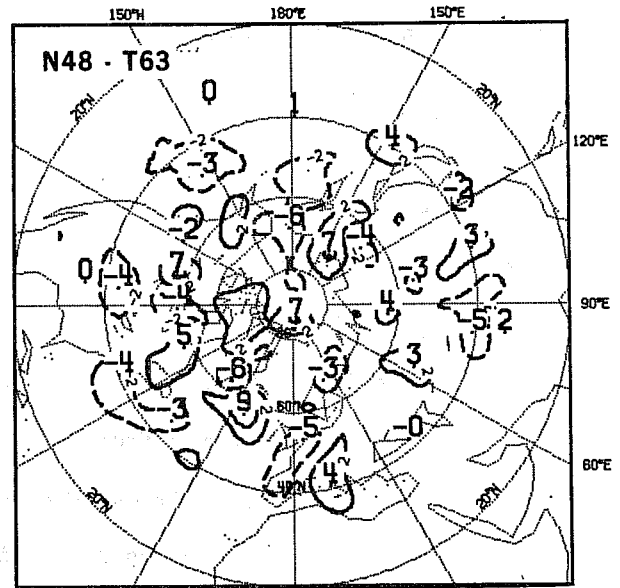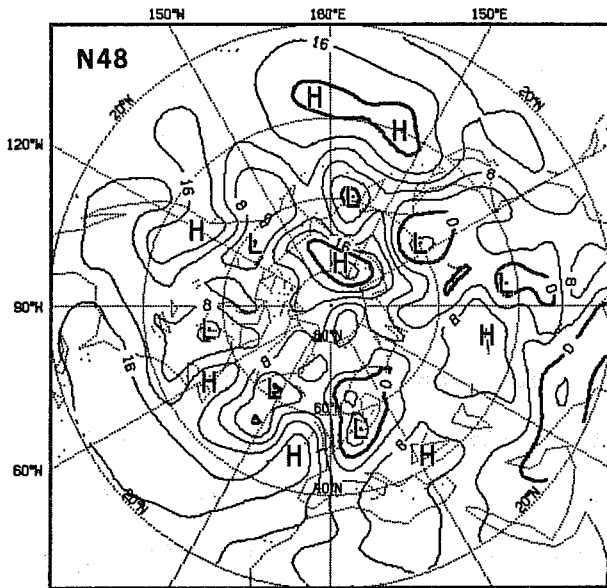
Fig.43b (Cont.)

Case of JUL 17

Fig. 43a: Z500, D+4

Fig. 43b: Z1000, D+4

The main FDs seen at 500 mb on D+4, across ATL, over EUR and in the ARC, are related mostly to the phase of three troughs in OBS. In all cases N48 troughs are lagging behind T63's which are also too slow. But the largest single difference is over Alaska, where T63 has intensified a trough in the OBS ridge position. N48 has a similar less intense feature and is thus less wrong.

At 1000 mb, FDs over ATL and EUR present a pattern similar to that at 500 mb in relation with depressions near latitude 60°N. In particular the position, intensity and shape of the low on 30°W is better forecast by T63. Elsewhere, differences are related to smaller OBS features. A trough over north WAM digs further inland in T63 than in N48. The lows north of EAS are too intense in the forecasts.

Case of JUL 24

Fig. 44a: Z500, D+2

Fig. 44b: Z1000, D+2

Here is another example of a type of FDs seen many times before. The interest of the case comes from the fact that relatively large differences occur already on D+2, while on other cases they were observed later on D+3 and D+4. T63 has more successfully simulated the strength and position of the short-wave trough on the east side of the larger trough at 500 mb and of the associated low at 1000 mb. It is also worth noting that the differences were ephemeral: on D+3 (not shown) the whole trough has much weakened and differences essentially vanished in this region.
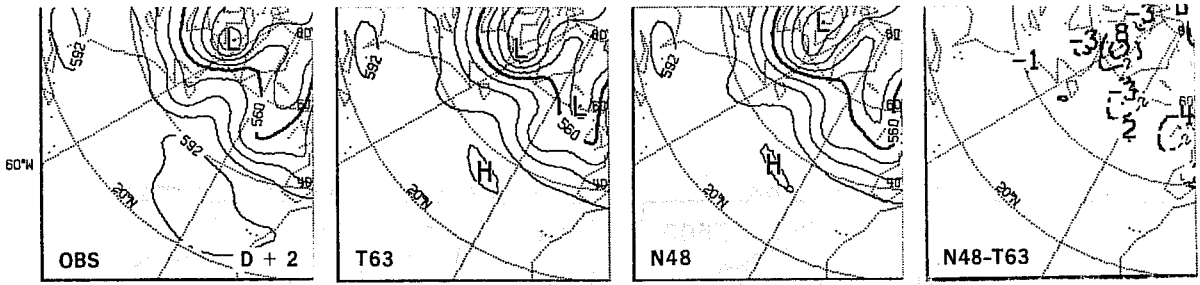
Fig.44a   JUL. 24, Z500: D+2
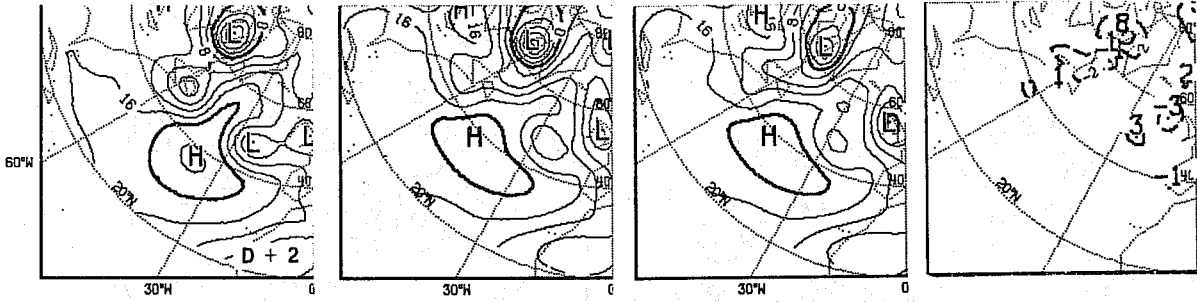
Fig.44b   JUL. 24, Z1000: D+2
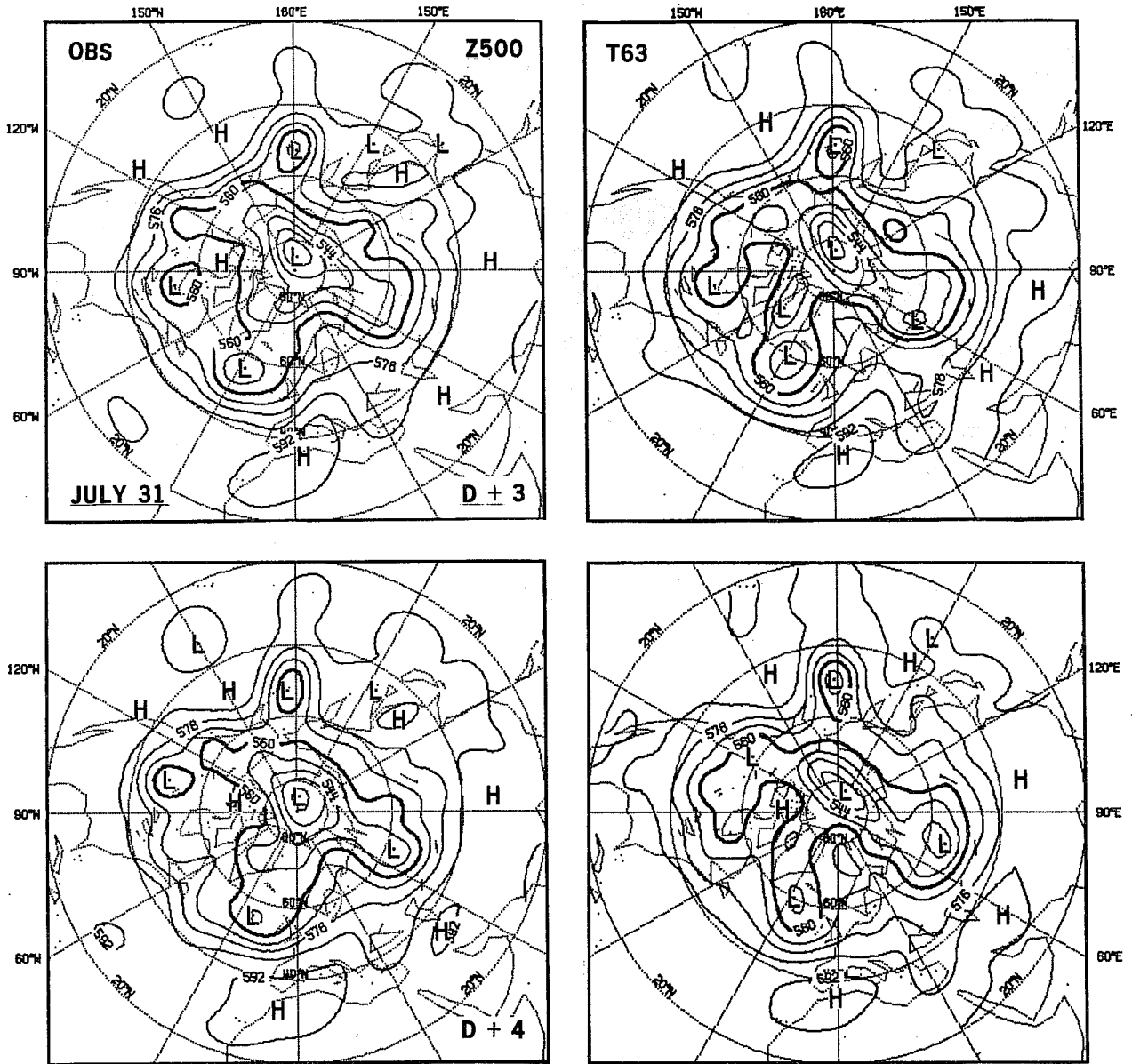
Fig.45   JUL. 31, Z500: D+3, D+4
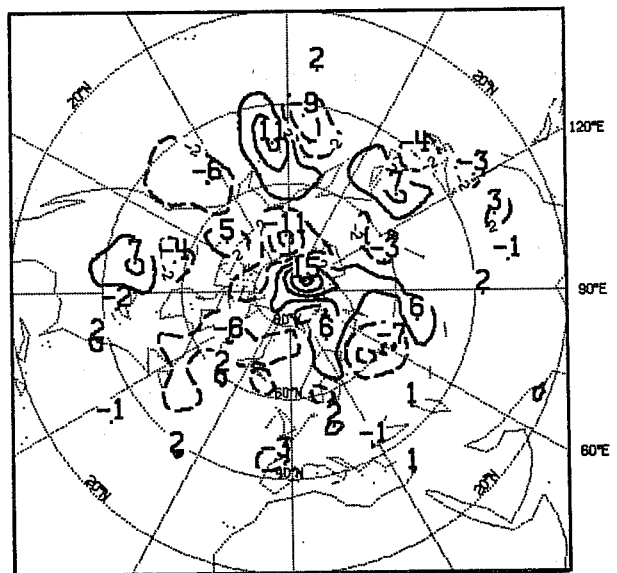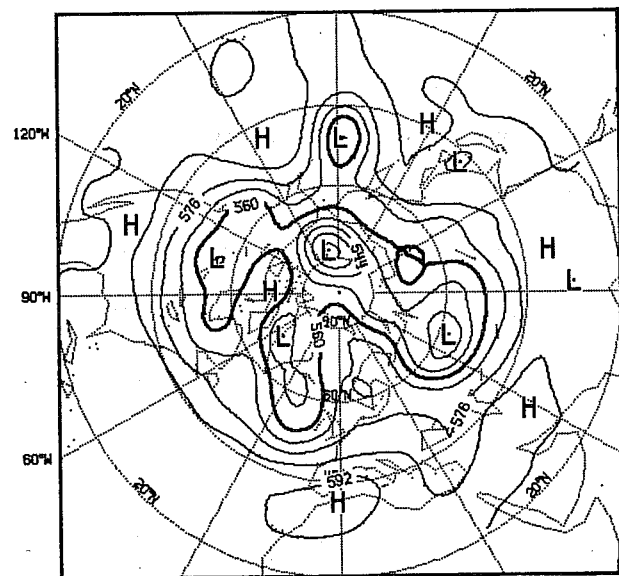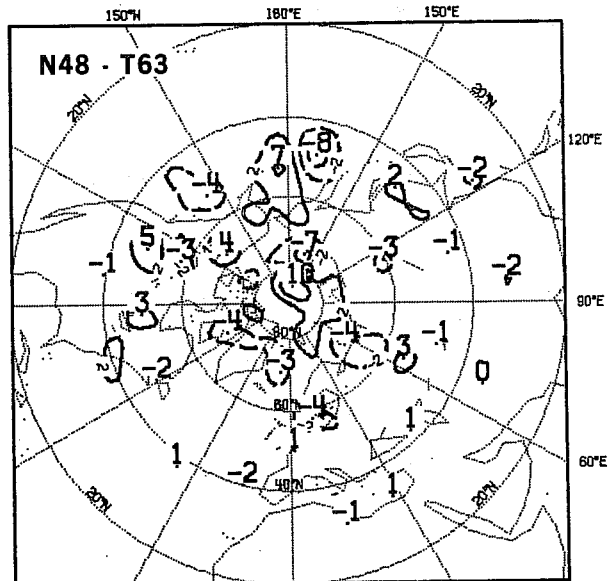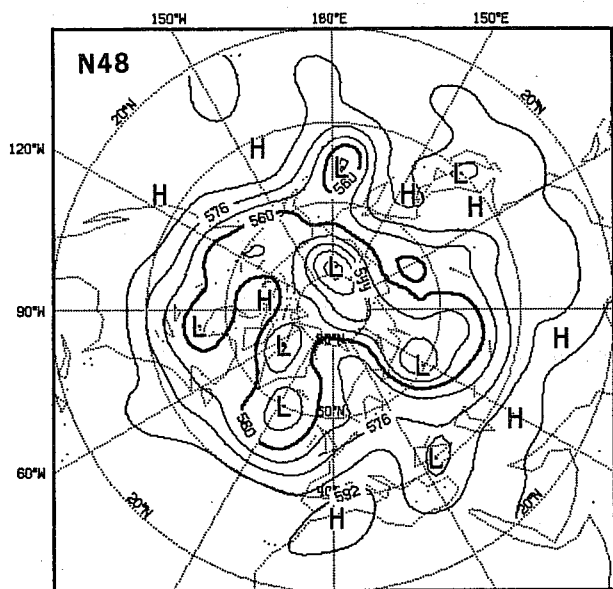
Fig.45 (Cont.)

Case of JUL 31

Fig. 45: Z500, D+3, D+4

In SUM, predictability is low and varies considerably from case to case. Largest AC score differences between forecasts on D+2, D+3 are found for good cases, JUL 31 being a good example of such a case.

The major FD at 500 mb is related to the average position and shape of the low near the pole. (N48 has dragged it unrealistically towards the South). The next largest FD occurs in the region of the cut-off low over PAC and is also a mixture of amplitude and phase differences. T63 is more successful in handling the evolution of both of these features.

Case of AUG 21

Fig. 46: Z500, D+3, D+4

Fig. 47: Z1000, D+3, D+4

The largest FD is once more over ARC in relation with the OBS large cyclone. T63 is again closer to reality, placing the low somewhere in between OBS and N48 with a better structure. The next largest FD is found over ATL where the cut-off is more accurately simulated by T63, digging further on the upstream side of the trough and climbing out faster on its downstream side. The shape of the trough over WAM is also better captured by T63.

At 1000 mb, on D+3, FDs show up mostly as amplitude differences and T63 is deeper in many places. On D+4, a phase component to the differences is also apparent over WAM and ATL. For once, N48 features appear ahead of T63. Nevertheless, over WAM at least the lows are more correctly positioned by T63. In ARC, the double low structure is better preserved by T63 although the lows are noticeably too deep.

3.6 Summary of the synoptic evaluation

We have presented maps or portions of maps for Z500 and Z1000 of OBS, T63, N48 and the differences N48-T63 in the range D+1 to D+6, but more often D+3 and D+4 and for 24 of the 53 cases involved in the comparison.

It was seen that, FDs occur preferably in regions of tight gradients (jets) in the height field and in relation to large scale flow features, such as troughs and ridges, lows and highs. The amplitude of differences and the scale of their patterns depend on the forecast range, but also on the amplitude and scale of the meteorological perturbations themselves, their speed of propagation and rate of change with time. Largest differences might be expected for fast moving or fast developing large disturbances. The type of atmospheric circulation that prevails over certain regions during a given period of the year determines preferred locations for the occurrence of large differences.

Various types of FDs may be recognised and classified in connection with the amplitude (over- or under-development differences) or position (phase speed and track differences) of meteorological disturbances. We must point out that both models show similar deficiencies in simulating the evolution of baroclinic disturbances, forecast errors being usually much larger than differences. Some of the outstanding and most systematic aspects of these deficiencies are: on the one hand, **both models tend to overdevelop large mature cyclones;** on the other hand, **speed and growth of new developments are frequently underestimated.** Overdevelopments are found to occcur almost exclusively over the north east PAC and EUR regions at all ranges: very often, they are observed already on D+2 but their effect on model performances is of course more catastrophic at longer range. The problem of underestimating the speed and growth rate of new cyclones can only be well diagnosed at short range when forecast new developments may be matched with observed ones. Preferred locations for these are along the east coast of the
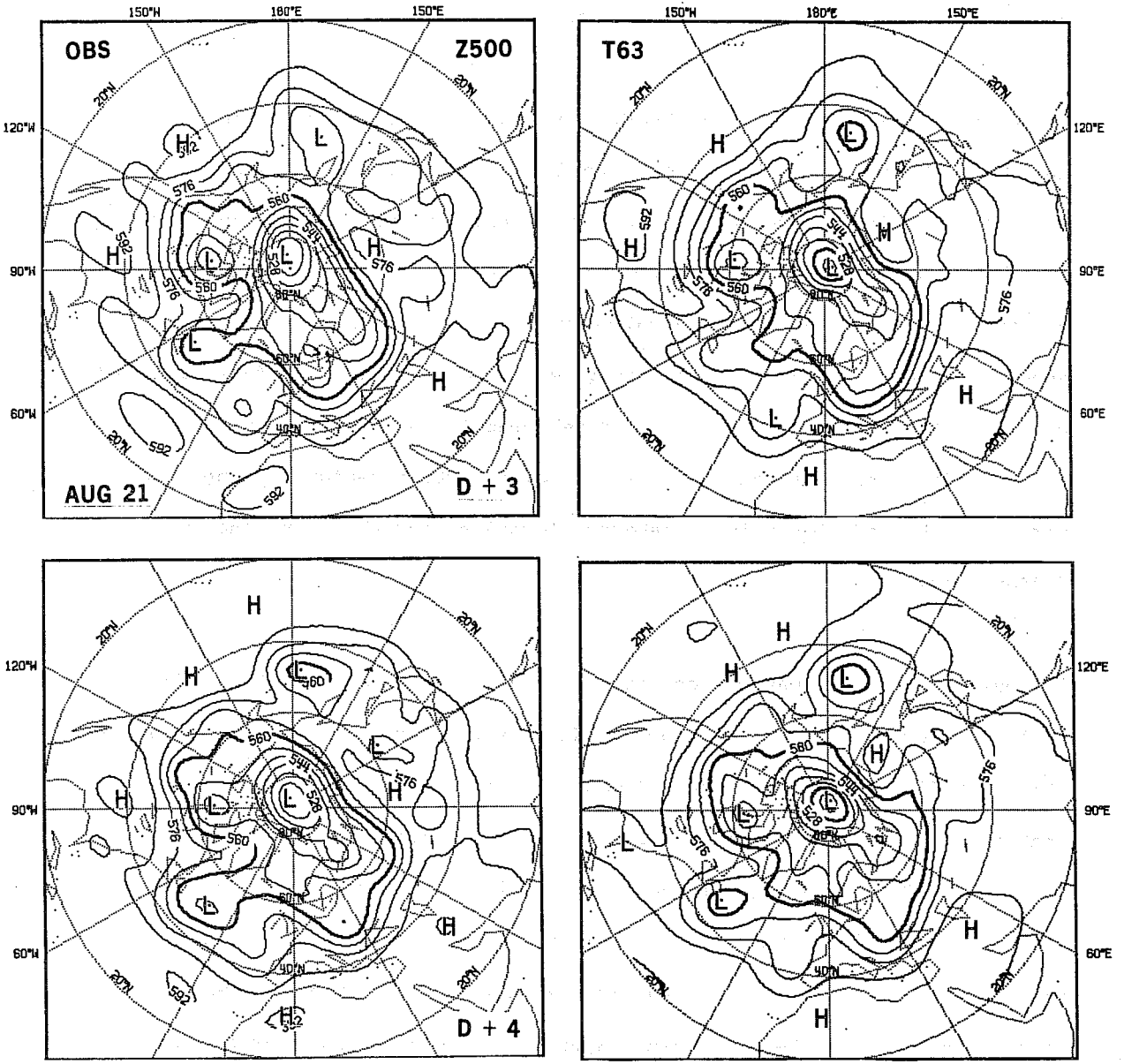
Fig.46　AUG. 21, Z500: D+3, D+4
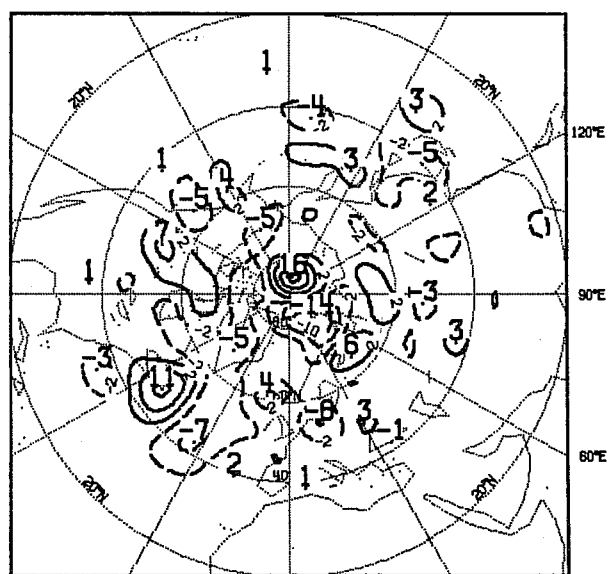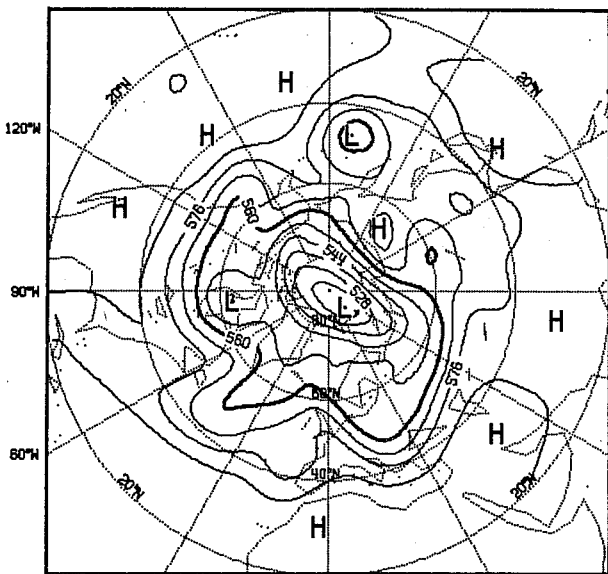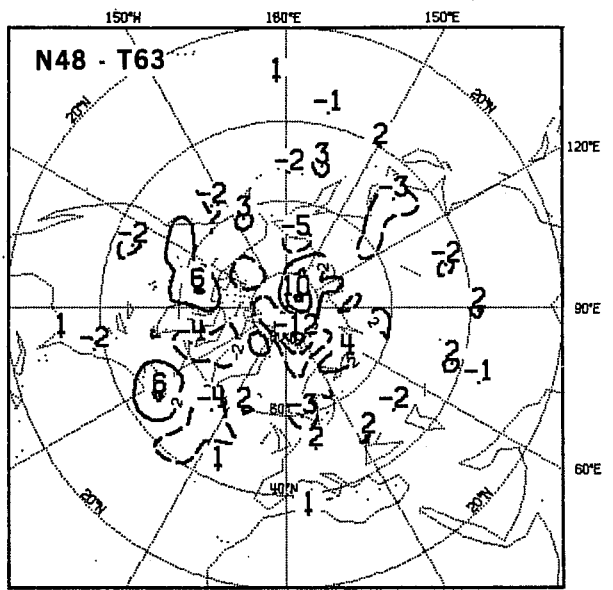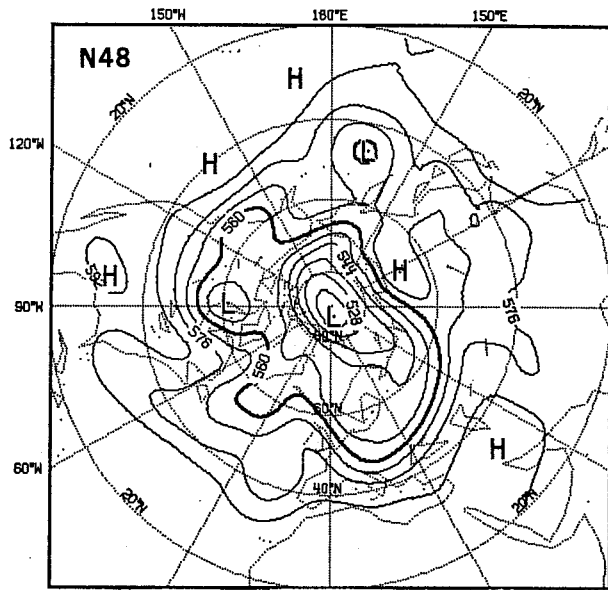
Fig.46 (Cont.)
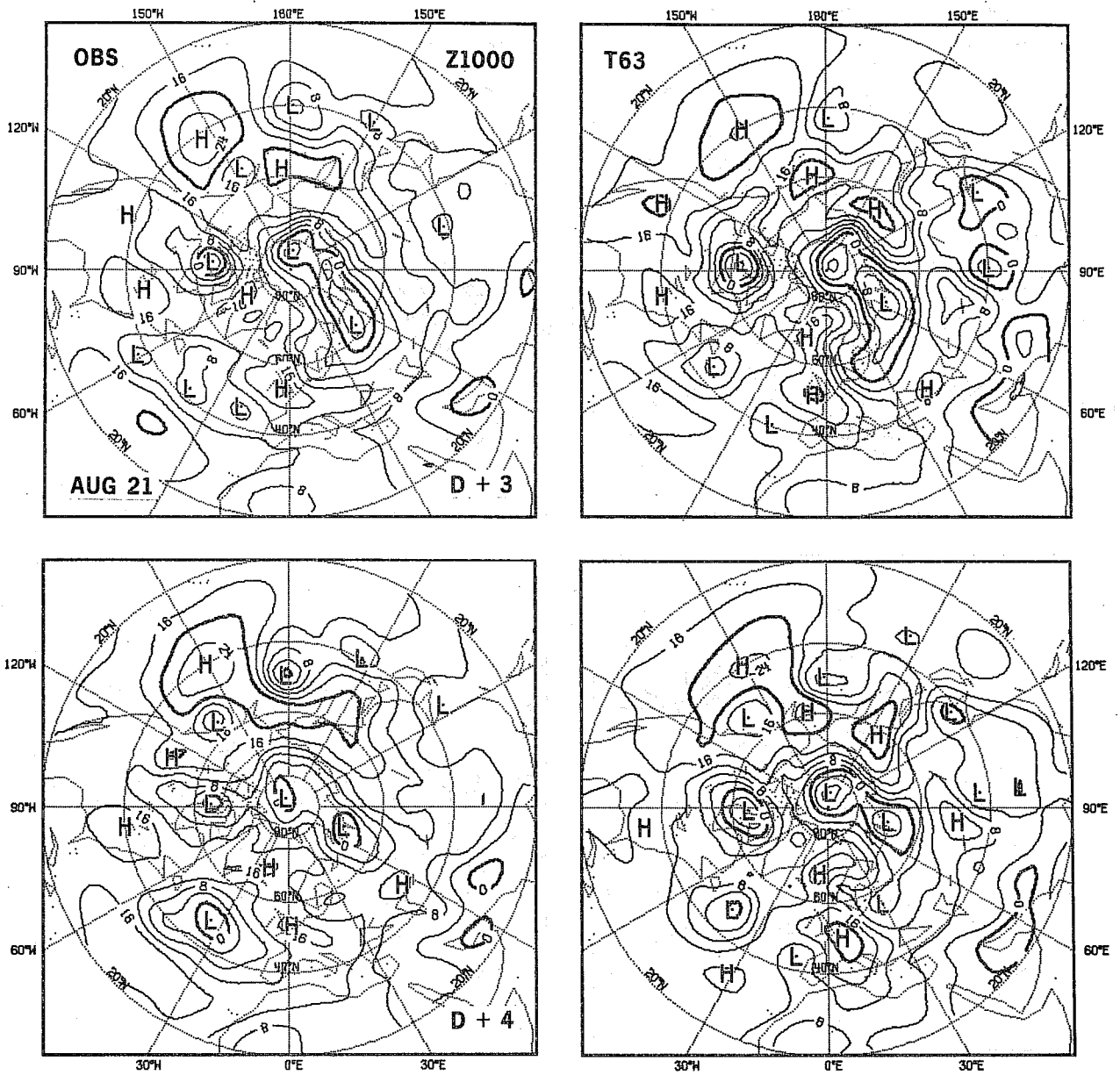
Fig.47   AUG. 21, Z1000: D+3, D+4
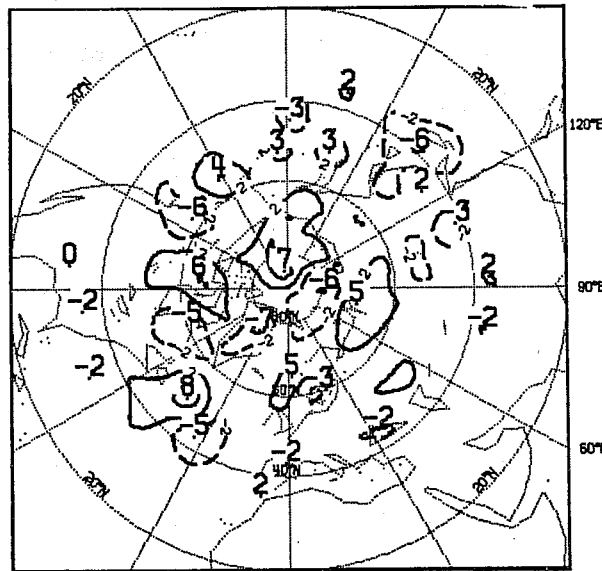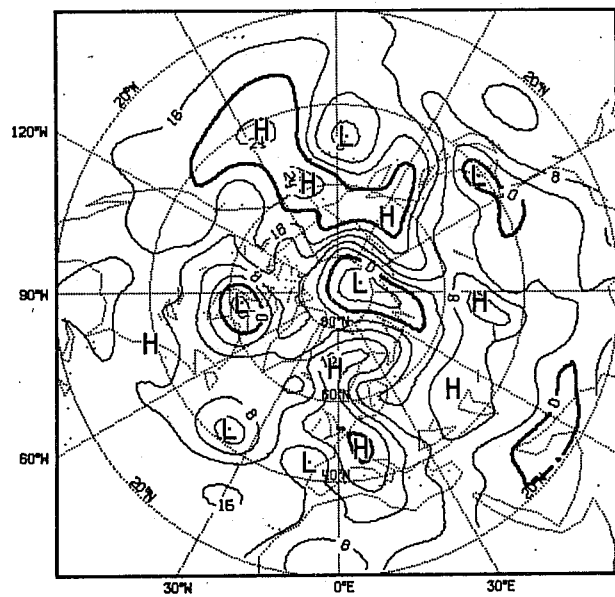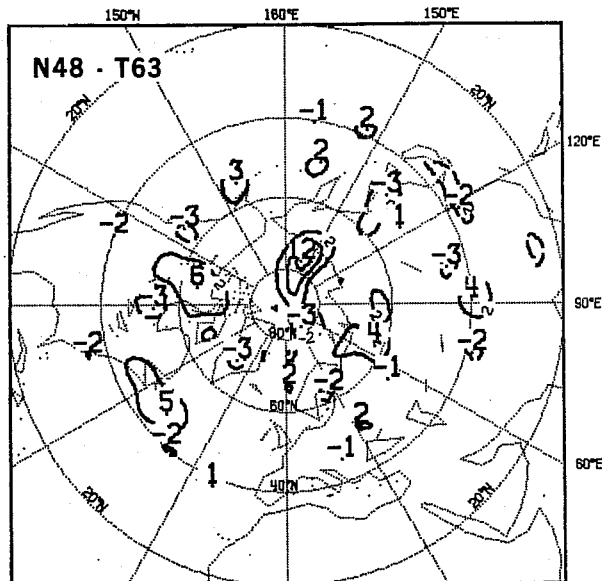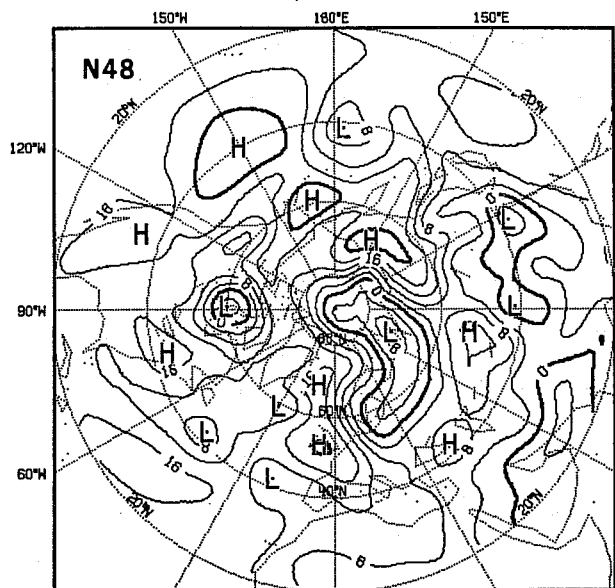
Fig.47 (Cont.)

