IMPACT OF HORIZONTAL RESOLUTION ON EXTENDED-RANGE FORECASTS AT ECMWF

S. Tibaldi*, Č. Branković, U. Cubasch**, F. Molteni
European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire

1. INTRODUCTION

As a result of the ongoing efforts to improve the ECMWF's operational model climate, the Centre's model mean errors are now lower in amplitude and their growth is slower. It has therefore become necessary to diagnose them over a longer forecast range and, in particular, to diagnose their dependence upon season and model resolution. An experimental programme is therefore under way to provide a suitable database of extended (30-day) integrations during different seasons and with a variety of model resolutions. Such a database provides not only a source of information for mean error studies, but it also affords the opportunity to examine the forecasting ability of the Centre's operational system in the extended range, and the dependence of forecast skill in this range upon different weather regimes and model characteristics.

This report describes results from an analysis of the first year of integrations. Section 2 describes the database, Section 3 is devoted to the analysis of the mean errors, while Section 4 discusses the objective skill scores of the extended range integrations. Section 5 summarizes the main results and draws conclusions to form the basis for future work.

2. THE DATABASE

The set of experiments consists essentially of a monthly series of integration "pairs", started from two consecutive initial conditions, separated by 24 hours, around the mid-month period. The first month considered was April 1985 and the last was March 1986. All integrations have been performed at four different model resolutions: T21, T42, T63 and T106 and with the operational 16-level, envelope orography version of the model, with the physical parametrization package as of March 1986. The shallow convection and modified

---

* Department of Physics, University of Bologna, Italy
**Max-Planck Institut für Meteorologie, Hamburg, Federal Republic of Germany

Kuo convection parametrizations (Tiedtke et al., 1988) are included, but gravity wave drag (GWD, Miller et al., 1988) and modified surface exchanges (Blondin and Böttger, 1987) are not. Some results from the first winter season of T106 runs (Oct 85 - March 86) were also discussed in Hollingsworth et al. (1987).

3. THE MODEL MEAN ERRORS

This section will be devoted to the diagnosis of the model's mean errors as a function of forecast time (the so-called "climate drift") for all model resolutions and for two "extended" seasons, the October to March (OM) period (Northern Hemisphere "extended" winter), and the April to