Asian and North American land masses and over adjoining oceans (EAS and ATL sectors). FDs often appear as modulation of these errors: T63 is seen to overdevelop mature lows slightly more than N48 on average, over nothern EUR in particular. T63 is, however, more succcessful than N48 in predicting the speed and the track of small, fast moving, developing cyclones over EAS and ATL.

In connection with overdevelopment, the models are seen to maintain at 1000 mb a higher level of variance than observed. T63 level is slightly above that of N48. An examination of FDs also shows that maximum values are predominantly positive (T63 deeper than N48). At 500 mb, forecast variance levels are below that of OBS but T63 maintains somewhat higher values than N48. The meridional or eddy component of the flow remains higher in T63. In Section 4, we shall show that in general T63 retains more eddy KE at all levels and keeps a lower ratio of eddy to zonal KE. Thus, in contrast to overdevelopment relative to N48 at 1000 mb, T63 appears more able than N48 to predict the correct amplitude of observed features at short range and to maintain the correct level of variance at longer range.

The most common and most characteristic type of FDs found in the range D+1 to D+5 is not related to the amplitude of systems but to their position (phase differences). When measured in the direction of their displacement, **the systems in N48 are almost consistently found to lag behind those of T63.** This is true for the overwhelming majority of cases. When compared to OBS, T63 systems are also found to be better in phase in the great majority of cases. Phase differences are of course easier to diagnose when the mean flow is fairly strong and straight and when the moving perturbations are fairly stable. This is why the best examples of phase differences are found over the PAC and ATL regions and along the cyclone tracks. And because these tracks have a SW-NE orientation, phase difference patterns usually have the

same orientation. However, a track difference is also obvious in many cases. On average, the storms in T63 move not only faster but also on more northern trajectories. Again, on average, this is nearer reality.

Usually mean FDs are small. Accordingly, mean score differences, mostly in favour of T63, remain relatively small. But, **on a few more sensitive cases, large local phase and amplitude differences develop fairly quickly, producing by D+4 and D+5 forecasts synoptically very different over large parts of the Northern Hemisphere.** It is important to note that such cases (e.g. SEP 30, OCT 25, NOV 1, NOV 15,) which were likely to produce large score differences were also usually very much in favour of T63.

4. DIFFERENCES IN MODEL MEAN BEHAVIOUR

In this section, we present results from ensemble averages computed on a seasonal basis using the convention adopted previously for AUT, WIN, SPR and SUM.

4.1 Height field

In Section 2, we saw from the scores (Fig. 1) that forecasts remain closer to each other than to reality for a long period. This is illustrated in Fig. 48 comparing the mean errors in WIN for Z500 and Z1000 on D+4. Both models suffer from almost identical mean errors at both levels. The synoptic differences noted in Section 3 were thus essentially transient (phase differences for example) and have essentially vanished after averaging. The mean error patterns on D+4 are remarkably similar up to their fine details.

At longer range, the forecasts diverge and larger mean differences are more probable. The next figures (49 and 50) present the mean errors on D+10 as in Fig. 48 for the 4 seasons. Pattern similarity remains high even at this range in all seasons, in spite of large seasonal changes. Quasi-permanent
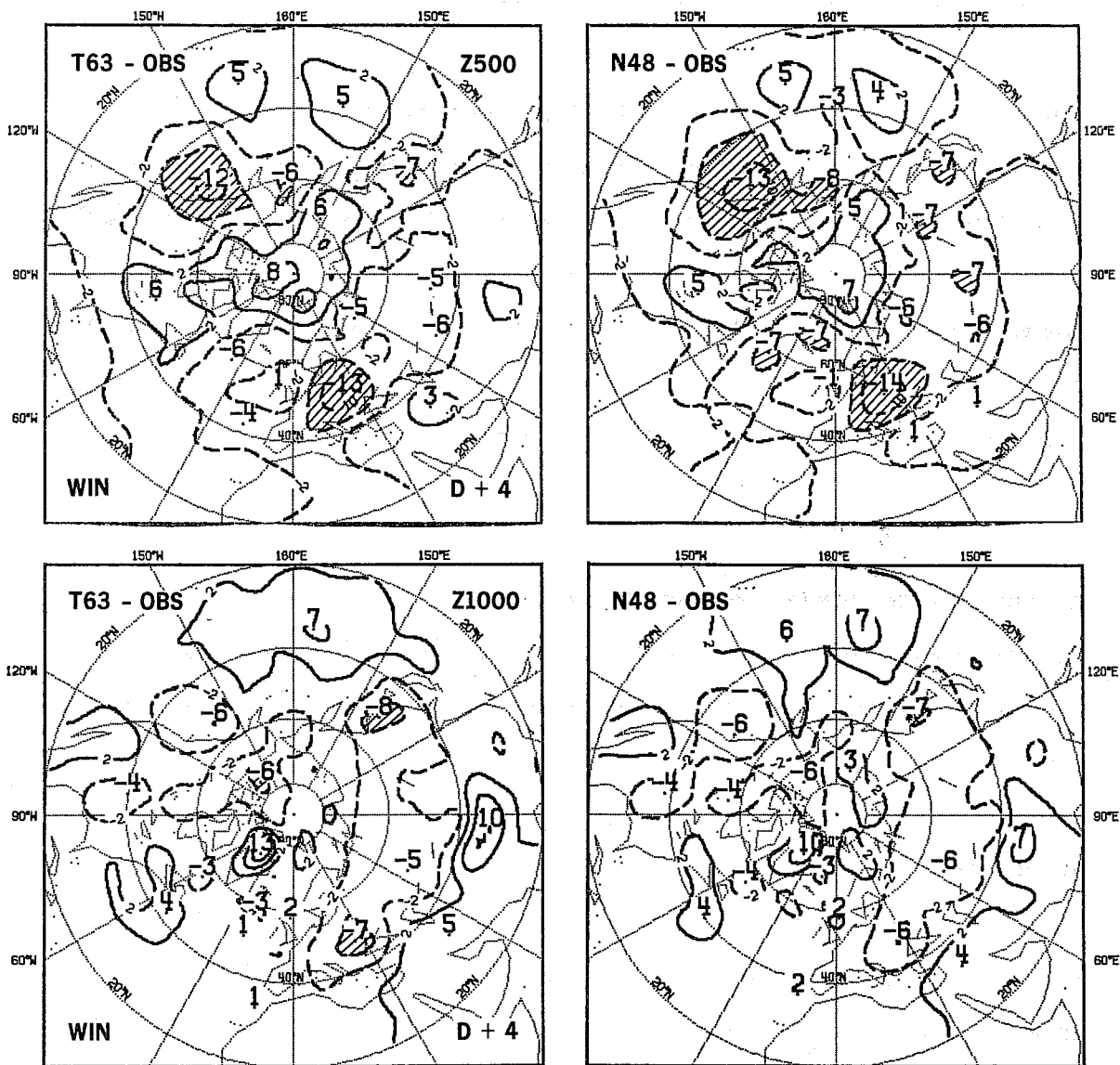
Fig.48   Mean errors: Z500, Z1000, D+4 for WIN (T63-OBS, N48-OBS)
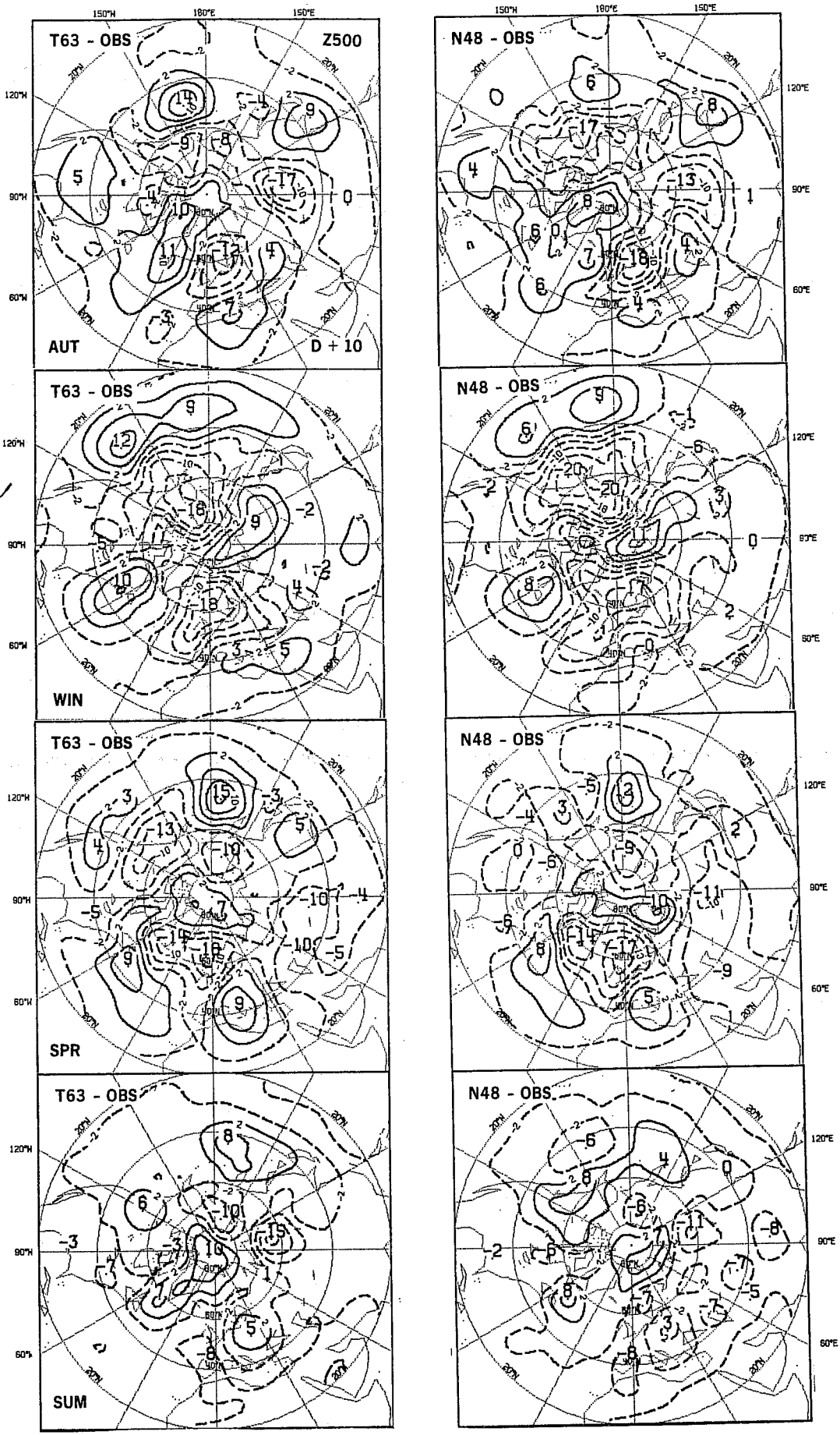
Fig.49 Mean errors: Z500, D+10 for AUT,WIN,SPR,SUM(T63-OBS,N48-OBS)
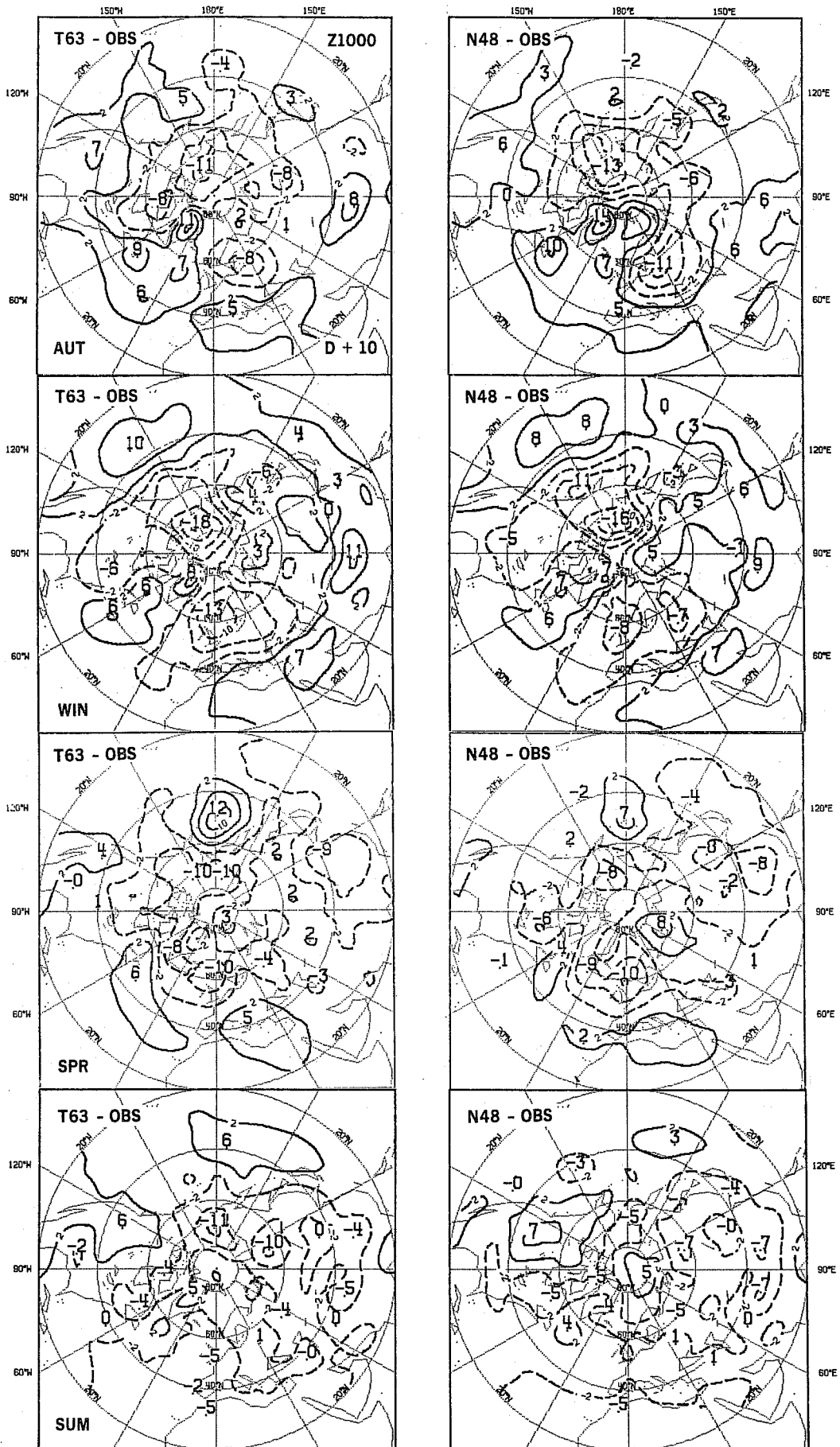
Fig.50 Mean errors: Z1000, D+10 for AUT,WIN,SPR,SUM(T63-OBS,N48-OBS)

features of both models are characterized by overdeepening over EUR and north

PAC (the region covering Eastern Siberia and Alaska). The latter region also

shows the largest model differences. At 500 mb for example, T63 has smaller

negative errors there in AUT and WIN but the reverse is true in SPR and SUM.

At 1000 mb, the errors and differences are smaller than at 500 mb but their

structure is not essentially different.

Note for future reference that in WIN, at 500 mb, the heights in T63 are

somewhat higher throughout most of the middle and high latitudes.

4.2 Kinetic and available potential energy levels

A notable mean behaviour difference between the models is to be found in the

levels of kinetic energy (KE) that they are able to maintain. Although both

models show a tendency to lose eddy KE with respect to OBS, N48 losses are

larger. This is illustrated in Fig. 51 comparing the medium (4-9) and short

(10-20) wavenumber bands of the mean tropospheric (82.5-20°N; 1000-200 mb) KE

levels. Note that differences do not develop gradually: they have already

reached a plateau by D+1. In Fig. 52 we present spectra at 4 pressure

levels of the time averaged (D+0 to D+3) values of KE for the SUM season. We

see that T63 has more KE in nearly all wavenumbers and at all pressure

levels. This is in good agreement with the synoptic impressions gained in

Section 3 that T63 has more variance. In relative terms, well emphasized by

the logarithmic scale of the graphs, the differences are larger in the

shortest waves. This higher level of short- wave tropospheric KE in T63 is

maintained at longer range. This is shown in Fig. 53 displaying the mean

spectra (D+7 to D+10 averages) for all 4 seasons. At low wavenumbers,

differences are smaller and not systematic.

The mean differences in available potential energy (AE) are more important

than those in KE. Both models lose a substantial amount of AE before D+1,
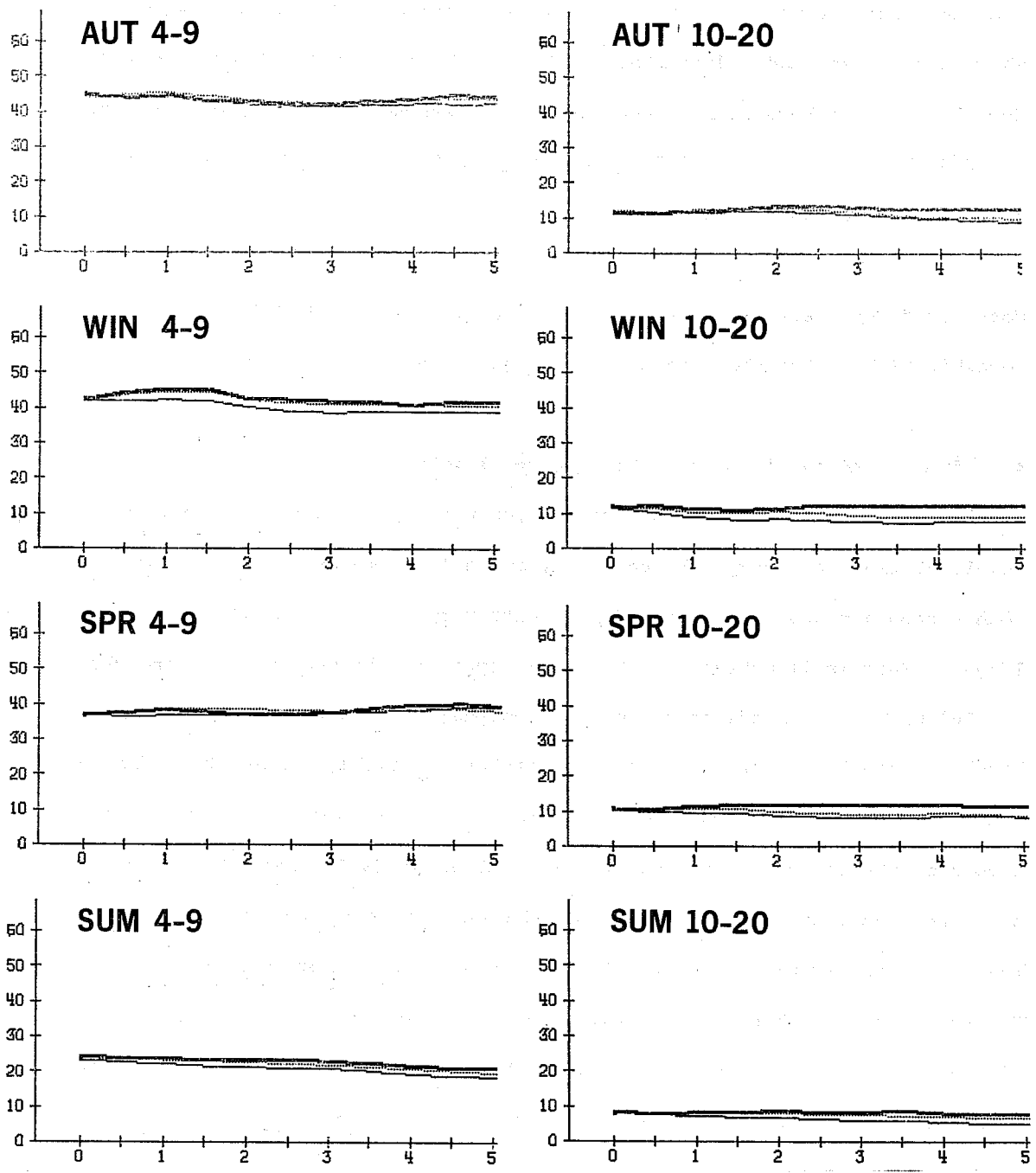
Fig.51 Time evolution up to D+5 of total averaged geostrophic KE(kJm^{-2}) in the troposphere (1000-200mb) between 20°N and 82.5°N for the wavenumber bands (4-9) and (10-20) in AUT,WIN,SPR,SUM. (OBS= ⎯⎯⎯ ; N48= ⎯⎯⎯; T63= ------)

110
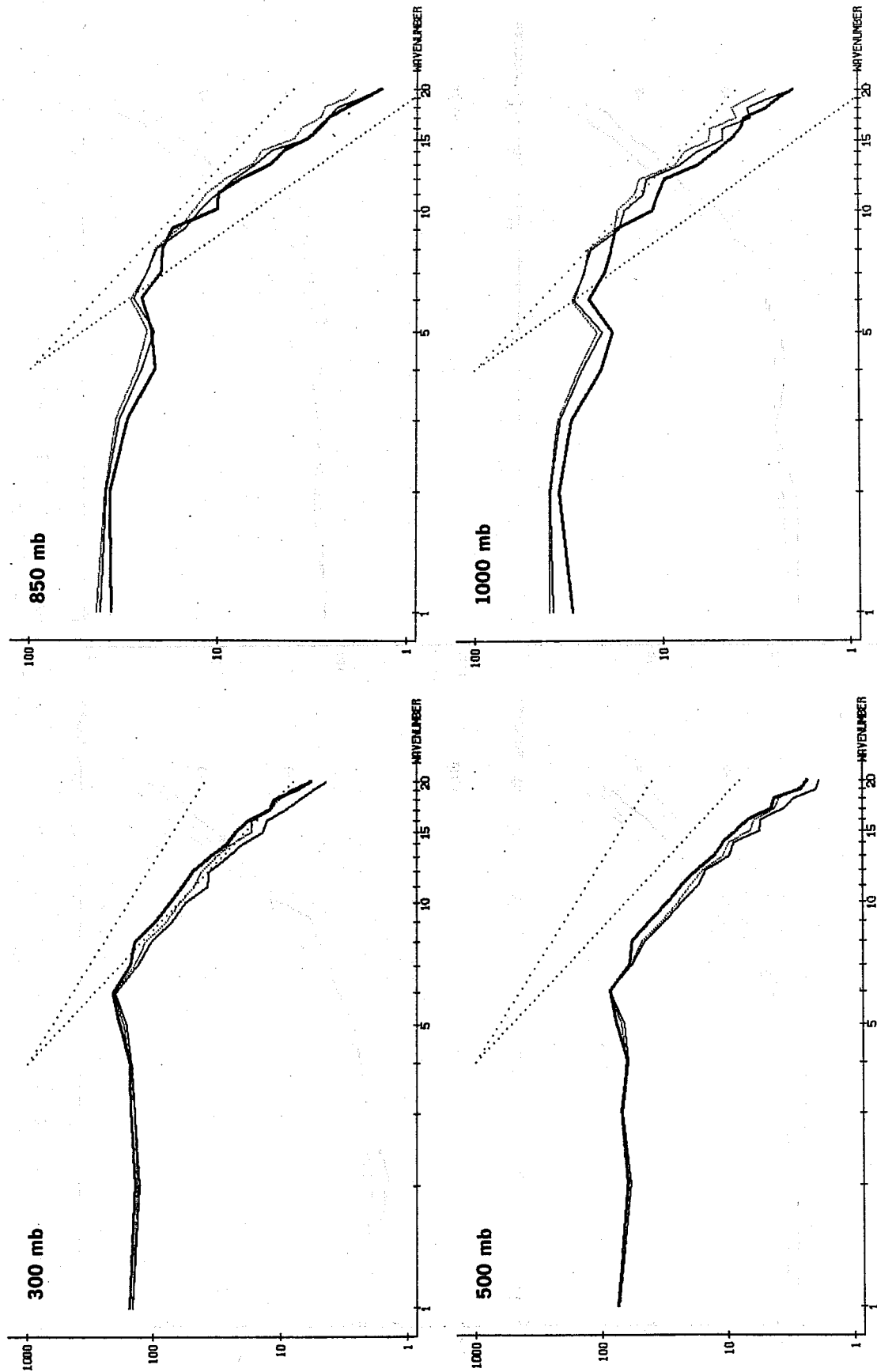
Fig. 52 Spectra of averaged geostrophic KE (kJm^{-2} bar^{-1}) at 300, 500, 850 and 1000 mb between 40°N and 60°N in SUM for the forecast period D+0 to D+3 (OBS= ———— ; N48= ———— ; T63= ----).
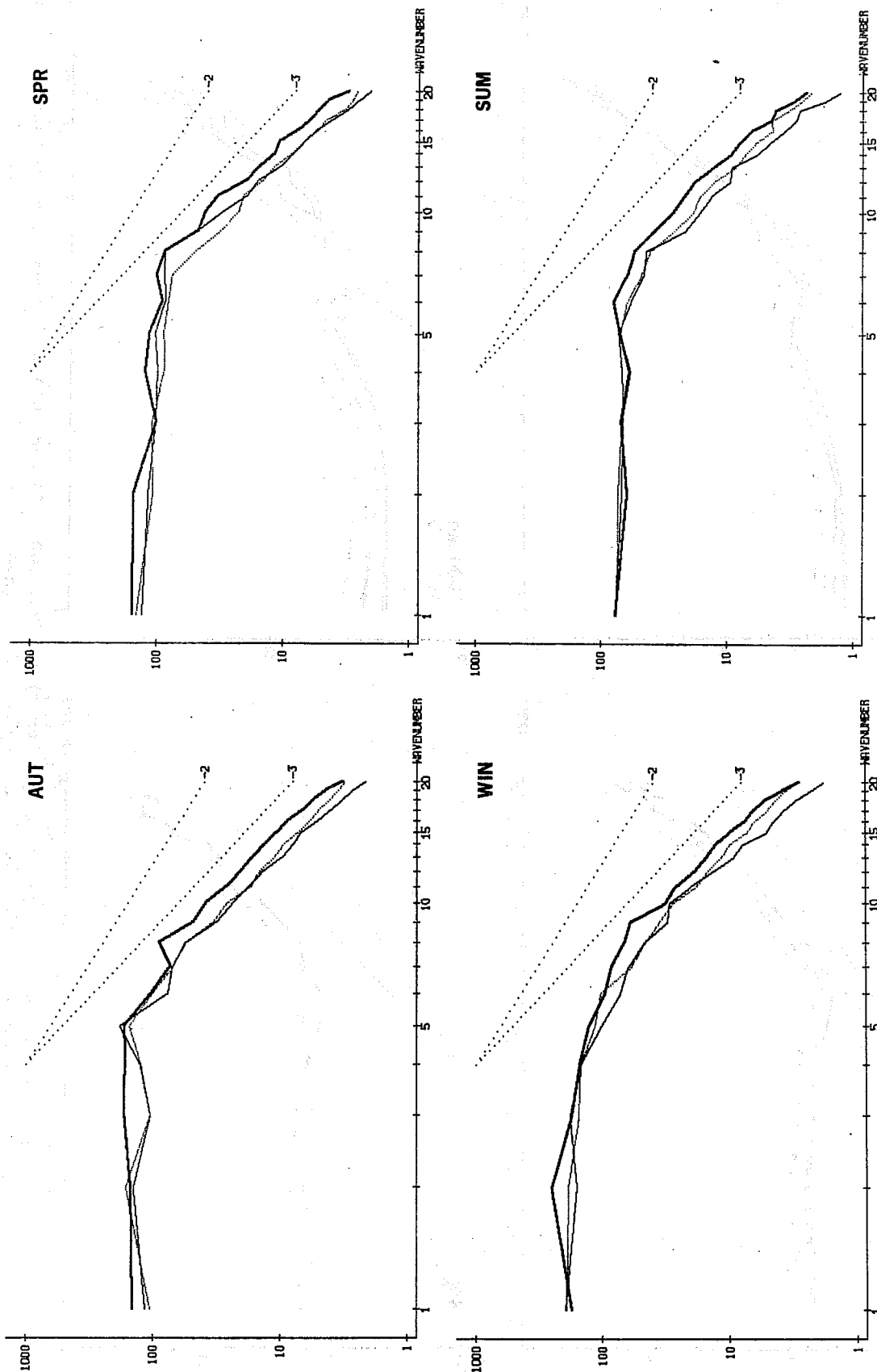
Fig. 53 Spectra of averaged geostrophic KE(kJm^{-2}) in the troposphere (1000-200 mb) between 40°N and 60°N in AUT, WIN, SPR and SUM for the forecast period D+7 to D+10 (OBS= ⸺; N48= ⸺; T63= ----).
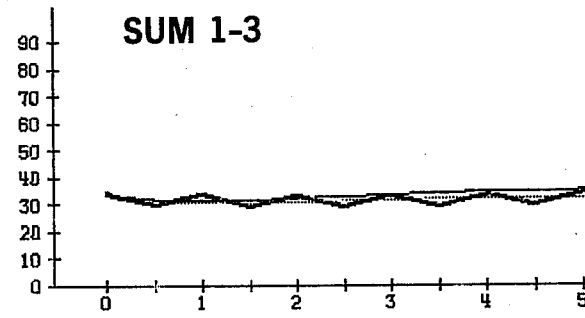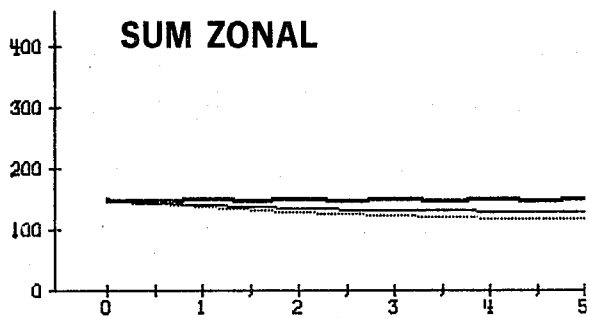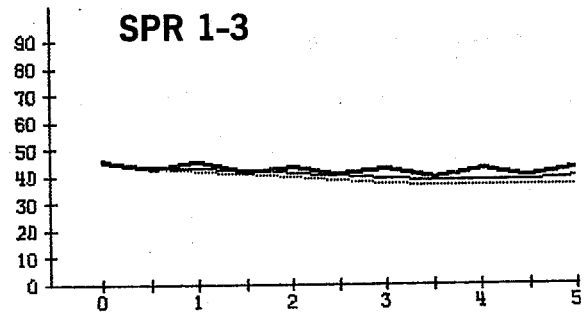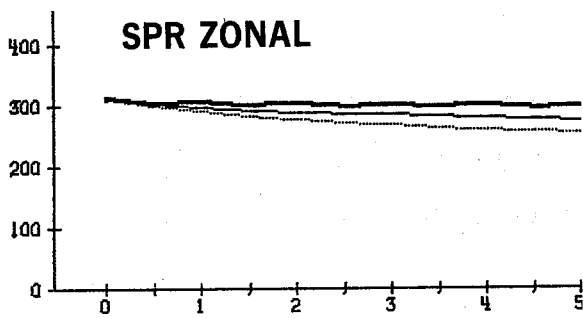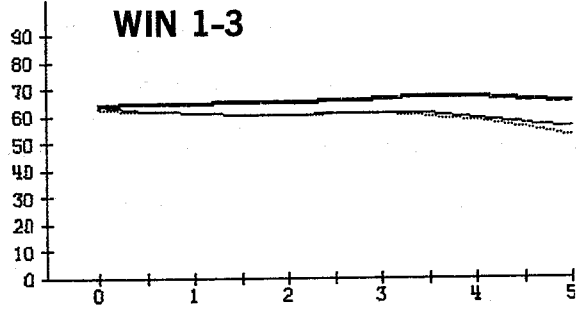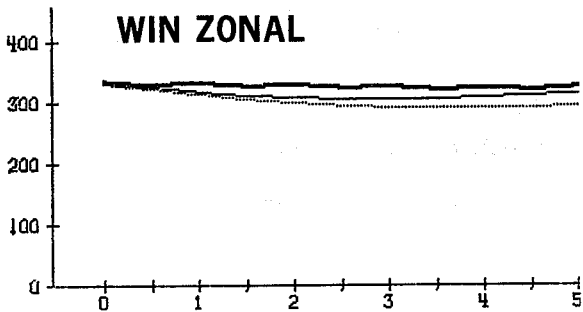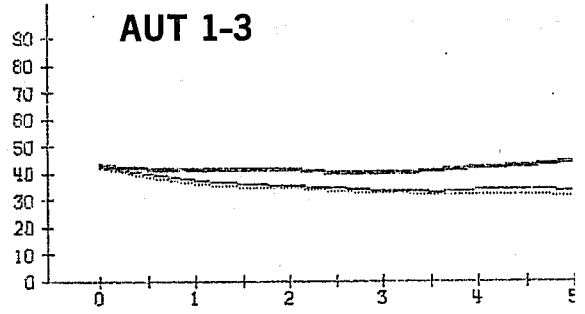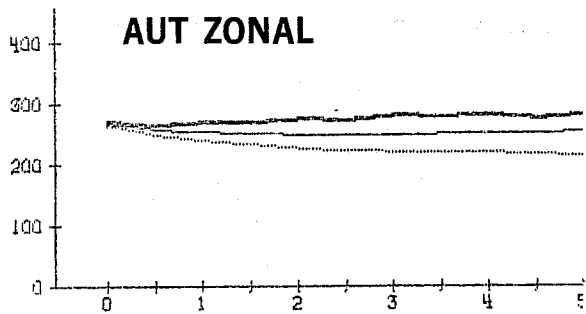
Fig. 54 Time evolution up to D+5 of the averaged AE(kJm^{-2}) in the troposphere (850-200 mb) between 20°N and 82.5°N for the zonal and small (1-3) wavenumbers in AUT, WIN, SPR, SUM (OBS=————— ; N48= ——— ; T63= ------).

particularly zonal AE. Superimposed on this initial loss, one also observes a gradually increasing model difference, representing a deficit of T63 with respect to both N48 and OBS. This is illustrated in Fig. 54 showing the trends in zonal and long-wave AE levels up to D+5 (these are not altered from D+5 to D+10) in all 4 seasons. For the long-wave eddies, the trends are similar to the zonal ones, but the relative deficit in T63 is not as pronounced.

At this point it is worth noticing that no such differences were observed in the experiments reported in TR23. We already mentioned that changes were made only to the physical parameterization of both models in the interval and of course the initial datasets came from different sources (NMC/ECMWF).

4.3 Temperature field

One of the most notable deficiencies of both models is their tendency to cool the atmosphere globally by as much as 2°K in a 10-day forecast. Most of this cooling occurs in the stratosphere at all latitudes and in the tropical troposphere but a certain amount of cooling is observed in the middle-latitude mid-troposphere. This is the region where we find a difference in the mean behaviour of the models, i.e. a larger cooling in N48. Fig. 55 shows mean pressure-latitude cross-sections of the temperature errors averaged for the 4 seasons and between D+7 and D+10. Note that model differences are somewhat larger during the first 2 seasons and particularly in WIN where a well defined negative centre (60°N, 500 mb) is apparent. In Fig. 56 we look at temperature error maps at 500 mb for that season at both D+7 and D+10 (the hatched areas correspond to a fictitious cooling of more than 3°K). Note that cooling increases from D+7 to D+10 and that the areas where cooling is intense correspond approximately to the typical cyclone-track regions.
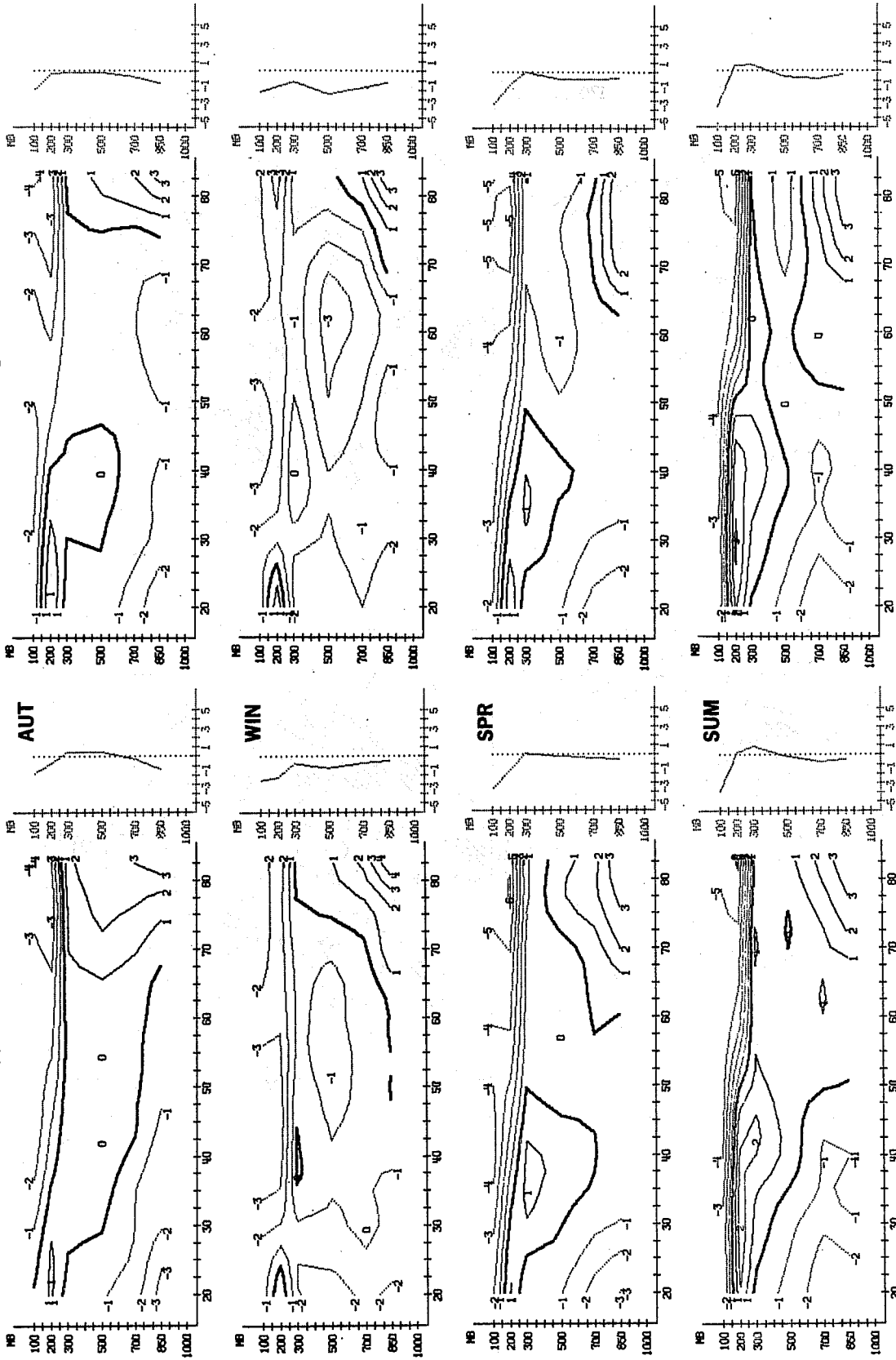
Fig. 55 Pressure-latitude cross-sections of the mean T63 and N48 errors of T(°K) in AUT, WIN, SPR, SUM for the forecast period D+7 to D+10.
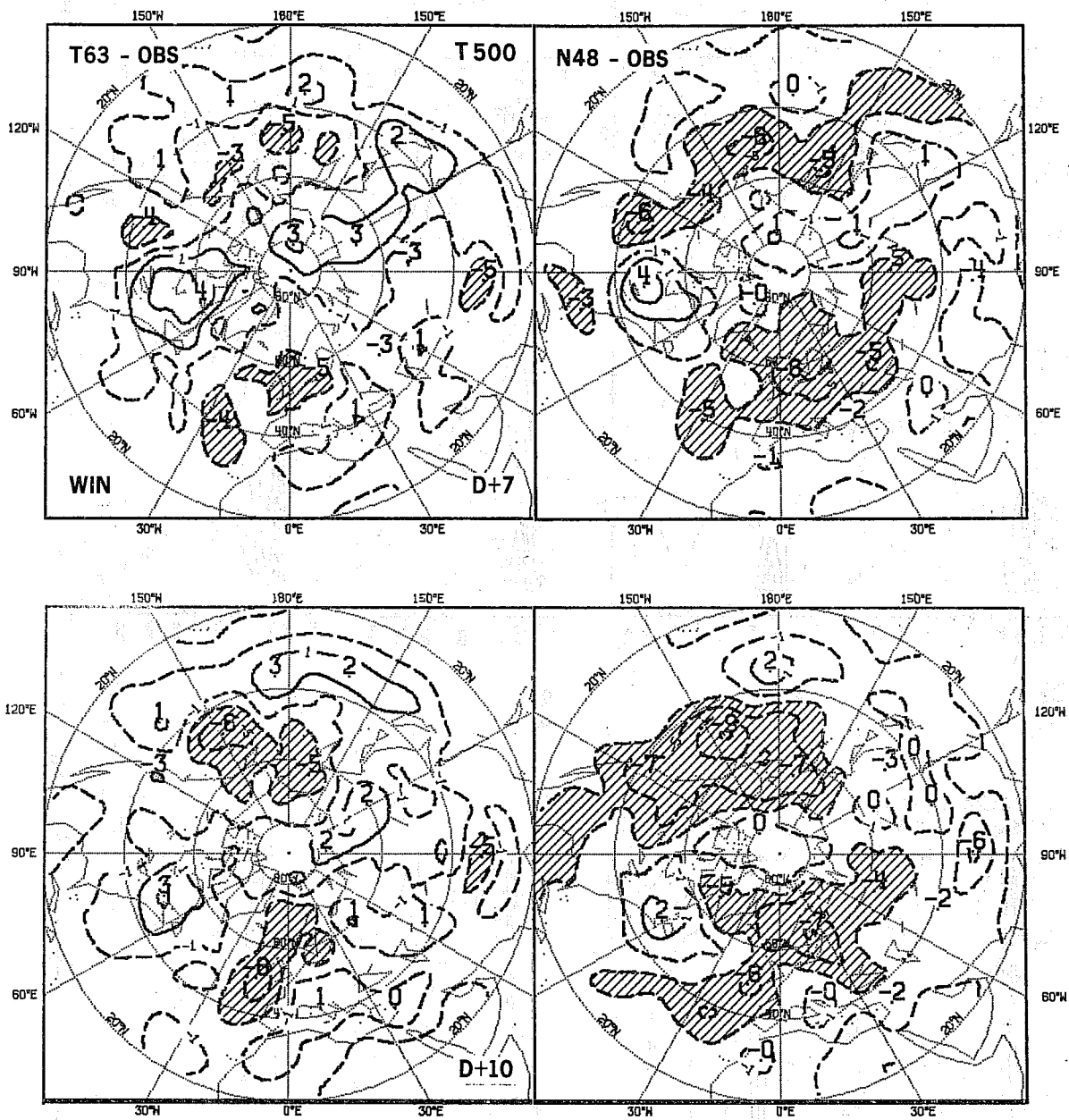
Fig. 56   Mean errors in T500 ($^{\circ}$K) for T63 and N48 on D+7 and D+10 in WIN.   The
hatched areas correspond to regions with cooling errors greater than
3°K.

In so far as 1000-500 mb thicknesses are affected by midtropospheric cooling, mean N48-T63 temperature differences should also cause mean height differences at 500 mb. This can be verified by comparing Fig. 48 (WIN) and Fig. 56.

The amount of middle-latitude mid-tropospheric cooling has been found to be sensitive to the level of horizontal diffusion applied to model variables. Furthermore a connection between T and AE differences can be made. Thus, it appears that most of the observed systematic differences, in AE, in T, to some extent also perhaps in KE (since KE is not insensitive to horizontal diffusion) and partially at least, in Z, are somehow connected.

The effect of horizontal diffusion on KE, AE and T will be studied in Appendix B and we shall further discuss its impact on the above mean N48-T63 differences.

5. NATURE OF THE FORECAST DIFFERENCES

5.1 Main model differences: techniques and other causes

In the previous sections we objectively compared the performance of the models but also brought into evidence the synoptic FDs as well as mean FDs. Before attempting to investigate their nature, it is worth recalling the model (and other) differences which might have led to them.

The fundamental difference lies of course in the numerical techniques and it is our main purpose here to analyse its effects. A few others have nevertheless been found to be of consequence: post-processing of the forecast fields, horizontal diffusion and coast-line definition. The first two items have been investigated thoroughly and will be the subject of two appendices. The third item generated important differences beyond D+5 on one case discussed in Section 5.3. At shorter range (one case is analysed in Section 5.2), the effect was considered (only 2 tests were made) to be of small

consequence and not further studied.

In practical comparisons of techniques, when spectral and grid-point models of fixed resolutions are involved, the question of resolution and associated truncation errors cannot be totally neglected. The resolution of a triangularly truncated spectral field is usually defined in terms of its maximum wavenumber, that of a grid point field in terms of the distance between two grid points (grid size). A definition which can be used for both techniques can however be achieved in terms of the number of degrees of freedom used by either of them in order to approximate the same continuous field. For a meaningful comparison of technique it is necessary that resolution be comparable in order to ensure that truncation error differences are small. The question whether this was true in the present case will be dicussed in Section 5.3.

For a given truncation, spectral methods to compute derivatives are exact and are usually aliasing free for linear and quadratic terms. Therefore in these contexts technique differences constitute in fact grid-point technique errors. These are generally classified according to the linear or non-linear character of the equations involved or of given terms or operators therein. Atmospheric dynamics is mostly dominated by the so-called advection terms. The main grid-point technique errors should thus be analysed accordingly.

5.1.1 Computation of derivatives. "Linear" phase errors

The computation of derivatives using finite differences usually leads to systematic linear advection phase errors (amplitude errors can be avoided) which depend on the scale of the phenomena in relation to the mesh size. The smaller the scale, the larger the error. For a physical system with many wave components there is therefore a fictitious dispersion of the "wave packets". If the mesh size is not uniform throughout the domain phase errors

will in addition be dependent on the location within the domain and on the direction of the flow.

Realistic atmospheric flows are highly non-linear but over relatively large regions of space (and time) where the mean wind is fairly constant, linear concepts are approximately valid and may be used to characterize grid-point technique errors in the displacement of well-organized perturbations for example. In this regard, it is already clear, from the observed systematic N48-T63 phase differences (Section 3) that N48 suffers from "linear" phase errors. Here the word "linear" refers less to the nature of the errors than to the type of mathematical analysis used to define them. More details about this type of N48-T63 FDs will be given in Section 5.2.

Of course, second or higher order derivative operators such as the Laplacian (∇^2) are also misrepresented by finite differences. An obvious candidate is also the horizontal diffusion operator $(K\nabla^4)$ whose finite difference equivalent can be shown (Bass and Orzag, 1976) to be much less effective for the highest wavenumbers. But the possible effect of this difference must have been completely hidden by larger differences caused by, at the beginning, the use of different operators and, later, the use of different values of the coefficient K.

5.1.2 Aliasing and coupling errors

A type of non-linear error characteristic of finite difference techniques is the aliasing error which occurs whenever the non-linear interactions give rise to terms which cannot be resolved by the finite set of grid points. The importance of this error is directly related to the magnitude of short-wave components. For horizontal resolutions such as N48's it is unlikely to be large due to the very small amount of energy in the smallest waves, compared to the one in the meteorologically significant longer waves.

As noted by Sundquist (1975) however, the aliasing error (in particular in the thermodynamic equation) will be largest near the tropopause in regions of steep orography, with models using a σ-coordinate system. But these are also the places where quasi-discontinuities appear in the temperature field inducing larger errors in the temperature advection terms because of the mistreatment of horizontal derivatives by finite difference schemes. (see Orzag, 1974). Besides aliasing, other errors, which can be given the general name of coupling errors (Lilly, 1965), occcur as a consequence of the finite difference representation of non-linear differential operators. Numerous schemes can be designed to handle these non-linear interactions satisfying a number of integral constraints and keeping average errors to a minimum, but individual interactions will always remain somewhat misrepresented (because of the errors in computing horizontal derivatives). In 5.1.1 we already mentioned the possibility of systematic "linear" phase errors for the interactions between the quasi-stationary large-scale and the more transient small-scale components of the flow. We would like to emphasize here the importance of the more "non linear" and less systematic (in its effects), part of these coupling errors. The distinction that we are making between "linear" phase and "non-linear" coupling error is arbitrary and qualitative: it is based for a large part on our subjective assessment of the N48-T63 forecast differences and our desire to categorize them. Quantitatively, N48 "non-linear" coupling errors should lead to forecast differences of the same order of magnitude as the "linear" phase errors. Evidence for the presence of the latter should be sufficient to guarantee the presence of the former. It is well understood here that we are speaking in both cases of FDs which lead on average to improvements of T63 over N48. This point, along with examples, will be further developed in Section 5.3.

5.2 Systematic phase differences and N48 "linear" phase error

5.2.1 Introduction

It is fair to affirm that the most obvious and frequent type of differences observed in Section 3, at short range in particular, was related to the position of well-defined systems. Moreover, in the majority of cases N48 features were lagging behind those of T63. Well-known theoretical considerations (Section 5.1) suggest that this is due in particular to the presence in N48, and absence in T63, of "linear" advection phase errors, implied by the type of second order spatial finite different schemes used by N48.

Besides pure phase differences - along a cyclone track - , one also noted cross-track differences. Among the possible contributing factors, we would like to isolate two of them, directly related to the "linear" phase errors in N48. First, the wave patterns are not isotropic and therefore different phase errors arise in the EW and NS directions. Second, the N48 grid is not isotropic, and for storms moving north eastward at about 45°N the phase error will be about twice as large in the north south than in the east west direction, leading to a more southern track for N48. This is true, assuming that such phase errors are proportional to the square (for a second order scheme) of the physical grid length. The horizontal diffusion used tends to attenuate this by increasing the EW phase error, but its effect at short range on the synoptic waves is likely to be extremely small.

It is also worth mentioning that no such tendency was observed between T63 and coarse resolution spectral models (T40 or T53).

In complex atmospheric models there are many causes of phase errors, besides horizontal finite difference errors. These errors will be common to T63 and N48. For example, both models suffer in a similar way from incorrect initial conditions, parameterization errors and inaccuracies of the vertical and time

discretization schemes. However phase differences between the forecasts must be attributed to model differences. Our main task here only consists in isolating a specific type of difference and emphasizing its origin.

In this respect, we would like in this section to complement the synoptic results of Section 3, from the point of view of phase differences during the life cycle of a cyclone by showing examples of trajectories with important differences. Then, a few statistics demonstrating the systematic retardation effect of negative "linear" phase errors in N48 will be presented. Thirdly an impact study of these errors versus other errors will be made on two cases. Finally, the impact of phase differences on the relative models performances will be examined.

5.2.2 Cyclone trajectories

In Fig. 57, some cyclone trajectories are presented, most of them taken from cases already discussed in Section 3. Dots and crosses correspond to the cyclone centre position (at 1000 mb) at one-day intervals (the days are indicated). The choice was made subjectively from well defined quickly-moving cyclones that could be tracked easily for many days in both models from nearly identical initial positions (this is why most cyclone tracks start no later than D+2) and showing thereafter large phase and/or track difference. The number of tracks per region is in approximate proportion to the geographical distribution of observed differences. Most tracks came from autumn and spring cases. In winter the models more rarely show simple phase difference patterns. In summer, the cyclones are rather weak and cannot be followed for long periods of time. This does not mean that phase differences are not important in summer. The many cases shown in Section 3 point to the contrary.
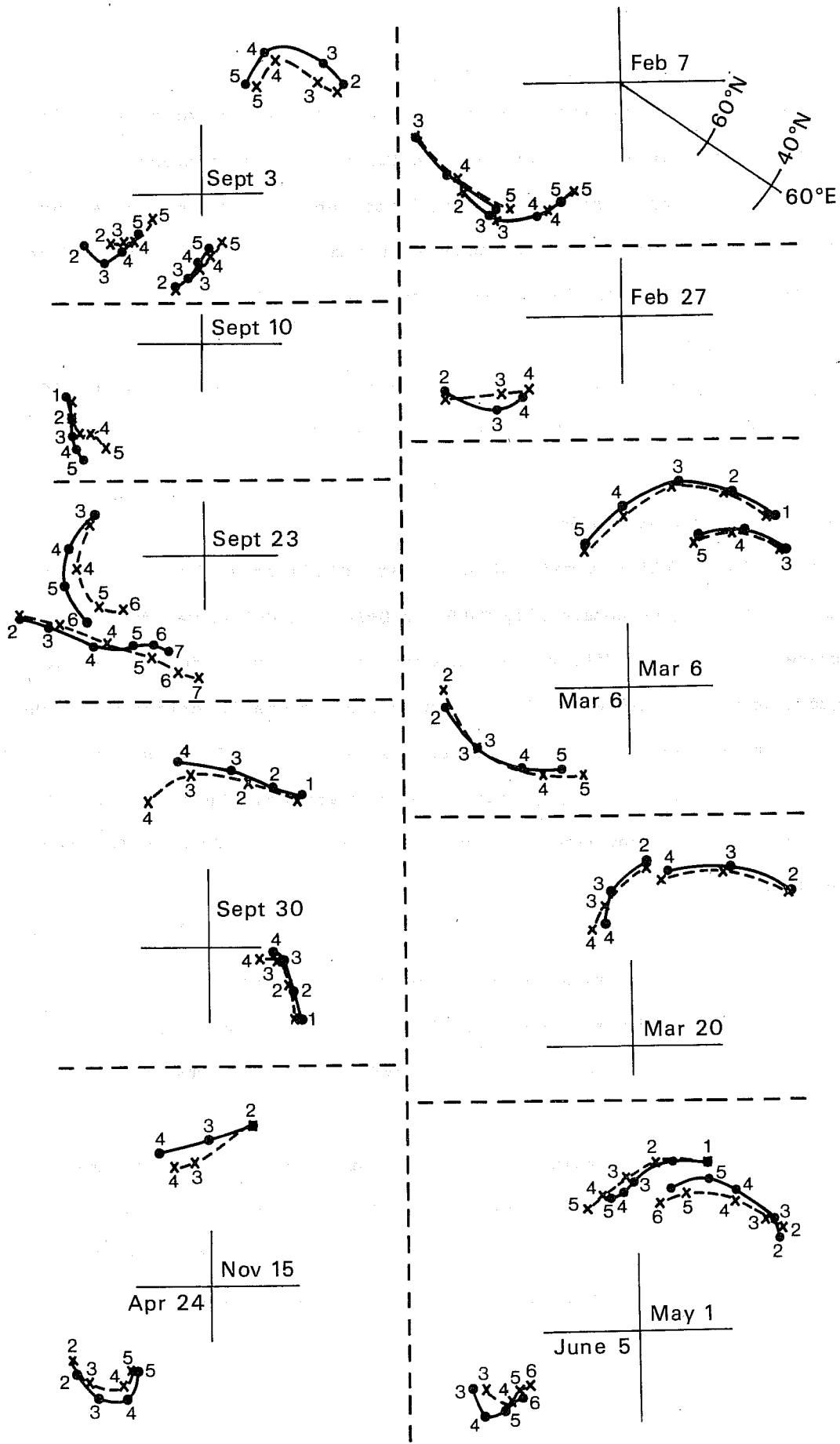
Fig. 57 A selection of cyclone tracks from various forecasts (N48= •——•) ;
T63= *---*). Crosses correspond to the days in the forecasts.

One characteristic of the tracks is that young cyclones move more quickly than mature ones. Phase differences are often largest at midtrack. Near the end, N48 cyclones appear sometimes to be "catching up". Most tracks curve northward. In a majority , N48 cyclones do not move as quickly northward and their tracks are on average to the south of those of T63.

Perhaps the best approximation to purely linear phase errors is depicted for the MAR 6 case (Figs. 36, 37 and further 58 and 59).

5.2.3. Phase error statistics

A statistically valid estimate of the phase differences between the two models was obtained by considering Z1000 on D+2. (On D+1 it was difficult to objectively measure the differences in practice.) We considered only well defined lows with a gradient sufficient to get a reasonable estimate of the centre of the low and that could unambiguously be identified as being the same in the forecasts and the verifying observations. In order to get a statistically significant number of such lows, we had to eliminate forecasts beyond D+3.

On D+2 we can expect a priori the "linear" phase error in N48 to be an important component of the overall phase error and assume the other components to have a statistically similar effect in N48 and T63.

In this way our sample contained 192 cases. For each of them we computed the observed eastward displacement between D+1 and D+2 and the resulting EW phase error (in degrees) on D+2 forecasts for both T63 and N48. We further classified the lows from slow eastward moving (or even retrogressing) to fast eastward moving lows, and converted the phase errors in terms of an equivalent distance error (in km) at 45°N.

The results are displayed in Table 5.

Table 5.   Phase error statistics

D	Number of lows considered	E(degrees)		L(km)	
		T63	N48	T63	N48
all lows	192	-.55	-.82	-43	-64
D < 5°	64	+.64	+1.02	50	80
5° ≤ D < 10°	39	+.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)

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

A few comments can be drawn from this table: the first one is that **for the 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 enough the cyclones in their decay stage and to have them penetrate too far into the land masses. The second one is that **for the fast moving systems both models are too slow but N48 substantially more than T63.**

A mean difference of about 120 km in the position of the fast moving cyclones in middle latitudes on D+2 is, though not very large, yet significant since these fast moving cyclones are generally also fast developing ones and the phase difference is often associated with an amplitude and track difference as shown above.
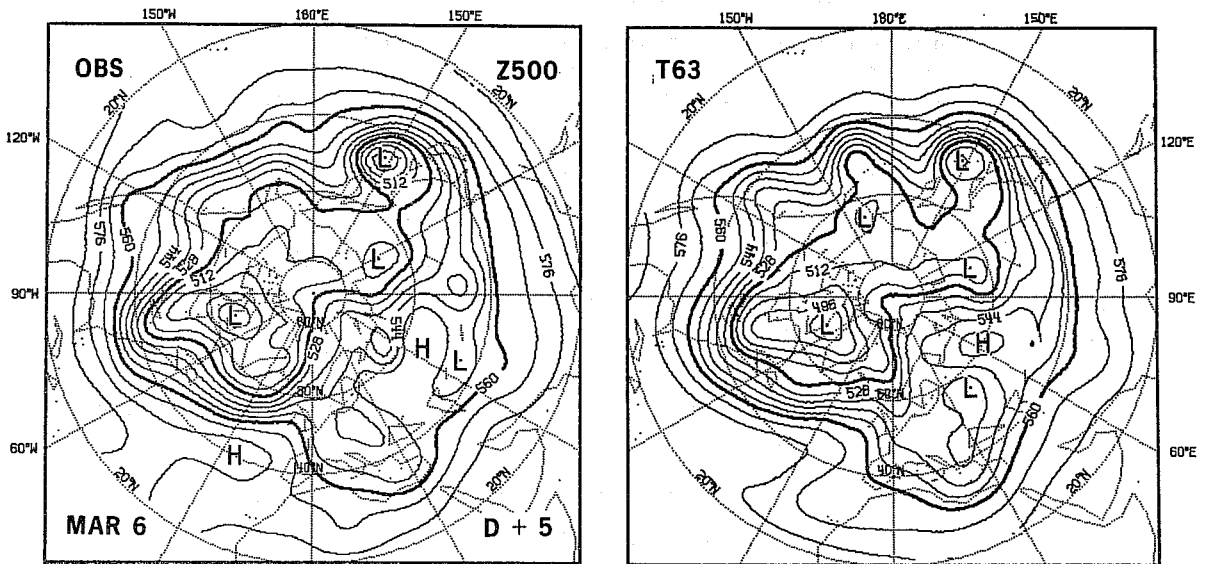
Fig. 58a   MAR 6, Z500, D+5
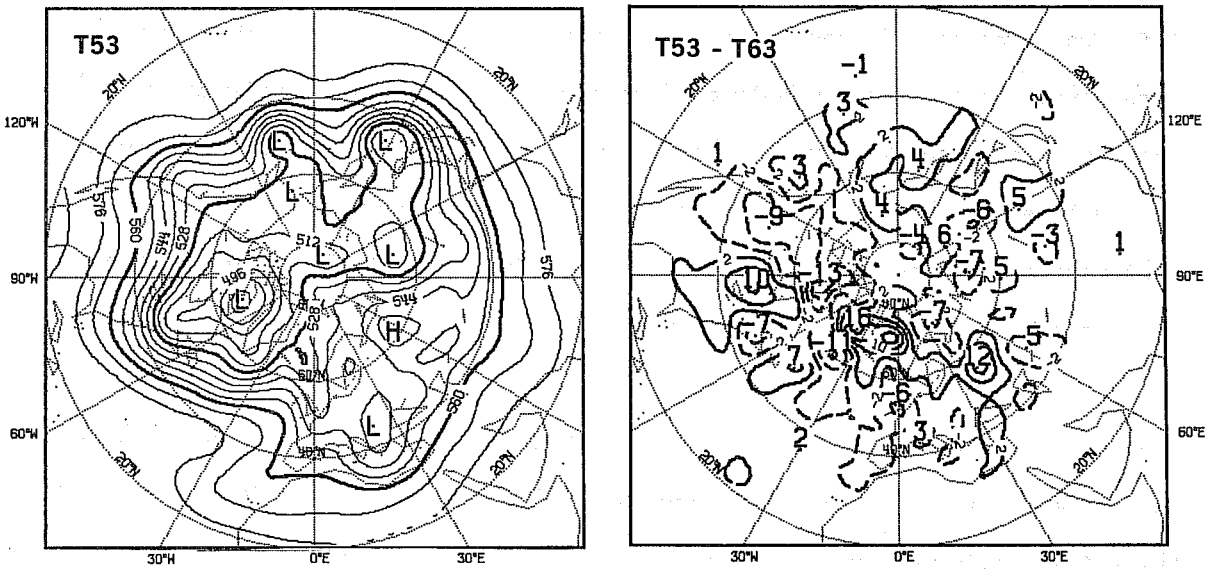
Fig. 58b   MAR 6, Z500, D+5 (T53, T53-T63, N48-T53)
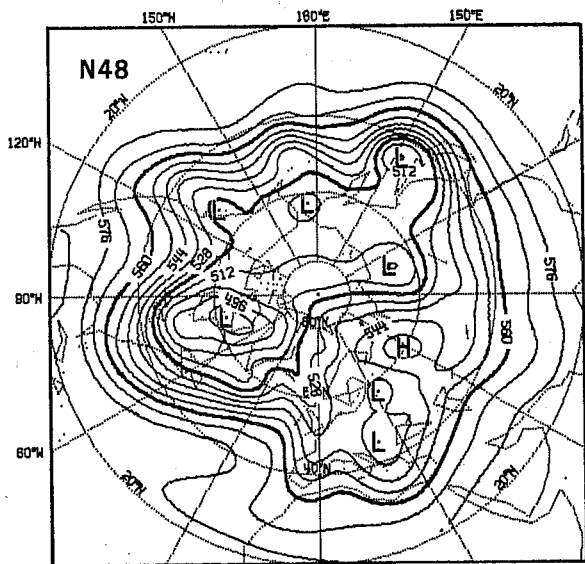
Fig. 58a (Cont.)
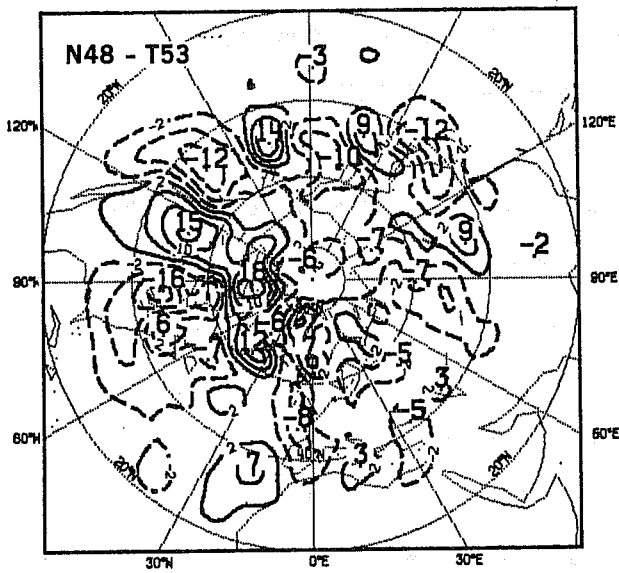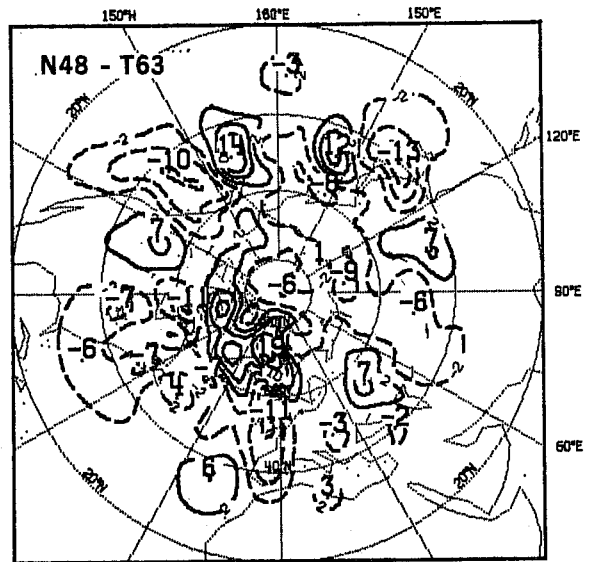
Fig. 58b (Cont.)

5.2.4 Case studies

In this Section we pursue further the discussion on MAR 6 and JUN 5 cases. These show extreme examples of what we believe is largely the result of N48 "linear" phase errors. To partially verify this hypothesis we also integrated lower resolution spectral models (T53 on MAR 6 and T40 on JUN 5). For JUN 5, we were also preoccupied by the fact that T63 and N48 did not use the same definition for their continental coast-lines and we performed a sensitivity test (experiment C48).

Case of MAR 6

Fig. 58a: Z500, D+5 Also, T53 forecast and differences T53-T63, N48-T53 are shown in the same figure.

Fig.58b: Z500, D+5, T63 and N48 forecasts with a simplified physics (and N48-T63 difference maps).

Fig.59: Z1000, D+5. And T53 forecast and differences as in the figure above. At both levels, the D+5 maps show a well-organized and definitely phase-like difference pattern between N48 and T63 along the 500 mb jet, from EAS to north ATL. When comparing the spectral models (T53 and T63) little phase differences are apparent over EAS and PAC. The main regions of difference, WAM and north ATL, have a more "non-linear" character. When comparing T53 and N48, the phase-like difference pattern is not only similar but perhaps even more convincing than the N48-T63 differences, with apparently less "non-linear" differences over WAM and north ATL. The phase difference pattern also includes typical NS (cross-track) differences for cyclones over EAS and north ATL.

In order to prove that the observed FDs arise mainly from the dynamical terms, it was decided to run two parallel experiments (one with T63 and one with N48) **with a limited physical package,** i.e. just surface friction and dry convective adjustment. At 500 mb, the D+5 T63 and N48 corresponding maps (Fig. 58b) present a number of differences with the ones displayed in

Fig.58a, but the difference pattern exhibits almost exactly the same

signature, especially along the 500 mb jet from EAS to WAM.

Over North ATL and EUR the forecasts with limited physics are very different

from the standard ones, so one cannot expect to have the same difference

patterns. However, some similarities can be observed, in particular along

the Greenwich Meridian.

Case of JUN 5

Fig. 60: Z1000, D+5. Also T40 forecast and differences T40-T63, T40-N48 and

C48 forecast (N48 with modified coast lines: see text below for further

information) and differences C48-T63, C48-N48 are shown.

For this case, we concentrate the discussion on D+5 for the North American

east-coast storm. We noted on D+3 and D+4 (Fig. 42) the considerable NS

phase difference on that system. It is maintained on D+5. Comparing the two

spectral models (T40 and T63), one notes differences of position and depth

but very much unlike N48-T63 differences. In fact, T40 predicts a better

position (further north) and depth (less deep) of the main depression than

either T63 and N48. The anticyclone to the SE is also better positioned.

C48 is an integration of N48 with modified coast lines (as has been

introduced in the operational version of N48, on APR 1st 1981 in closer

agreement with T63 coast lines). During the course of the present series of

experiments, coast lines in N48 were defined to be consistent with the

mountain field used then but later found to be inacccurate in many regions of

the globe. In particular, it was noticed that the North American east coast

thus defined was shifted unrealistically eastwards into the Atlantic. In

T63, because the mountain field is spectrally analysed there results
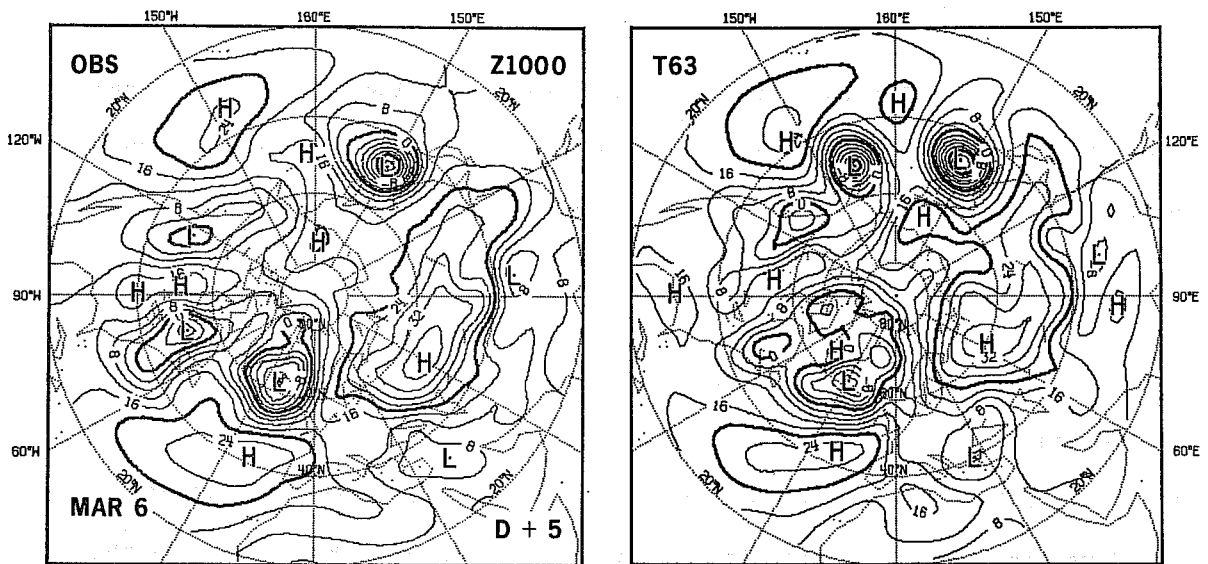
Fig. 59a   MAR 6, Z1000, D+5.
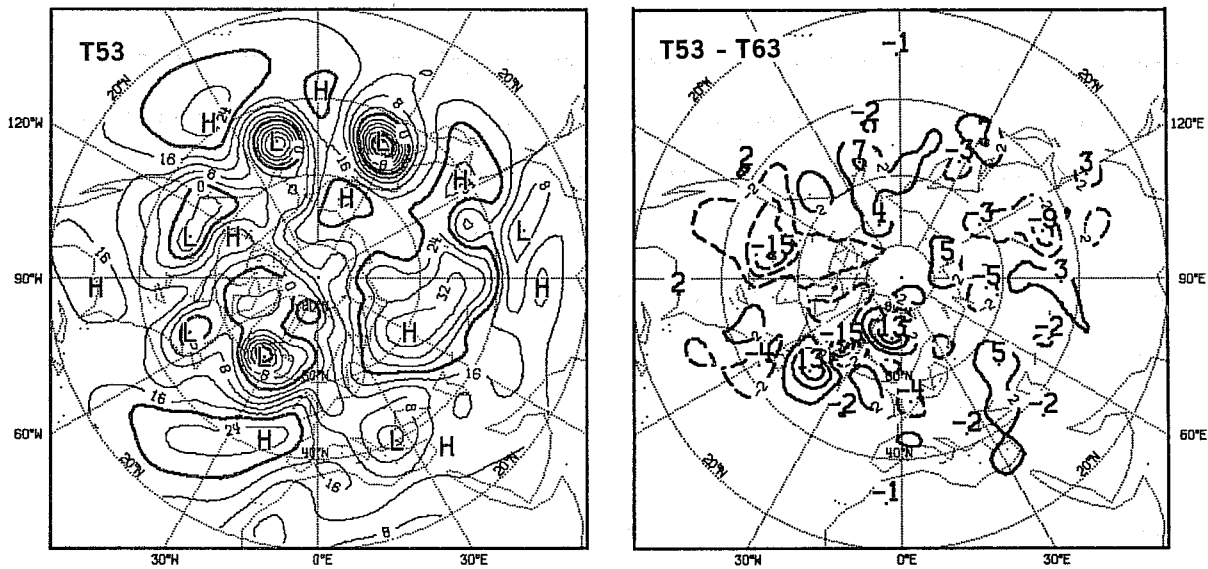
Fig. 59b   MAR 6, Z1000, D+5 (T53, T53-T63, N48-T53).
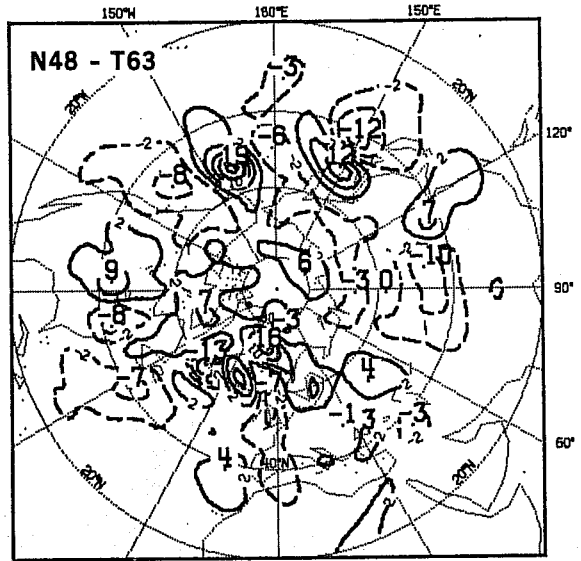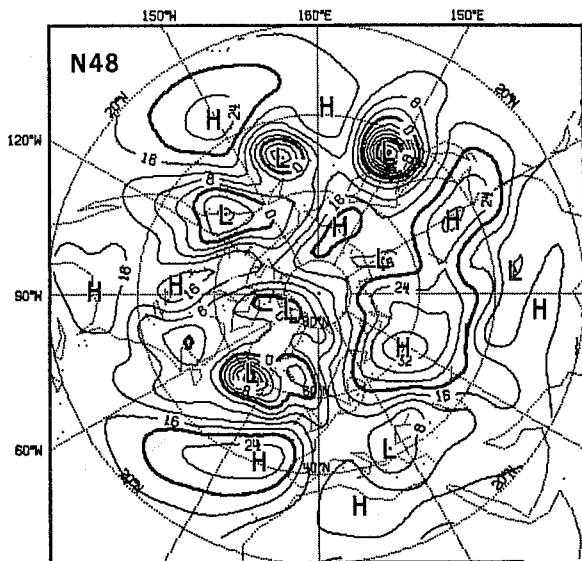
Fig. 59a (Cont.)
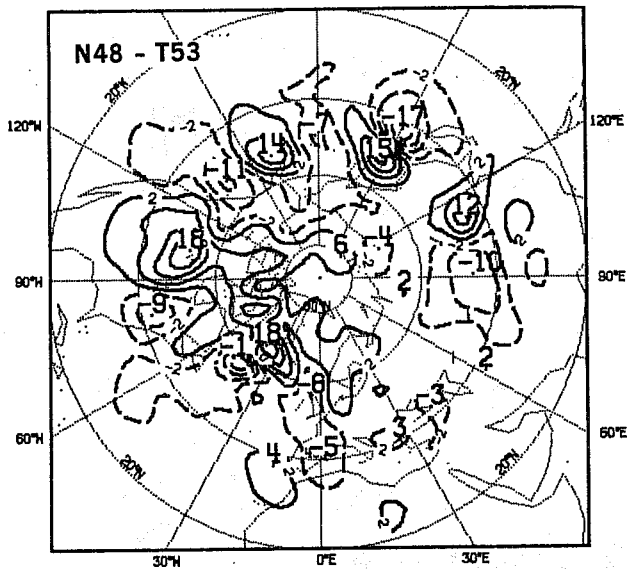
Fig. 59b (Cont.)
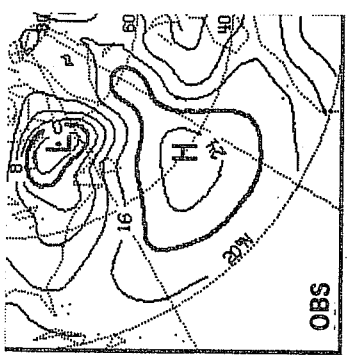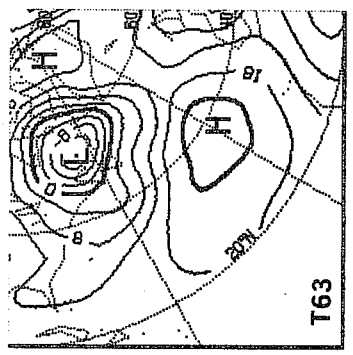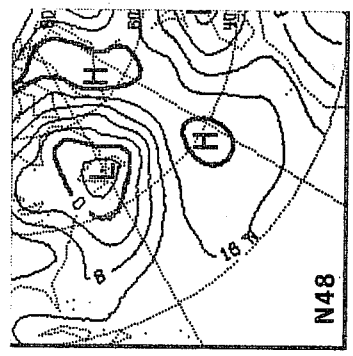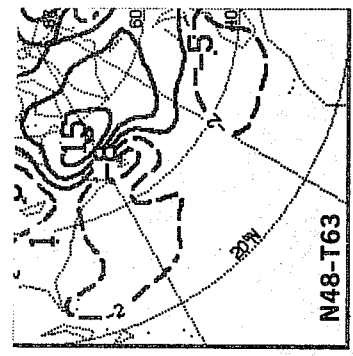
Fig. 60a JUN 5, Z1000, D+5.
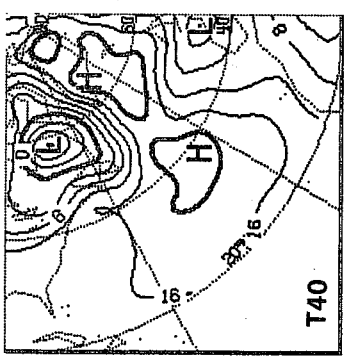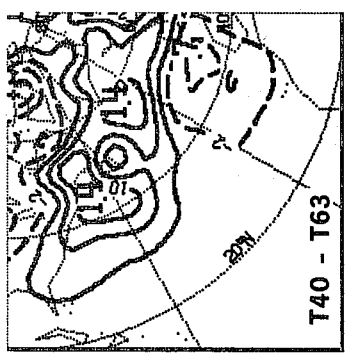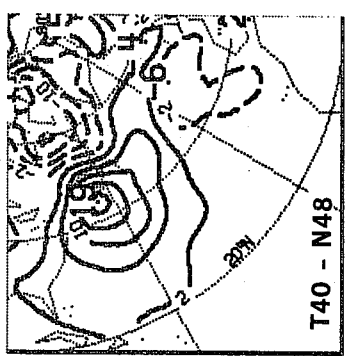
Fig. 60b JUN 5, Z1000, D+5 (T40, T40-T63, T40-N48).
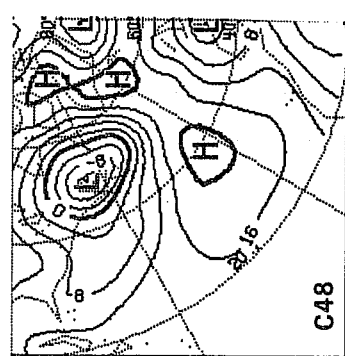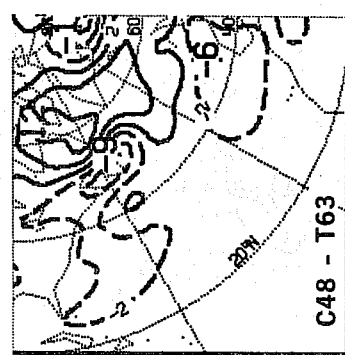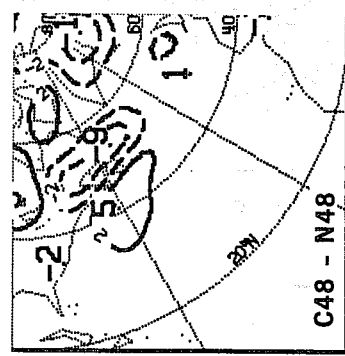
Fig. 60c JUN 5, Z1000, D+5 (C48, C48-T63, C48-N48).

non-vanishing terrain heights over sea. The above definition is not suitable and coast-lines in T63 are defined independently. The effect of coast-line differences was not thought to be important but the relatively large N48-T63 differences, early in the forecasts of this case, needed to be examined more thoroughly.

The results of C48 indeed show an impact of a shift of the east coast line of North America to a more correct position. The new low position is more correctly further inland but also incorrectly deeper. Both changes, not surprisingly, bring T63 and N48 in closer agreement. However, arguments on the cause of the north-south phase difference do not suffer from the change. The difference dipole remains well defined.

The impact of resolution on this particular storm is not negligible but also not especially in favour of the higher resolution models (AC scores, calculated hemispherically, Fig. 61, nevertheless show that T63 is better than T40 which performs as well as N48). Resolution and technique (linear phase error) differences had opposite effects with respect to the phase of the particular storm. Coast-line definition is one of the very few minor model differences. On this case, it had some impact.

5.2.5 Score improvements and phase differences

If "linear" phase differences constitute, as we are claiming, an important part of the total differences, and we have shown that this is true in many cases, they should also be responsible for an important part of the T63 improvement over N48. Can we reconcile this statement with the facts that the improvement in AC is larger in the long waves and larger for Z1000?

Let us examine the phase difference patterns in terms of scale and in relation to the vertical structure of atmospheric lows. First we must point out that the long waves in question are the planetary long waves (EW Fourier

133
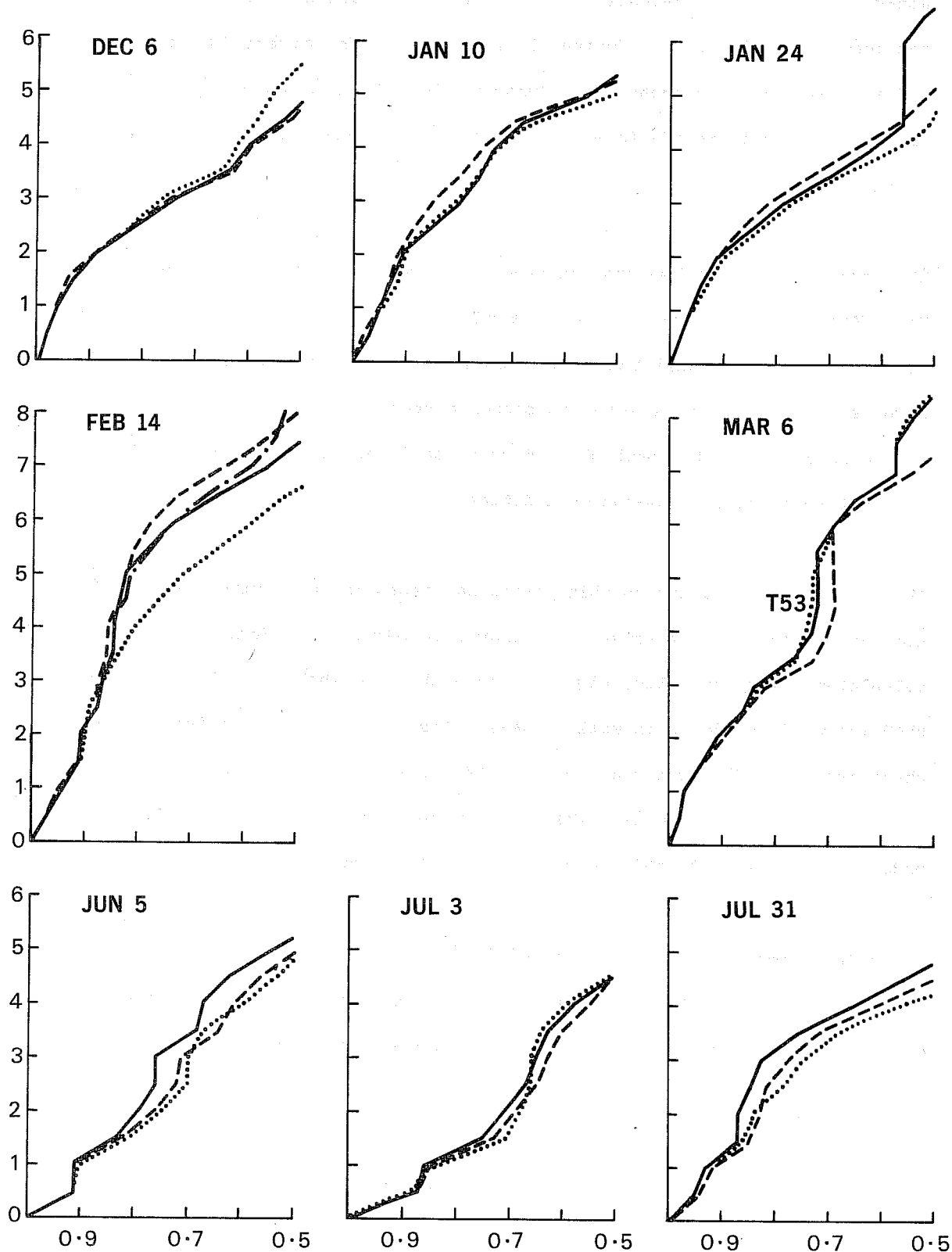
Z1000

Fig. 61 Predictability (days) in terms of AC averaged between 20°N and 82.5°N
for Z1000 on DEC 6, JAN 10, JAN 24, FEB 14, MAR 6, JUN 5, JUL 3 and
JUL 31.
(T63 = ———;   N48 = ———;   T40 (T53 on Mar 6) ......;
C48 (N48 with more correct coast lines on Feb 14) .—.—.—.)

decomposition) containing the full "north-south" spectra. We also observed the large differences to occur in specific geographical areas and to depend on the stage in the life cycle of storms, as well as on the forecast range. In general, phase differences are relatively small even where large amplitude difference dipoles develop. In other words, the differences are usually very local. The more local the differences are, the more widespread is the "wave packet" representing them spectrally. When, for example, in a fast moving low over the Atlantic a phase difference develops while the rest of the forecasts by T63 and N48 remains very similar, the difference pattern must have a relatively large wavenumber 1 component attached to it. Local and long wave differences are thus not necessarily inconsistent. (e.g. OCT 18, NOV 15, NOV 29, Figs.21, 24, 25). On the other hand the phase errors are only locally "linear". It in fact corresponds to a global non-linear problem. In other words, it is in part the very long waves that define the basic flow for the "linear" advection of shorter waves. "Linear" advection errors in the motion of a cyclone (particularly its NS component) is equivalent to coupling errors between very long and shorter waves.

Table 6 compares AC improvement (%) of T63 over N48 for the total height field, long (1-3) and medium (4-9) waves for some of the maps in Section 3 that showed large local (often phase) differences. We also give the figure numbers to allow for easy reference.

135

Table 6. AC improvement (%) of T63 over N48 for the height field (in some selected cases).

Dates	Forecast day	Z500			Z1000			Figures
		Total	1.3	4.9	Total	1.3	4.9	
SEP 30	4				5	8	4	20
OCT 18	4	10	15	9				21
OCT 25	4	5	7	-3				22
NOV 1	5	10	15	3	4	8	-10	23
NOV 15	3				3	3	1	24
NOV 29	3				3	9	1	25
MAR 6	4	2	0	6	3	1	3	36, 37
APR 3	4	4	3	9				38
APR 24	4	4	6	-3				39
MAY 1	4	5	6	6				40
JUN 5	4	2	6	-1	6	2	6	41, 42
JUL 17	4	2	1	5	5	2	11	43a, 43b
JUL 31	4	3	0	4				45
AUG 21	4	2	1	3	4	5	2	46, 47

This table needs some comment. OCT 18, D+4 may be interpreted as an extreme example of phase errors with a major impact on long waves, and as can be checked from Fig. 21 there is a clear imbalance in the difference pattern: most of the major FDs are concentrated over ATL and ARC, explaining why the long waves have been so much affected. MAR 6, D+4, on the other hand, is a more classical example of almost pure linear phase error with a well spread out and regular (about wavenumber 7) pattern of positive and negative differences. Most of the improvement occurs in medium waves.

Amongst the other cases with large unbalanced error patterns, associated with large improvement in the long waves, one can quote: NOV 1 (phase and amplitude differences), NOV 15 is almost exclusively phase differences), NOV 29 (phase and amplitude), APR 24 (mainly phase) and JUN 5 (structure differences).

For the JUL and AUG cases, no such imbalance can be seen, because the meteorological fields themselves are much less structured: there is comparatively less energy in the long waves and therefore the improvement in phase speed affects mostly the medium waves.

With regard to the vertical structure of phase-type differences we observe that they are remarkably similar at 1000 and 500 mb and often larger at 1000 mb. This is consistent with the constant phase difference brought about by advection phase errors in N48 relative to deep "solid" systems. On the other hand total variance is smaller at 1000 mb and thus in relative terms, differences are larger there. Therefore, not surprisingly the impact is also larger. Alternatively, anomalies are smaller at 1000 mb and differences are therefore more likely to lead to larger AC differences. A similar structure for the improvements due to increased resolution from T40 to T63 was noted in TR23.

Therefore, the qualitative evidence to the effect that (i) N48-T63 systematic phase differences are due to the presence of "linear" phase errors in N48 and that (ii) these account for a large part of the performance differences between the models appears overwhelming.

5.3 Other differences. N48 "non-linear" coupling errors

5.3.1. Introduction

One of the practical problems of comparing spectral and grid point atmospheric models is that of the choice of resolutions to use in either of them. The impact of this choice on the results is likely to be very much dependent on the accuracy of the initial conditions as well as the sophistication and similarity of the other aspects of the models.

Except for a few details, such as horizontal diffusion discussed in Appendix B and coast-line definition mentioned in Section 5.2 for the JUN 5 case and

again in the present section for the FEB 14 case, the other aspects of the models are identical. If we can satisfy ourselves that our choice of resolution is such that initial and model truncation error differences (including the interaction with the physical parameterization schemes) are small, then the other major FDs between N48 and T63, besides the systematic phase differences discussed in Section 5.2 are also related to grid-point technique errors. Moreover aliasing errors in N48 are likely to be small for the first few days of the forecasts at least. There is indeed no reason to assume that they will be much larger than the truncation error differences between the models. Then the only model differences to account for the observed FDs are N48 "non linear" coupling errors. In fact systematic phase differences are but one outstanding manifestation of the total coupling errors as mentioned in Section 5.1.

5.3.2 The resolution problem

N48 and T63 are comparatively high resolutions higher than those used so far for models in similar comparisons, such that truncation errors should here play a smaller role. The choice of T63 in relation to N48 was made empirically and for practical reasons. It can be said to minimize the truncation errors in the spectral model for a computer usage comparable to that of N48. As this usage is approximately proportional to the cube of horizontal resolution, higher resolution models quickly become much more expensive while lower resolution models become much cheaper (by 40% for T53 and 75% for T40) but with the risk then that truncation error difference to N48 become important preventing a meaningful comparison of techniques. When, for example, we compared in TR23, forecasts of N48 to those of the spectral model with resolutions T40 and T63, we found that T40 could compete successfully only up to D+2, D+3, suggesting that T40 truncation errors contributed substantially to its loss of skill at longer range. Unfortunately, in these experiments, resolution was reduced in both the model

and the initial data and it is not possible to distinguish the effects.

The problem of initial truncation also arises for the N48-T63 comparison
since the number of degrees of freedom in both analysis and N48 corresponds
to a spectral model with a higher resolution than T63. There is then an
initial difference between N48 and T63. To visualize its importance we may
ask what is the equivalent spectral resolution for a field given on the N48
grid. A perfect equivalence (no aliasing) is only possible if the grid-point
field is itself exactly expressible in terms of a truncated spherical
harmonic series. If this is the case, the rhomboidal truncation R96 provides
for a maximum number of independent degrees of freedom exactly equal to the
number of grid points. In particular R96 preserves the NS variation of
effective resolution in the N48 grid. On the other hand, the triangular
truncation T96 is sufficient to represent a N48 field with uniform resolution
over the sphere and it only requires half as many independent degrees of
freedom. The initial information content of wave components between T63 and
R96 is generally considered to be negligible (That of wave components between
T40 and T63 may not be so however). In order to avoid any ambiguity we will
define resolution according to the number of degrees of freedom used both for
defining the initial state and integrating the models.

According to that definition, N48 is very nearly equivalent in terms of
resolution to R96 and T63 thus has (slightly) larger truncation errors. On
the other hand, it has no other type of error due to horizontal
discretization (except for aliasing in cubic and higher order terms) and
judging from the models performance, this advantage more than compensates for
the former disadvantage. The difference N48-T63 can thus be defined as the
resultant of the grid-point technique errors N48-R96 minus the truncation
error differences T63-R96, providing no other model difference exist. Only
if the truncation error and other differences are small, we can safely assume
that the differences N48-T63 are representative of the grid-point technique

errors.

In the following we examine a few resolution experiments to see if the above assumption is reasonable for the models.

5.3.3 Resolution experiments

In Fig. 61, we compare the performance of N48 with T63 and either T53 or T40 integration for a sample of cases: DEC 6, JAN 10, JAN 24, FEB 14, MAR 6, JUN 5, JUL 3 and JUL 31, presenting the curves of P(AC) for Z1000. Each experiment has been motivated by the desire to investigate particular aspects of the differences, but not all interesting cases could be rerun (in particular AUT cases). Thus the sample is certainly biased. We shall nevertheless briefly discuss the results.

Case of DEC 6

For this case, we refer to Section 3.3.(and Fig. 26) where we noted the superiority of T40 over both T63 and N48. (As indicated in Fig. 61, T40 scored also better at 1000 mb). We mentioned that this was in all probability due to more (T40 shows the same erroneous tendency but with less intensity) excessive depth in both high resolution forecasts of the features over WAM, south of Greenland and over Russia. Phase differences between N48 and T63 on all three features were also visible but had no impact on their performance possibly because the amplitude errors were already so large in both models.

Case of JAN 10

Additional Fig. 62, Z1000,D+5, including T40 forecast and T40-T63, N48-T40 differences.

The D+2, D+3, D+4 forecasts of Z500 and Z1000 by T63, N48 and T40 have already been shown (Figs. 27 to 30) and discussed in Section 3.3. We noted the large FDs over ATL, EUR and ARC between T63 and N48 and mentioned that the structures in T40 and N48 were much more similar and closer to OBS, particularly over ATL. At 1000 mb, on D+5, the poor quality of the forecasts over the same regions makes no doubt. Over ATL, the forecasts are in fact too poor to warrant a discussion. Over ARC, one notes the presence in the models of an erroneous anticyclone along 30°W, particularly intense and unrealistically steep in T63. Elsewhere, the forecasts are characterized by overestimation of the cyclonic circulation over northern Europe coupled with its underestimation over the Mediterranean. The scale of the low is particularly excessive in T63, still too large in N48 and about right in T40 which provides the most realistic forecast. Nevertheless, the similarity of N48 and T40 forecasts in constrast to their common large scale differences to T63's remains the most striking feature of this case. The most important smaller scale differences between them act in fact to increase the divergence of T40 from T63 permitting N48 to stay closer to T63.

Case of JAN 24

This forecast is poor at 1000 mb for all models. N48 does slightly better objectively than either T63 and T40. The forecast differences (not shown) are not anywhere very large and it is very difficult to interpret them.

Case of FEB 14

In the present state of the art, these forecasts by N48 and T63 must be considered as outstanding. We note a large improvement of T63 over T40 with a large gap in AC already by D+5. It must be pointed out that in terms of its own mean performance, the forecast by T40 is also above average and that
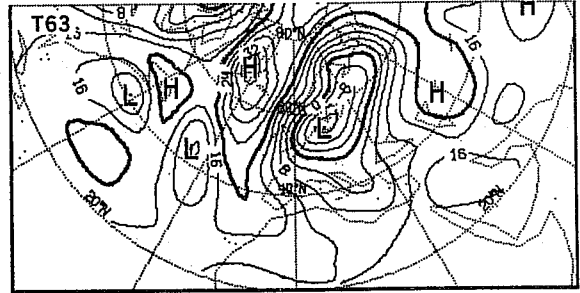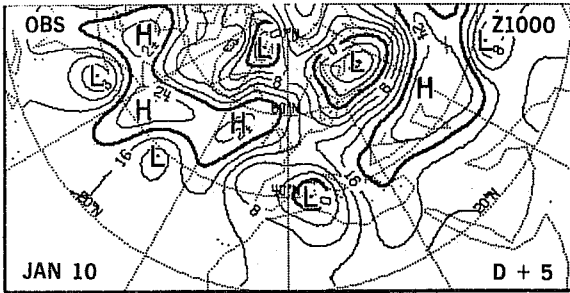
Fig. 62a   JAN 10, Z1000, D+5.
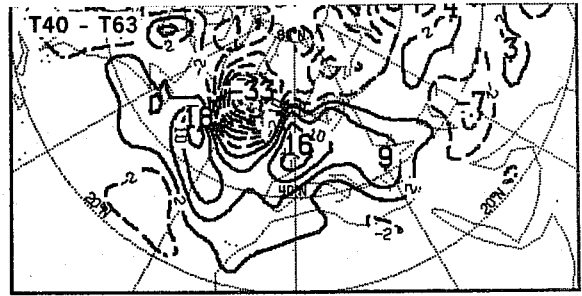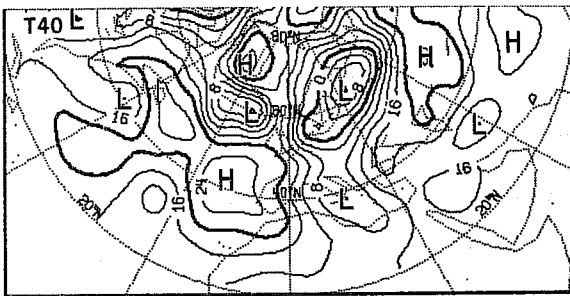
Fig. 62b   JAN 10, Z1000, D+5 (T40, T40-T63, T40-N48)
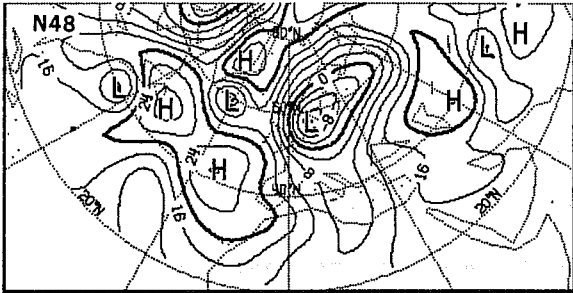
Fig. 62a (Cont.)
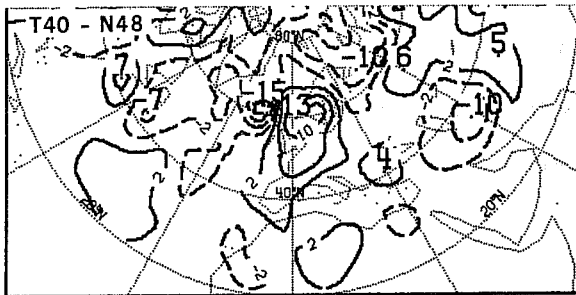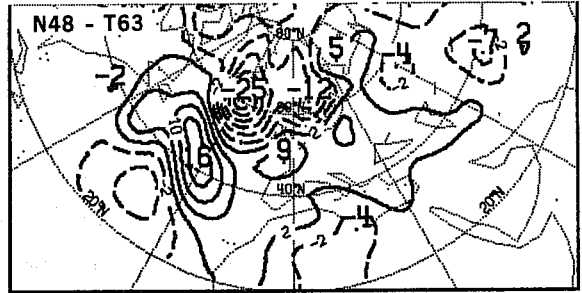
Fig. 62b (Cont.)

if T40 does often almost as well and sometimes even better than T63, it does so on cases where all models have generally average and often low skill.

No exceptional T40 forecast was found so far. In particular P(60%) was never found to be larger than 7 days. In other words T40 can probably not be expected to be useful beyond D+7, which means that truncation errors are likely to become predominant around D+7 for T40 and confirms that they must be usually small for T63 up to D+10. However, comparing N48 and T63 on the FEB 14 case, one observes that both models perform equally well up to D+5 but that N48 maintains a higher score thereafter. This case is unique from the present series of experiments (including those in TR23). Could it be that the resultant errors being smaller than usual in this case, we are observing the effect of lower truncation errors in N48?

On Fig. 63 we show the D+6 analysis and forecasts by T63, N48 and C48 (described below) of Z1000. Note first that the forecasts are remarkably successful on the large scale for such a range. They are also very similar except over EUR where N48 correctly predicts anticyclonic flow. In T63, the flow is mostly cyclonic with a large low centred near the Black Sea. In Appendix B, we discuss the effect of horizontal diffusion on the average rate of large scale latent heat release in T63 and the choice of the FEB 14 case for the experiments was not a fortuitous one. We in fact observed (not shown) a sensitivity to diffusion for the structure and intensity of the low systems in T63 near the Black Sea in this case. But the major features of the circulation over EUR were not substantially altered by a change in K. Now, the only other known model difference, to which we are aware (Section 5.2) of model sensitivity is the coast line definition. Indeed when comparing the coast lines of the Mediterranean, we noted a much smaller area of sea in N48 old coast-line definition and by integrating N48 with improved coast lines (experiment C48, lower right panels of Fig. 63), the grid point
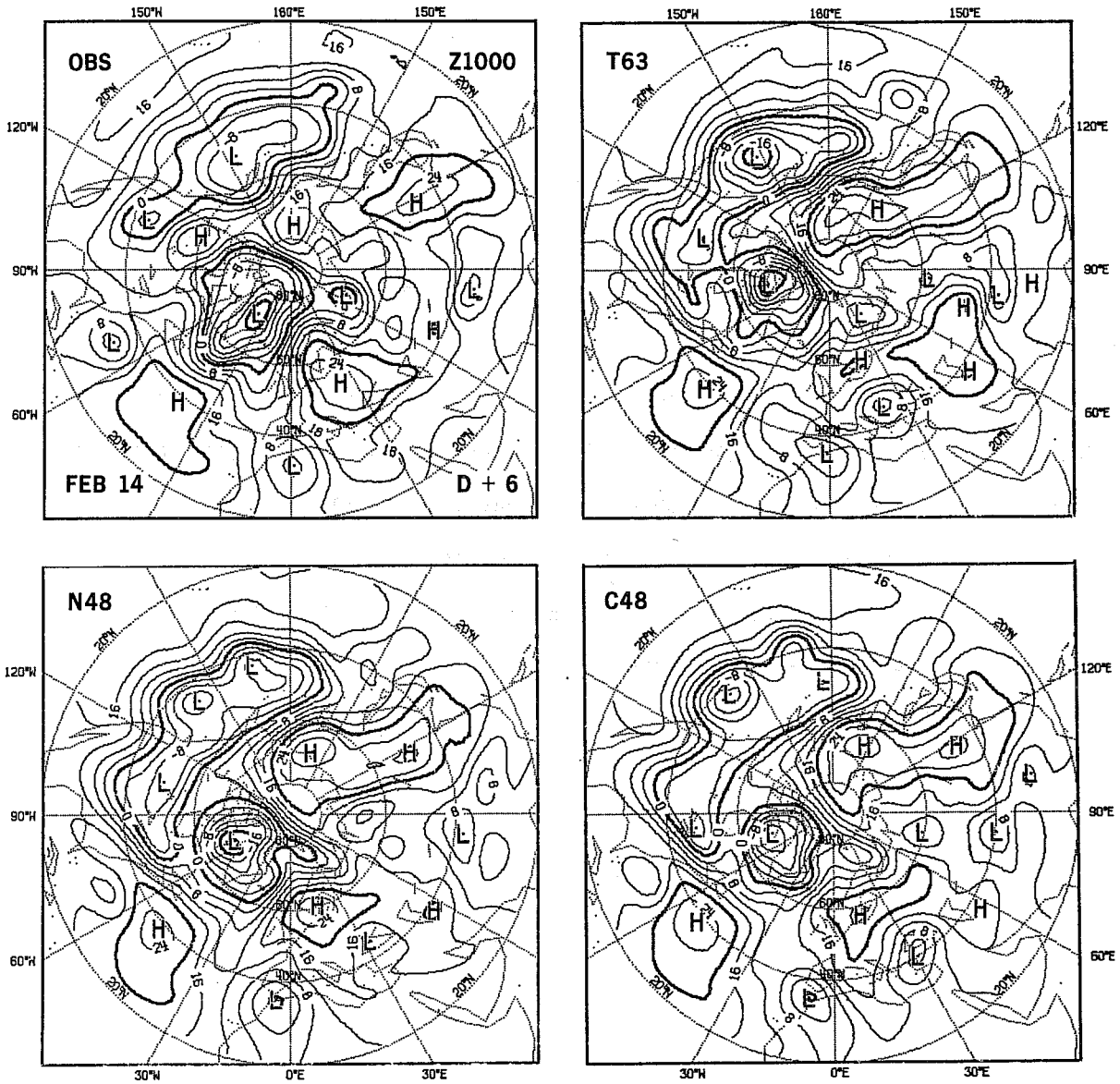
Fig. 63 FEB 14, Z1000, D+6 (OBS, T63, N48, C48)

and spectral results became more similar over EUR (the low and near the Black Sea is much deeper in C48 particular and the Z1000 large AC difference in favour of N48 has essentially vanished (Fig. 61).

Case of MAR 6 and JUN 5

We extensively discussed these cases in Section 3 and again in Section 5.2. Most of the improvements of T63 and of T53 over N48 on MAR 6 were related to phase differences (Figs. 36, 37 and 58, 59). In regions where T63 and T53 differed, N48 structures appeared more similar to T53's once phase differences were taken into account. On JUN 5, we emphasized a difference between N48 and T63 in the position of the low on the east coast of North America (Figs. 41 and 42) and showed that T40 predicted quite a different structure for the same low (Fig. 60). The predictability curves (Fig. 61) show that T53 performed as well as T63 and better than N48 on MAR 6 and that T40 performed as well as N48 on JUN 5.

Case of JUL 3

Although T63 performs better than N48 up to D+4, its advantage over T40 is less well marked after D+2. No maps are shown on this case.

Case of JUL 31

The forecasts are good for the summer season. The scores of the spectral models are well ordered according to resolution. N48 produces the worst forecast on D+1 and D+2. Later it is better than T40, becoming as good as T53 by D+4 but remaining inferior to T63. Differences between all models were particularly marked for the low over ARC (Fig. 45) which appeared very sensitive to spectral resolution. In particular its shape and position in T53 (not shown) were more similar to N48 than to T63.

In summary the truncation error differences between N48 and T63 (which in principle should favour N48) do not seem to play a major role in N48-T63 FDs at least up to D+5. At longer range it is possible to imagine that large FDs are due to truncation error, aliasing or other model differences (e.g. coast lines as shown on FEB 14 case). Furthermore, besides N48 "linear" phase errors which are the most frequent cause of FDs, the other differences,up to D+5 whether or not they are in favour of T63 often appear as a lack of effective N48 resolution as seen from the few examples discussed here and other cases from TR23.

5.3.4 N48 "non-linear" coupling errors

In Section 5.2 we related the observed systematic phase differences to N48 "linear" phase errors arguing on the strength of linear theory. From the very structure of the differences, there is no doubt however about the net non-linear character of the differences, since locally linear phase errors in the displacement of one or two large meteorological systems over the hemisphere indirectly affect the very large scale, as is well demonstrated by score improvements concentrated on average in the long waves. If we now look at cases with large FDs which cannot easily be classified amongst cases of systematic phase differences, we note that these occur on situations where the advecting current changes very quickly or abruptly in strength and/or direction. In these situations, the non-linear advection terms in the dynamical equations are not particularly smaller than they are in situations where systematic phase differences may be observed. N48 coupling errors must also be as large but rapid changes of strength and direction of the advecting flow make the errors more non-linear and prevent us from observing systematic effects.

Here we do not wish to present any additional maps. Rather we refer the reader to all previous examples.

SEP 30, OCT 18, OCT 25, NOV 1 (Figs. 19 to 23) and JAN 10 (Figs. 27 to 30)

show such examples of large local differences none with a clear cut

systematic phase character.

Other cases, such as MAR 6, show both types: systematic phase differences in

places of strong constant flow and other differences in regions of fast

changing flow. In the cases for which we also performed resolution

experiments, sensitivity to resolution was generally found to be larger in

these same regions (JAN 10, MAR 6, JUN 5). It is this type of results which

gives the impression that T63 has more effective resolution since both

spectral truncation and N48 "non-linear" coupling errors often have similar

net effects. Table 7 summarizes the main FDs in each case.

Table 7. Examples of large local differences with no clear cut systematic phase differences

SEP 30	Quasi-stationary long wave pattern at 500 mb. FDs are mostly in terms of amplitude (D+2, D+4). T63 preserves the intensity of the large scale troughs better (clear on D+2).
OCT 18	Rapid modification of the long wave pattern over ATL and EUR. T63 predicts a more correct structure on D+4.
OCT 25	A blocking high over northern EUR and Russia. T63 predicts a much better position of the high on D+4.
NOV 1	Large FDs over EAS and PAC. Phase and structure of large scale flow is more correct in T63.
JAN 10	Large modification of the jet structure across ATL at 500 mb between D+2 and D+4. Large FDs but N48 solution is more similar to that of T40 even though it scores better.

5.3.5 Differences in temperature forecasts

We have seen in Section 2 (Fig. 9) that T63 systematically (on a seasonal

basis) obtained better AC(T) scores than N48, near the tropopause level (300

mb). This is an important result. However, in the bulk of this report, we

more or less deliberately chose to concentrate our attention on height FDs.

To our defence, we must first say that no systematic improvement in T was

observed in the previous set of experiments (TR23) and we did not expect it

here (indeed the effect is least apparent during WIN). It must also be

admitted that it is much more difficult to study temperature FDs than height

FDs: temperature fields having more variance in the smaller scale,

temperature forecasts are less skillful. We believe however that upper

atmosphere temperature FDs are not fundamentally different in both character

and origin from height FDs.

To illustrate these points we show maps of T300, but also Z300, on D+4 of JUL

31 (Fig. 64, in addition to Z500 , Fig. 45). It is a case with large AC(T)

as well as AC(Z) differences in favour of T63 in upper levels in particular

(Fig. 65). In spite of that, when looking at the T difference map, we find

it difficult to diagnose T63 improvements, let alone interpret them. In

contrast,height FDs show a remarkable homogeneity of structure at all levels.

But it is reasonable to assume that temperature FDs derive from the same N48

"non linear" as well as"linear" coupling errors. These are enhanced near the

tropopause because of the artificially large gradients in the models'

$\sigma-$ coordinate. Recently Simmons and Strufing (1981) integrated a version of

N48 with a modified vertical coordinate system (hybrid between σ and

p-systems) in which horizontal T gradients are appreciably smaller near the

tropopause. Improvement of AC(T) were obtained in this region, which would

tend to confirm our assertions.

As for Z the T FDs are mainly transient, (especially up to D+5), although it

is worth noting that at longer range the largest differences are located near

the largest mountain masses as illustrated in Fig. 66 which displays D+9

N48-T63 differences in T300 for the same JUL 31 case.

We must also point out that in the case of T, the anomalies (compared to

climate) are minimum near the tropopause and AC are thus more sensitive
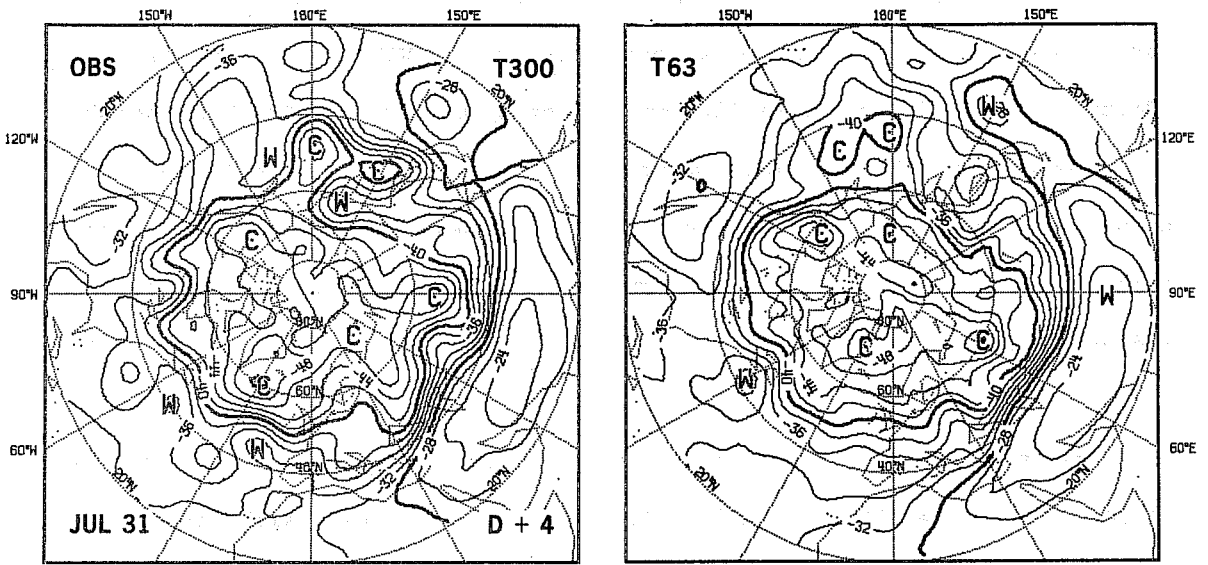
Fig. 64a   JUL 31, T300, D+4.
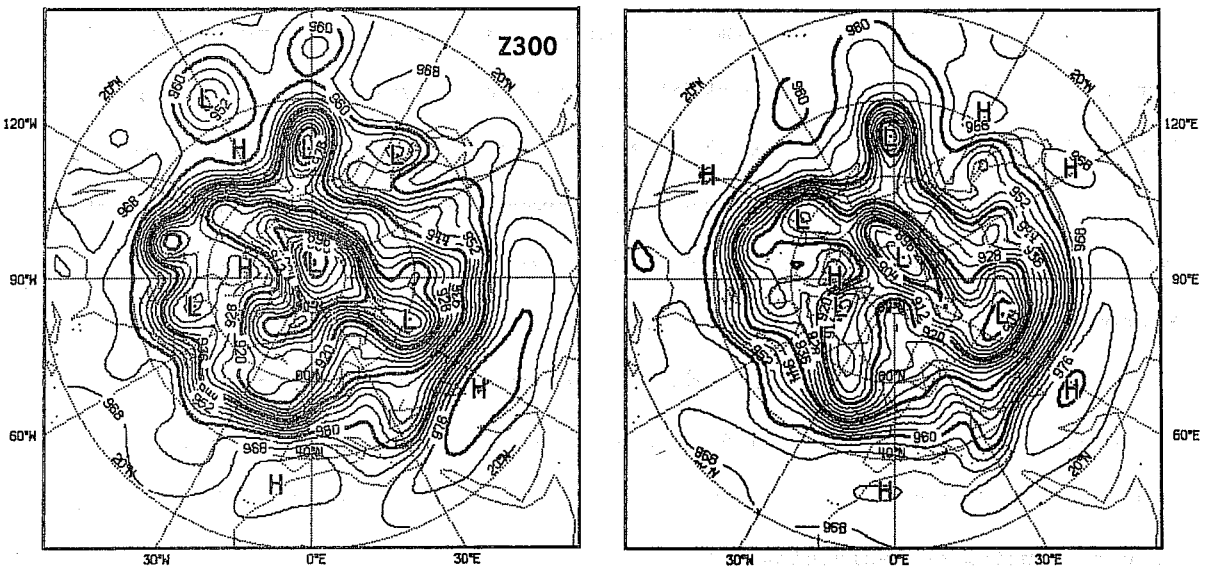
Fig. 64b   JUL 31, Z300, D+4
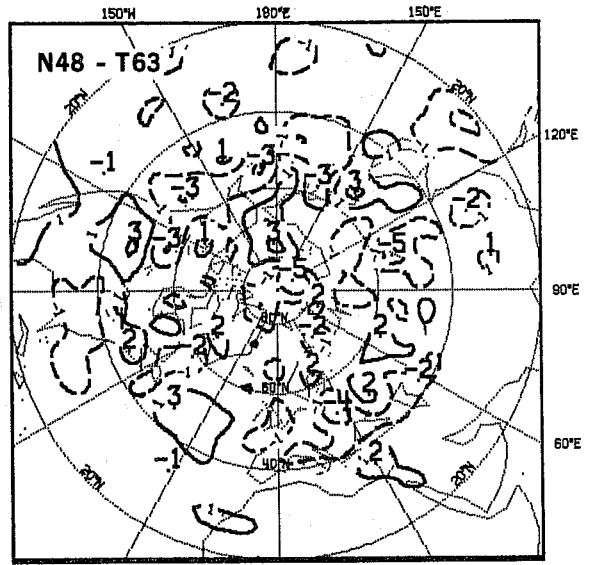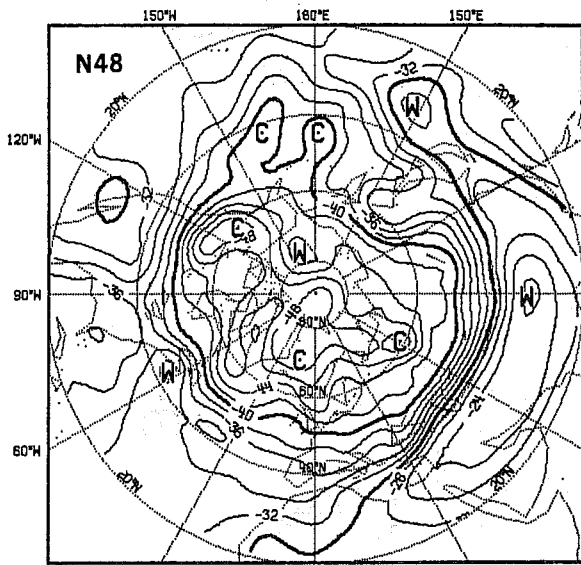
Fig. 64a (Cont.)
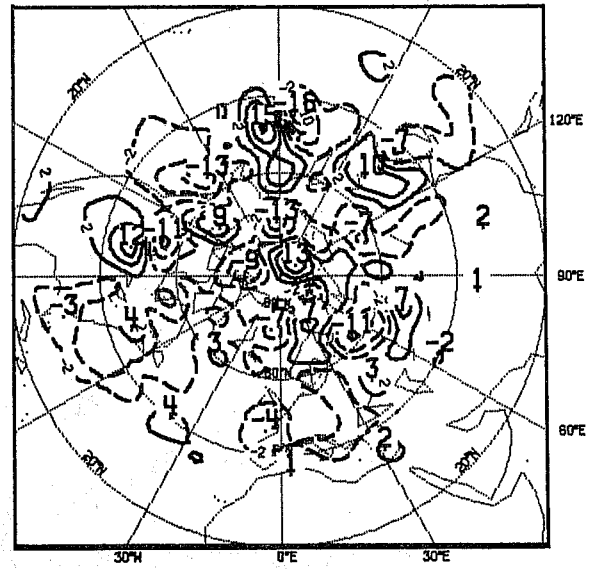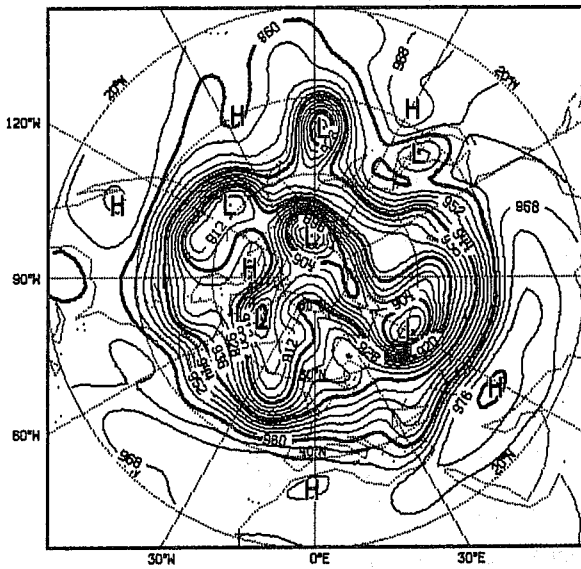
Fig. 64b (Cont.)
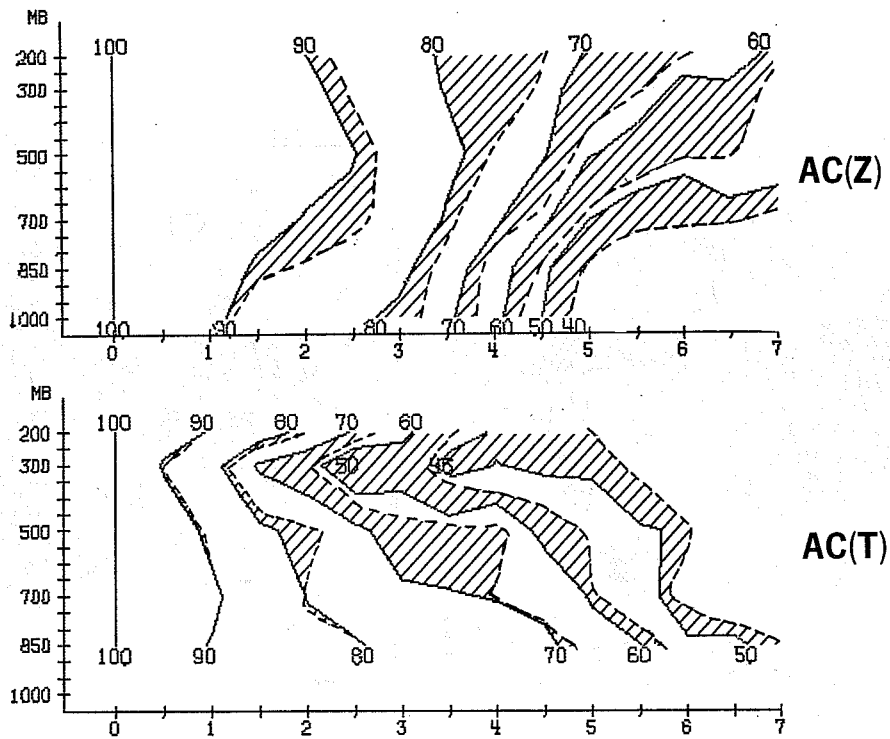
Fig. 65   Pressure-time cross-sections of AC(Z) and AC(T) for JUL 31
forecasts (N48= ———; T63= ----).
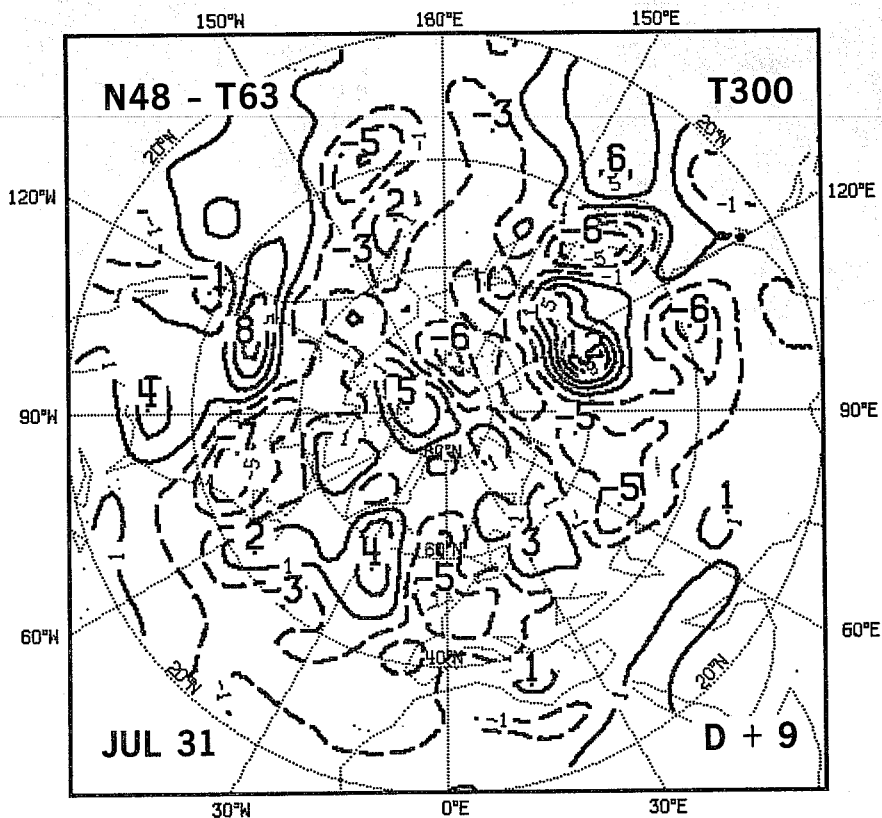
Fig. 66   JUL 31, T300, D+9 (N48-T63).

there.

Finally we mentioned in Section 5.3.4 that N48 height FDs with T63 were often similar to a lack of effective solution. The same was observed here for T300 for cases integrated also with T40. (results not shown).

6. CONCLUSIONS AND COMMENTS

The medium range forecasting ability of the ECMWF spectral T63 model has been demonstrated. The results show that both Z and T forecasts by T63 were superior to those of N48. On average, this was true at all atmospheric levels with a larger impact, in terms of AC, near the surface for Z and near the tropopause for T. It was also true at all scales but predominantly for the long waves and at all forecast ranges with benefits increasing up to D+4, D+5 and then decreasing to become statistically insignificant beyond D+7.

T63 mean objective improvements were arguably small. For Z1000, for example, predictability P(60%) was increased by about 6 h on average. They derived from generally small FDs,none the less of significant meteorological interest and almost systematically in favour of T63. Thus less than 20% of N48 forecasts showed higher predictability (P(60%) for Z1000) than T63 by more than 3 h while the contrary was true for more than 60% of T63 forecasts, irrespective of the season, the meteorological situation (the size of the improvement may however depend on it) or the mean forecast performance of the models.

The FDs were small in spite of the complexity of the models, because model differences were small. Physical parameterizations were nearly identical. Exceptions were the definition of coast-lines, clearly an oversight, and the value of the horizontal diffusion coefficient whose influence was at first underestimated (in view of TR23 results). Neither had an important impact on Z forecasts up to D+5 at least. The effect of diffusion only became obvious

on mean T forecasts at relatively long range and had little impact on mean Z forecasts. Other differences were minimized by using the same procedures for model data pre- and post-processing. At the beginning of the comparison, however, the post-processing procedures were not identical and this was shown to generate important biases in favour of N48, but these only affected part of the scores.

The impact of differences in numerical techniques has also been kept small by using a relatively high horizontal resolution in both models and, therefore, working in an area of convergence of the numerical solutions. A fortiori the resolution of the initial conditions and of the models cannot have played a significant role. Similarly aliasing errors in N48 cannot have become an important source of differences before several days of integration, unless the noise level in N48 was very large. The main source of FDs must therefore have been pure technique differences, that is, errors in the calculation of non-linear terms in N48 (coupling errors), the principal component of which must be due to the finite difference estimation of horizontal derivatives. Physical parameterization involves highly non-linear terms but since these are free from horizontal derivatives only truncation error and aliasing differences can be generated by them. This only leaves non-linear advection terms which are in any case usually larger as the main cause of differences.

The type of FDs observed on daily Z maps tends to confirm this. The most common type of Z differences was characterized by a phase-lag of N48 systems with respect to those of T63. They occurred preferably and in a more pronounced way when and where strong steady rectilinear jets (Z gradients) were present. In reference to the mathematical analysis which describes it we have called this type of difference N48 "linear" (meaning quasi-linear) phase error. Systematic phase differences were also usually small, amounting to mean displacement differences of about 60 km/day on the fast moving

systems. The actual magnitude of the local Z differences depended largely on the life history of the individual cyclones: for cyclones moving relatively quickly and steadily in their early stages, differences after 5 days could reach a few hundred meters of either signs. Systematic track differences, typified by a more southern trajectory in N48 for the lows moving in a SW-NE direction, were also observed. (The anisotropy of the phase error being one of the most likely contributing factors).

Non-systematic FDs also developed in and outside regions of rectilinear jets. For small FDs, it is difficult to determine unambiguously any particular cause. It is however particularly characteristic that large FDs appeared on many occasions where there occurred larger than usual intensity or direction changes in either space or time of the large scale flow pattern. On any given situation, it would be hazardous, even then, to single out a cause for FDs. But the fact that in most cases T63 performed better, the fact that N48 often behaved on a local basis like a spectral model of resolution lower than T63 and thus did not really provide for a different solution, the fact that errors in the same non-linear advection terms are already responsible for systematic differences of similar magnitude in certain types of situations leads us to believe that they may also produce the non-systematic differences.

If we consider individual wave-to-wave interactions and compare the exact and finite difference formulations (Lilly, 1965) of the non-linear advection terms one notes that errors are such as to underestimate these interactions in the same manner and for the same reason that phase speeds are underestimated. For very high wave numbers, the interaction cofficients nearly vanish. It is thus not surprising that spectral models can compete with grid-point models of higher resolution. But the FDs appearing early in the forecasts must remain due to errors in still relatively low (although not necessarily as low as for phase differences) wavenumber interactions.

Non-systematic FDs are thus also very likely to be generated by the non-linear advection terms and we have attributed N48 "non-linear" coupling errors (to distinguish them from the more linear phase errors) as the cause of a large part of them, particularly in situations of large (rapid or local) changes in the flow.

It therefore appears that the superiority of T63 over N48 is mainly due to technique differences, in particular its more accurate treatment of non-linear advection, the errors in N48 coming in large part from the second order finite difference estimation of derivatives.

At this point we would like to make a few closing comments. The first concerns the design of the experiments and relates to the problem of making meaningful comparisons of complex atmospheric forecasting models. Comparisons between experimental and operational models are relatively common practice in numerical weather prediction. The models compared usually differ in many aspects while their difference in performance is never very large. A clear indication from the present series of experiments is that great care must be taken to ensure that differences in performance are real and that genuine forecast improvements are not spoiled or hidden by biases and inconsistencies. It is worth emphasizing again that in the beginning of the comparison differences in post-processing alone accounted for important score differences. (Appendix A).

The spin-up and horizontal diffusion problems are also certainly worth recalling. The first one refers to initial imbalances leading to fictitious forcing. In the present case, a large thermal imbalance has been diagnosed which caused mean and, without doubt also, large scale differential cooling of the models with respect to the atmosphere. A large sensitivity of latent heat release to lateral diffusion has also been shown to exist in the present

models. (Appendix B). Both problems point to deficiencies of the present ECMWF system, deficiencies which must have significant impact on medium range weather forecasting.

Amongst the other numerical forecast problems common to both models, we have also noted negative phase errors and underdevelopment of young and fast moving lows and positive phase errors and overdevelopment of the more mature systems.

Finally, in reference to technique differences, it is reassuring to note that, at high resolution, different numerical solutions tend to converge and that differences can still be explained simply. For instance, N48 finite difference formulation is second order only. The deficiencies of second order schemes have been extensively studied in the literature and the use of higher order schemes has been advocated by many authors. In practice, however, an approximate equivalence is admitted between a higher resolution model with a second order scheme and a lower resolution model with a higher order scheme. The first option is often taken for simplicity. A spectral model corresponds rather to the second option (infinite order scheme) but without the inconveniences, since it is simpler and more efficient for integrations over the sphere with equivalent results as demonstrated here.

References

Basic model documentation

Baede, A.P.M., Jarraud, M., Cubasch, U. 1979 Adiabatic formulation and organisation 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.

Temperton, C., Williamson, D.L. 1979 Normal mode initialisation for a multi level grid point model. ECMWF Tech.Rep.No.11, pp.91.

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

Other references

Bass, A., Orszag, S.A. 1976 Spectral modeling of atmospheric flows and turbulent diffusion. Available from Environmental Sciences Research Laboratory. Office of Research and Development, US Environmental Protection Agency. Research Triangle Park, North Carolina 2771, USA.

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

Deron, E., Hollingsworth, A., Hoskins, B.J., Simmons, A.J. 1974 A comparison of grid point and spectral methods in a meteorological problem. Quart.J.Roy.Met.Soc., 100, 371-383.

Eliasen, E., Machenhauer, B., and Rasmussen, E. 1970 On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. Report No.2 available from: Institut fur teoretisk meteorologi, Kobenhavens Universitet, Copenhagen, Denmark.

GARP Publication Series No.17, 1979. Numerical methods used in atmospheric models Vol.2 WMO/ICSU, Geneva, Switzerland.

Jarraud, M., Girard, C., 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.

Lilly, D.K. 1965 On the computational stability of numerical solutions of time dependent non-linear geophysical fluid dynamics problems. Mon.Wea.Rev., 93, 11-25.

Orszag, S.A. 1970 Transform method for calculation of vector coupled sums. Application to the spectral form of the vorticity equation. J.Atm.Sci., 27, 890-895.

Orszag, S.A. 1974 Fourier series on spheres. Mon.Wea.Rev., 102, 56-75.

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 Strufing, R. 1981 An energy and angular momentum conserving finite difference scheme, hybrid coordinate and medium range weather prediction. ECMWF Tech.Rep.No.28, pp.67.

Sundquist, H. 1975 On truncation errors in sigma-system models. Atmosphere, 13, 81-95.

APPENDIX A: DATA PROCESSING DIFFERENCES

The model forecasts are provided on different horizontal grids and because of spectral filtering, the mountain field which determines the σ-surfaces is not the same. Data-processing differences cannot, therefore, be made to vanish completely. They can however be greatly minimized. As mentioned in the introduction, two different post-processing (PP) packages (a research version, RPP, used in TR23 for both models and a newly developed operational version, OPP) were available at the beginning of the present series of experiments. The impact of the differences on the objectively calculated model performance was not originally recognized and T63 forecasts were initially post-processed using RPP.

Fig. 67 shows maps of Z500, Z1000 and T850 differences (RPP-OPP) due uniquely to the use of either PP packages in obtaining the p-level fields on the verification grid from the same σ-level fields on the model grid for the DEC 27 T63 forecasts. Largest differences in T are observed over mountainous regions (the Rockies, the Himalayas and Greenland) as a direct consequence of differences in (unnecessary) extrapolation (below ground) procedures in PP packages. Differences in Z1000 also show a tendency to be largest over mountains (the Himalayas, in particular). At 500 mb, a large area of negative differences is seen in the South Pacific. Over Z maps a difference centre is observed right at the pole. Other difference centres appear at random on all maps. A major inconsistency between the two PP packages, besides extrapolation, was found to be due to an error in OPP (corrected since then), which used the true temperature instead of the virtual temperature to recover heights by integration of the hydrostatic equation. This leads to a negative bias in the height field obtained by OPP, particularly in the tropics and it explains in particular, the differences at 500 mb over the South Pacific.
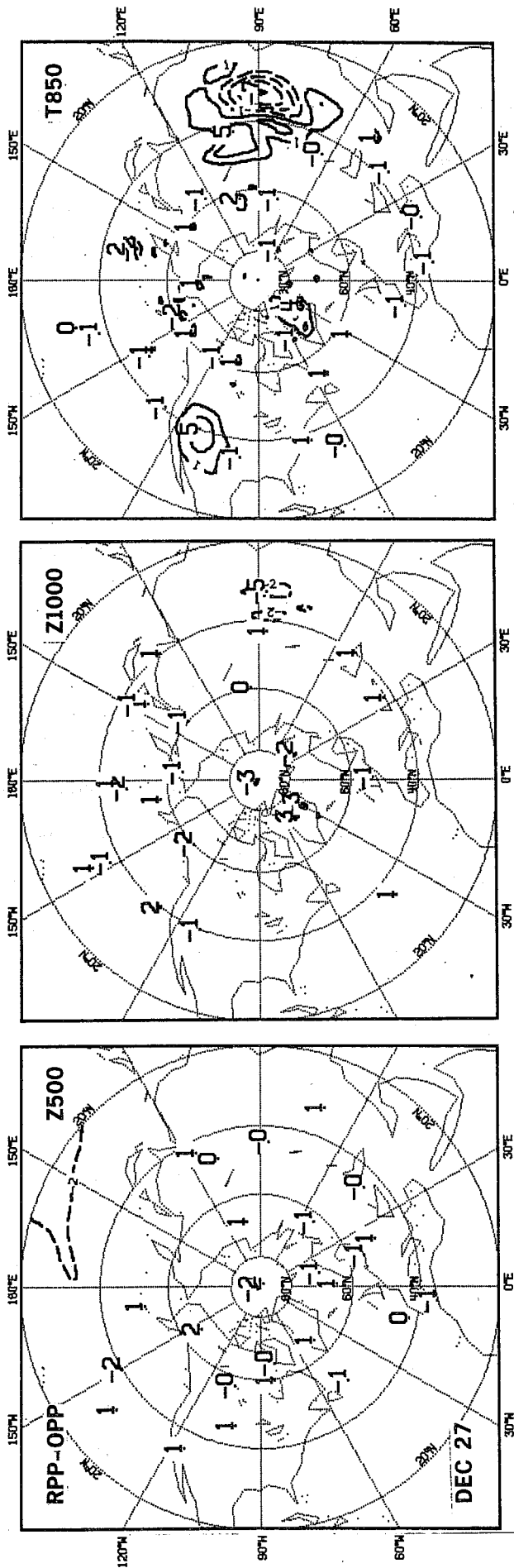
Fig. 67 Examples of post-processing differences (RPP-OPP) for Z500, Z1000 (dam) and T850 (^{o}K) taken from DEC 27 forecast.

In addition, both the horizontal and vertical interpolation procedures differ in more minor details. In the vertical, for example, heights are not recovered at the same σ-levels prior to the σ-to-p interpolation. In the horizontal, OPP does a straightforward spectral interpolation while RPP uses linar interpolation in the NS direction to go from the model Gaussian latitudes to the uniform 2.5°-latitude grid used for verification. Moreover, the pole value in RPP for T63 is taken to be the mean of the values on the nearest Gaussian latitude circle. This explains the large differences there.

Note that the differences are likely to be systematic from forecast to forecast. The few forecasts for which both PP packages were used, indeed proved this for the larger differences.

The impact of the differences shown in Fig. 67 on the evaluated performance of T63 is demonstrated in Table 8, giving single level values of the differences in AC and SE (DAC and DSE) observed between T63 forecasts post-processed with OPP or RPP and between T63 and N48 forecasts, and in Fig. 68, showing pressure-time cross-sections for the same scores. Clearly, the impact is relatively very large for temperature in the lower troposphere, with mean hemispheric DAC and DSE as large as 2% and 1°C respectively and possibly even larger. The impact on Z500 and Z1000 is also far from negligible (DAC and DSE as large as 1% and 2m). For example (Table 8), on D+4, T63 with OPP is better than N48 by 0.6, 1.2, 3.0 points compared to -0.9, 0.5, 2.3 with RPP respectively for T850, Z500, Z1000. These differences have to be compared to the mean improvements by T63 (Fig. 5).

Table 8. Score differences due to post-processing

		D+0	D+1	D+2	D+3	D+4	D+5	D+6	D+7
	DAC(PP)	.2	.0	.0	.3	.7	.8	.3	.0
Z500	DAC(T63-N48)	.0	-.1	.5	.5	1.2	.1	6.4	13.3
*	DSE(PP)	6.1	0.2	0.5	0.8	1.2	1.2	0.9	0.4
	DAC(PP)	.1	.6	.8	.6	.7	.7	.4	.1
Z1000	DAC(T63-N48)	.0	-.2	.0	.1	3.0	.1	2.0	8.9
*	DSE(PP)	2.8	0.9	1.0	0.8	0.6	0.5	0.4	0.3
	DAC(PP)	1.8	1.9	1.4	1.4	1.5	.6	.2	-1.0
T850	DAC(T63-N48)	.0	-.4	-1.0	-.3	.6	.3	3.1	8.6
*	DSE(PP)	7.4	2.1	1.0	0.6	0.3	-.1	0.2	-.4

* DAC are given in %

DSE are given in m for Z and in ^{o}C for T

The reasons why these observed PP diferences produce such a large impact on scores are: first, the differences are systematic, second, any inconsistencies (including possible errors in OPP, not in RPP) between PP packages become additional forecast errors in T63 when using RPP, because forecasts are compared to analyses which have been similarly post-processed with OPP.

All forecasts by T63 from SEP 3 to DEC 13 inclusively were post-processed using RPP and, as mentioned in Section 2, a clear bias exists in the scores in favour of N48 for these cases. For cases from DEC 20 to JAN 24, both PP files were produced and kept for analysis but T63-OPP files were used for verification. For the remaining cases only OPP files were produced.

In comparing the models' performance, no attempt was made to correct for this bias.
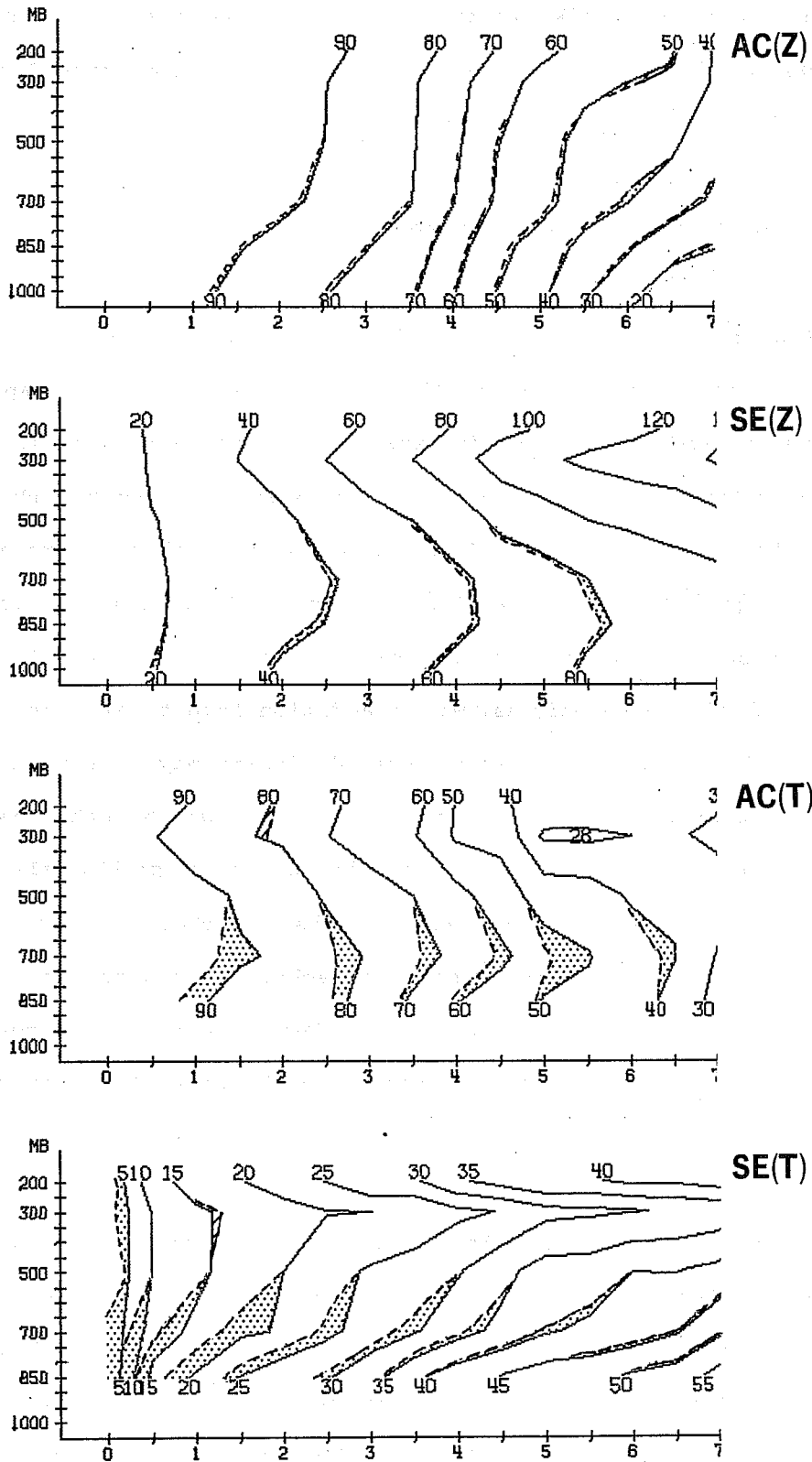
Fig. 68   Pressure-time cross-sections of AC(Z), SE(Z), AC(T), SE(T) for the
T63 forecast of DEC 27 using two different post-processing procedures.
(OPP= ──── ; RPP= ────).   Hatched (stippled) areas correspond to
better scores by RPP (OPP).

In conclusion, we would like to emphasize again the need, in forecast

comparisons, to take great care in designing data processing procedures in

order to avoid biases and spurious discrepancies.

APPENDIX B: IMPACT OF HORIZONTAL DIFFUSION

B.1 Introduction

Some form of horizontal diffusion (henceforth referred to simply as

diffusion) is generally introduced in large scale atmospheric models. Its

theoretical justification is the parameterization of the effect of sub-scale

motions on the resolved ones and this is often done in analogy to molecular

diffusion. In practice, it is useful to prevent the accumulation of (noise)

energy in the smallest resolvable scales of a model (spectral blocking)

which, for non-Galerkin techniques, may in addition lead to an undesirable

level of aliasing errors. In certain cases, diffusion may even be used to

maintain the stability of the numerical solution. On the other hand, a

diffusion scheme has to be selective to effectively remove small-scale noise

without unduly affecting the larger scales. Thus, the choice of scheme may

depend on the properties of the numerical technique used and is often a

compromise between the desired and the side effects. For this reason, the

impact of diffusion is of particular interest to the present comparison of

techniques.

Both N48 and T63 use relatively selective diffusion schemes of similar (∇^4)

type (initially though, N48 was using a non-linear type of diffusion). The

value itself of the diffusion coefficient (K), always larger in N48 (see

Table 1 for changes made during the experiment), is the major difference.

The relative loss in selectivity brought about by the finite difference

technique, as mentioned in Section 5.1, has of course, also played a role.

B.2. Possible impact on KE levels

A direct effect of using a higher K value is to reduce total KE and the effect should be more pronounced at higher wavenumbers. In Fig. 51, we compared the time evolution, up to D+5, of the seasonally averaged KE for the 4-9 and 10-20 zonal-wavenumber bands. In Figs. 52 and 53, we presented KE spectra at the beginning and end of the forecast period respectively. These graphs clearly indicated that T63 maintains on average a higher level of KE at high (10-20) wavenumbers throughout the forecasts. Up to D+5, the same tendency is observed for the medium (4-9) waves. Note that the differences appear at the beginning of the forecast period (D+1) and do not seem to grow substantially thereafter.

Could a difference in the value of K be responsible for these results? To answer this question, a realistic sensitivity test was made with T63, on the Feb 14 case. Fig. 69 shows the time evolution of the total KE, up to D+5. The curves for N48 and T63 are not atypical in the light of our experience. The curves labelled T63- and T63+ correspond to T63 (standard $K = 7.10^{14}$) with respectively lower ($K- = 1.10^6$, this is very nearly equivalent to no diffusion at all) and higher ($K+ = 3.10^{15}$, this is roughly equivalent to what is used in N48) diffusion. All models suffer a loss of KE, rather more quickly than could be explained by diffusion alone which, as observed, operates on a much longer time scale. In fact the short-term (up to D+3) effect of increasing K in this case was to also increase KE. The difference between T63 and N48 is already maximum on D+1, in agreement with the mean observed difference but in complete disagreement with the diffusion experiment. It is thus very unlikely that a larger K in N48 is directly responsible for the resulting difference.

Concerning the initial rapid growth of the differences, there is more qualitative evidence. Indeed, one of the major deficiencies of the present series of forecasts by both models is related to an imbalance in the thermal
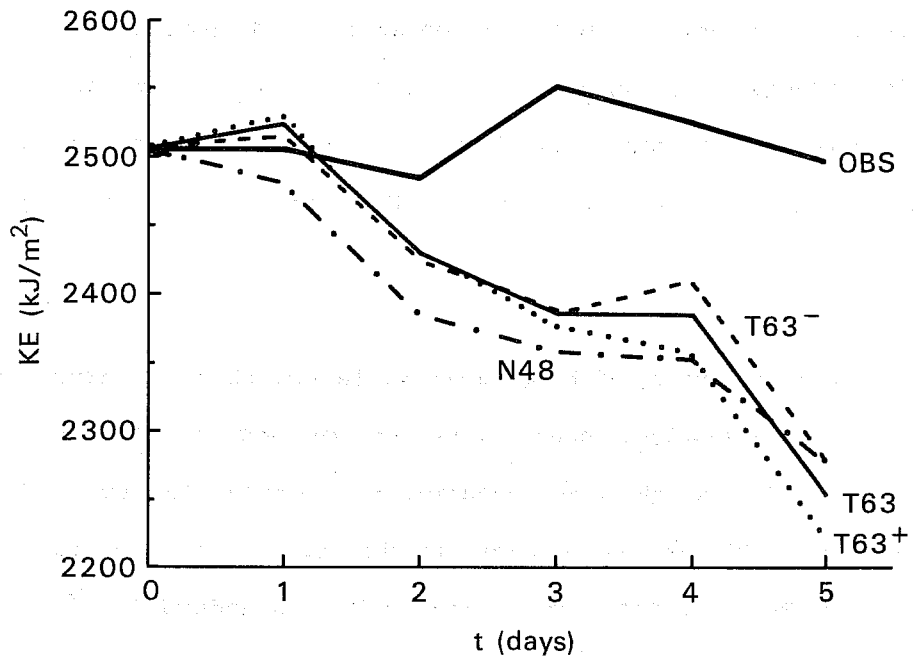
Fig. 69 Time evolution up to D+5 of globally averaged
KE (kJ/m^2) in OBS and for FEB 14 forecasts by N48
and by T63 with various horizontal diffusion
coefficients: k (T63)=7.10^{14} m^4 s^{-1};
k (T63$^-$)=10^6 m^4 s^{-1}; k (T63$^+$)=3.10^{15} m^4 s^{-1}.

forcing leading to a global mean cooling of the model atmosphere when compared to observations (Fig. 55). Moreover, as will be shown below, the imbalance is greater at the beginning of the forecast period with the maximum cooling rate occurring at initial time (spin-up). Due to this spin-up problem, the early stage of the forecasts becomes a critical period for their future evolution and any differences in the models' response are likely to be amplified. A detailed analysis of this problem with respect to the KE levels maintained by the models appears beyond the scope of this study.

B.3. Impact on models' thermal balance

An example (the Jan 16 case) of the time evolution of the global heating rates by surface fluxes (F), large scale condensation (L), convection (C), and (cooling by) radiation (R) is given in Fig. 70. One can see clearly the net imbalance between heating sources and sinks coupled with spin-up problem. There are differences between the models mainly in L and R (but R is not unrelated to L since it is affected by clouds which are correlated with L). These are already large on D+1. Both terms contribute to produce a smaller net cooling rate in T63.

In Section 4 (Figs. 55 and 56) we also observed that the mean temperature differences between the models were largely limited to the middle-latitude midtroposphere and over geographical areas corresponding to the mean cyclone tracks. These facts agree with the present finding that the main cooling differences are mostly due to differences in L. This is further confirmed by Fig. 71 showing global σ-latitude cross-sections of the 10-day mean L-heating for both models (the case of JAN 16). Clearly the L-heating occcurs almost exclusively in the middle latitude midtroposphere (apart from cooling by evaporation of rain) and the differences account for a good portion of the temperature differences. These, in turn, as will be seen later, are not unrelated to the differences noted in zonal AE (Fig. 54).
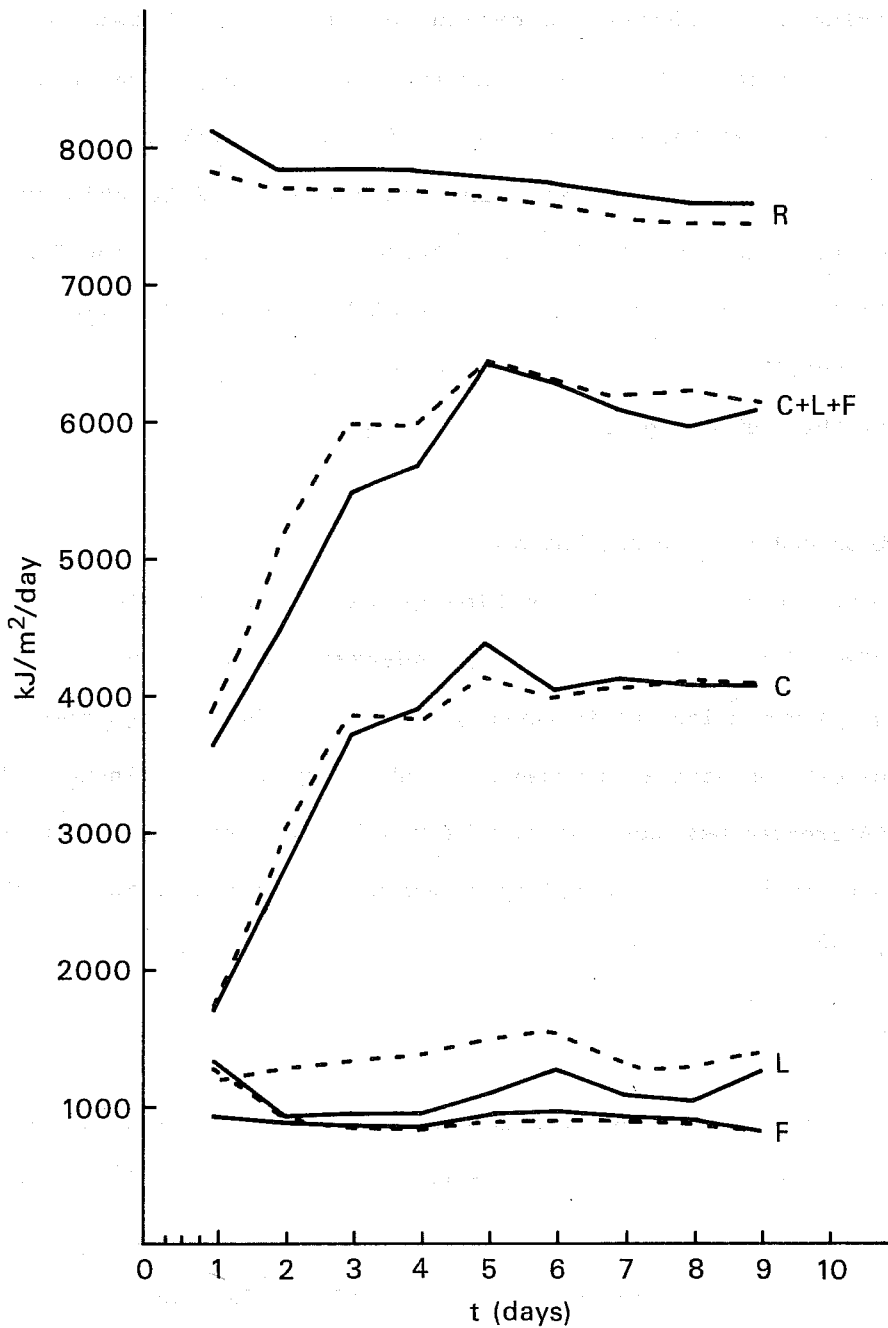
Fig. 70 Time evolution up to D+9 of the global radiative
cooling (R) and the global heating rates
(kJ/m^2/day) by surface fluxes (F), large scale
condensation (L) and convection (C) for T63 and
N48 forecasts on JAN 16.
(N48= ———— ; T63= -----).

T63
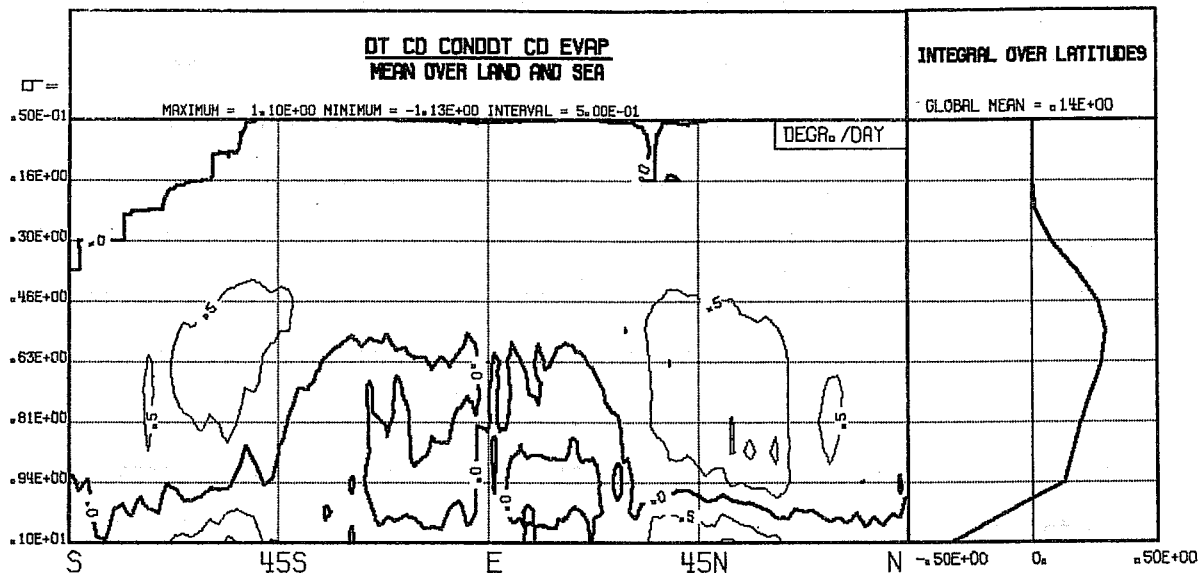
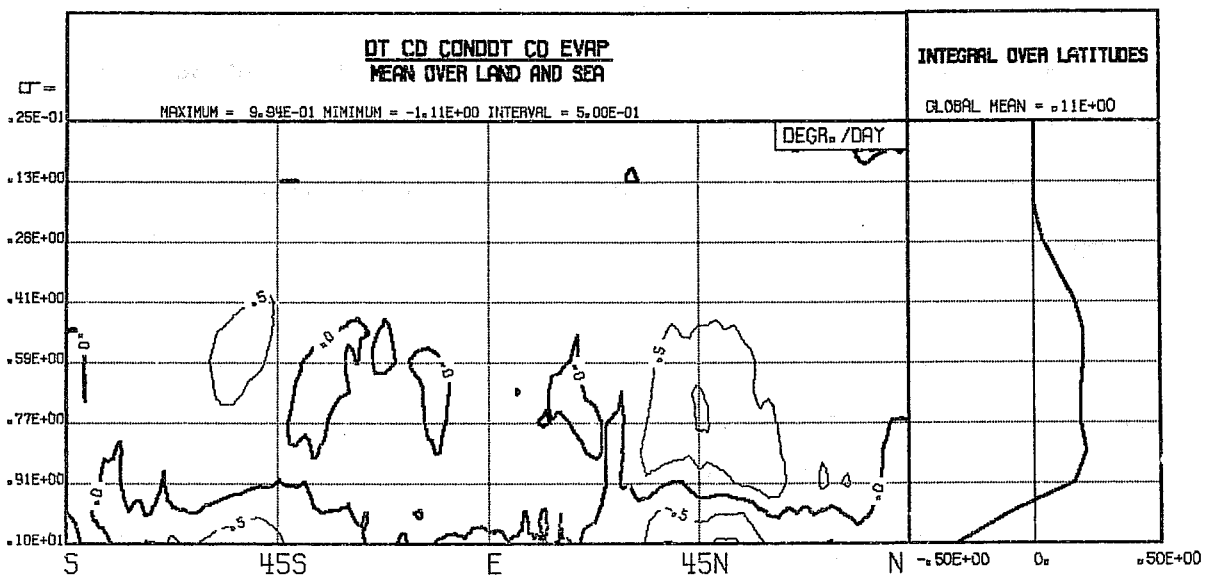## N48

Fig. 71 Global and latitude cross-section profiles of 10 day mean heating rates ($^{\circ}$K/day) due to large scale condensation and evaporation by T63 and N48 on JAN 16.

169

At this point, it is important to mention that none of these thermodynamic differences were really apparent in the first series of experiments reported in TR23, but then also none of the models appeared to be loosing an appreciable amount of zonal AE. The conditions in the two series of experiments were however very different and it would be hazardous to try and blame the differences on any single cause. The mean thermodynamic (AE) differences are in fact larger, relatively, than the mean dynamic (KE) ones and in the following we show that they are clearly related, however indirectly, to the amount of horizontal diffusion imposed on the forecasts.

Fig. 72 shows the σ-profiles of the global 10-day averaged cooling incurred by T63 for different values of K (case of FEB 14: T63, T63+, T63-). Note that substantial changes accordingly occur only in the midtroposphere. Since the T63- curve corresponds effectively to the no-diffusion profile, it is clear that diffusion is not entirely responsible for model cooling, but a large diffusion greatly enhances the cooling problem. Already, at standard diffusion T63 produces a colder forecast by about half a degree in the midtroposphere than it would with no diffusion.

Fig. 73 shows the time evolution,up to D+5, of total AE (KJ/m^3) for the same case comparing N48 and T63 (T63+ and T63-) to OBS. **As K is increased the AE curves for T63 tend towards that of N48.** Note that the models exhibit maximum divergence from OBS by D+1 with a tendency to recover afterwards. Note also that a large difference between the models already exists by D+1, suggesting a large impact again (as for KE) of the spin-up. However, as seen in Fig. 70, T63 maintains a higher level of L-heating and lower level of R-cooling throughout the 10-day forecast period.
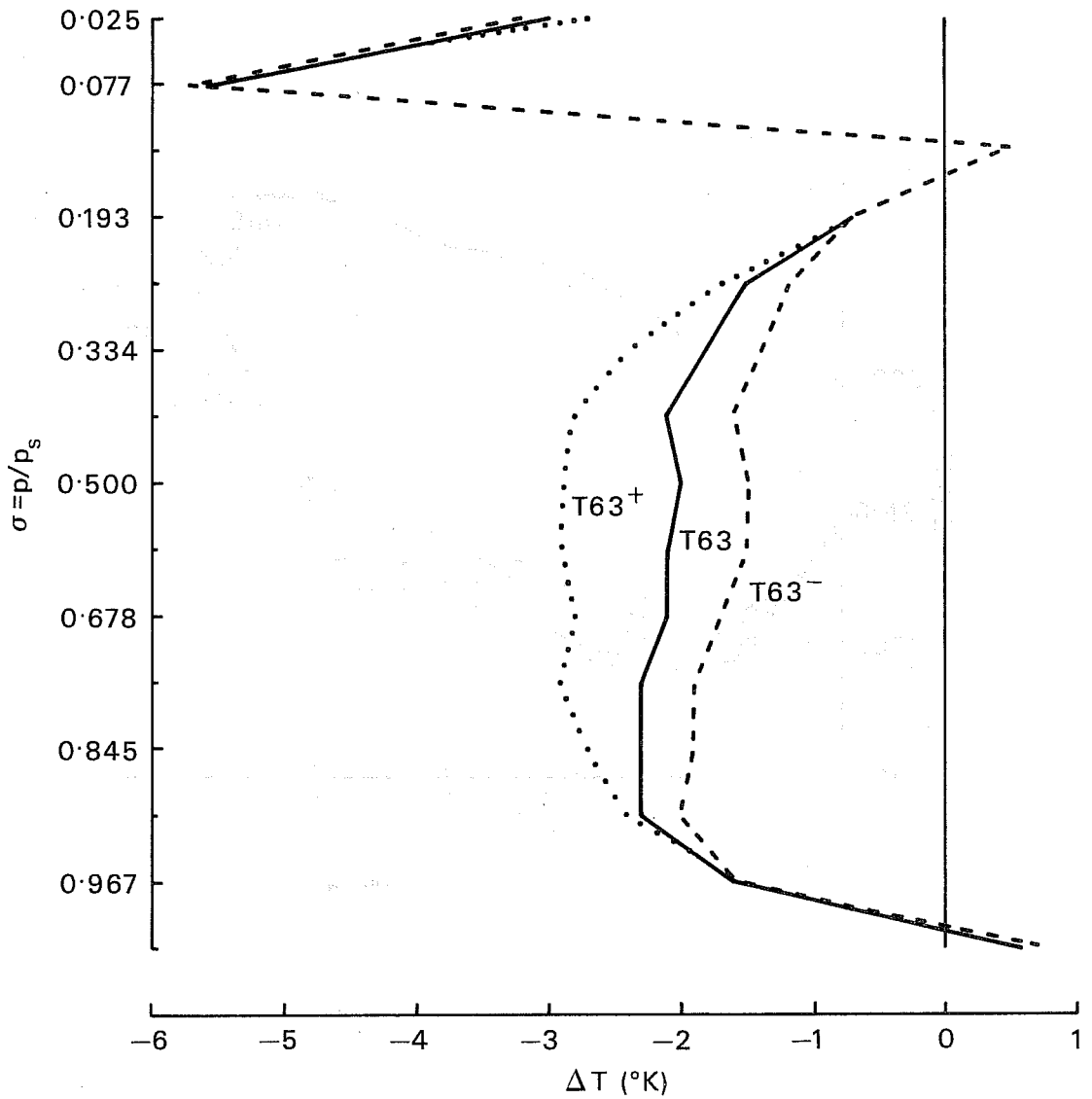
Fig. 72 Global σ-profiles of the 10-day mean temperature changes
(ΔT) forecast on FEB 14 by T63 with different horizontal
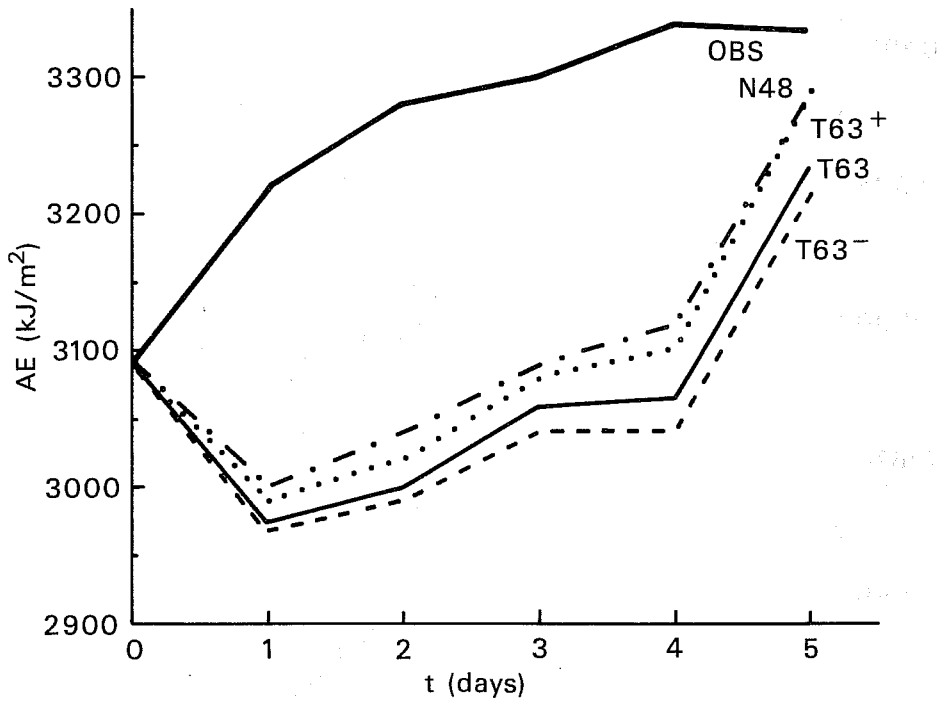diffusion coefficients (as defined in Fig. 69).

Fig 73   As Fig. 69 for globally averaged AE (kJ/m^2).

B.4 Additional sensitivity experiments.

The cooling influence of diffusion on global mean temperature is a puzzling feature of the present ECMWF models and is probably related to the thermal imbalance itself. But this problem is also beyond the scope of this study. However, some understanding of how diffusion affects large scale release of latent heat has been obtained by performing another set of sensitivity experiments using a low resolution (T21) spectral model. At present, diffusion is applied uniformly on all 4 model variables: vorticity (ξ), divergence (D), temperature (T) and specific humidity (q). To find out the effect of an increased K on any of the variables, T21 was first run with no diffusion at all (K = 0) and a weak diffusion ($K_1 = 2x10^{15}$) on all variables. These 2 runs were almost indistinguishable for our purposes.

Then a higher value ($K_2 = 3 \times 10^{16}$) was applied in turn to different subsets of the variables, and the integration repeated. The results are summarized in Fig. 74, displaying the time evolution (up to D+9) of the global mean temperature (from an initial value of $T_o = 254.19$) for the different forecasts. Global cooling is found to be sensitive to diffusion of any one of the variables but in particular to q-diffusion and to a lesser degree to ξ-diffusion. T- or D-diffusion have negligible impact (less than 0.1°K after 9 days). When applied to more than one variable the resulting effect is approximately equal to the sum of the individual effects. **A large influence of q-diffusion on L-heating is not unreasonable since L-heating is dependent on moisture availability.** That the global influence of T-diffusion is small is reassuring. As for ξ - and D-diffusion their role may be easily understood once it is recognised that **L-heating is also dependent on moisture convergence and that large scale convergence, small in comparison with vorticity, is more effectively controlled by ξ- than by D-diffusion.** Table 9 establishes this fact. It gives the D+10 asymptotic level of RMS divergence (10^{-6} sec^{-1}) achieved in some of the above mentioned T21 forecasts. In this instance, ξ-diffusion is shown to be about 400% as effective as D-diffusion
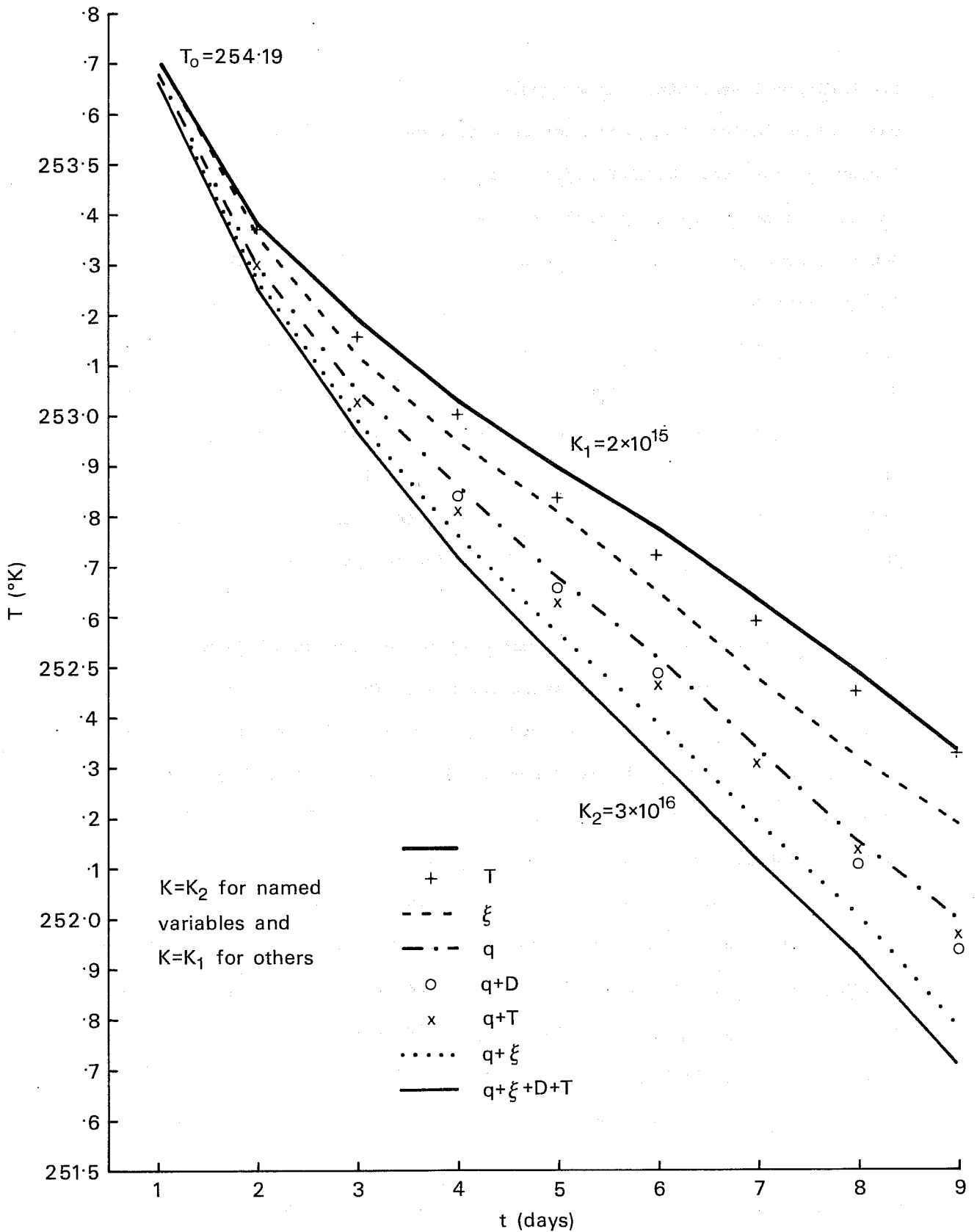
Fig. 74   Time evolution up to D+9 of the global mean temperature, forecast by T21
when two different values of the horizontal diffusion coefficient
($K_1 = 2.10^{15}$ m^4 s^{-1} and $K_2 = 3.10^{16}$ $m^4 s^{-1}$) are alternatively applied on the
four model variables: vorticity (ξ), divergence (D), temperature (T) and
specific humidity (q).

174

in lowering the level of RMS divergence.

Table 9. Effect of increased diffusion on mean RMS divergence

$\sqrt{\bar{D}^2}$	Variables for which $K=K_2$ ($K=K_1$ otherwise)	
5.0	none	
4.7	q	$K_1=2 \times 10^{15}$
4.5	q,D	
3.9	q,ζ	$K_2=3 \times 10^{16}$
3.7	q,ζ,D,T	

B.5 Summary of the diffusion experiments

The impact of diffusion differences on FDs can be summarized as follows. On

the one hand, the different levels of mean eddy KE in T63 and N48, at short

range cannot be due to them and there is evidence to show that they derive

instead from differences in the overall dynamical response of the models to

the large fictitious forcing implied by initial imbalances (spin-up problem).

It appears that T63 remains slightly more correct on average in this respect.

On the other hand, the thermodynamic differences are related to diffusion,

although not unrelated to the spin-up problem. Basically, T63 maintains a

warmer middle-latitude midtroposphere by achieving a higher mean L-heating

throughout the forecast period and particularly at the beginning. At the

same time, however, its level of AE is depressed away from OBS, particularly

zonal AE.

By increasing K fourfold in T63 it was possible on one case (FEB 14) to make

the differences in zonal AE essentially vanish. On the mean temperature

forecasts at medium range, the impact of diffusion is important (it explains most of the mean differences noted in Section 4) but it is overshadowed by much larger temperature errors. Until a better thermal balance is obtained for the parameterization and unless the spin-up problem is satisfactorily resolved, it will be difficult to ascertain the correct level of diffusion. Meanwhile note must be taken of the high sensitivity of L-heating to K.

ECMWF PUBLISHED TECHNICAL REPORTS

No. 1   A Case Study of a Ten Day Prediction

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

No. 3   Mixed-Radix Fast Fourier Transforms without Reordering

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

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

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

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

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

No. 9   On Balance Requirements as Initial Conditions

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

No.11   Normal Mode Initialization for a multi-level Gridpoint Model

No.12   Data Assimilation Experiments

No.13   Comparison of Medium Range Forecasts made with two Parameterization
         Schemes

No.14   On Initial Conditions for Non-Hydrostatic Models

No.15   Adiabatic Formulation and Organization of ECMWF's Spectral Model

No.16   Model Studies of a Developing Boundary Layer over the Ocean

No.17   The Response of a Global Barotropic Model to Forcing by Large-Scale
         Orography

No.18   Confidence Limits for Verification and Energetics Studies

No.19   A Low Order Barotropic Model on the Sphere with the Orographic and
         Newtonian Forcing

No.20   A Review of the Normal Mode Initialization Method

No.21   The Adjoint Equation Technique Applied to Meteorological Problems

No.22   The Use of Empirical Methods for Mesoscale Pressure Forecasts

No.23   Comparison of Medium Range Forecasts made with Models using Spectral or
         Finite Difference Techniques in the Horizontal

No.24   On the Average Errors of an Ensemble of Forecasts

No.25   On the Atmospheric Factors Affecting the Levantine Sea

No.26   Tropical Influences on Stationary Wave Motion in Middle and High Latitudes

ECMWF PUBLISHED TECHNICAL REPORTS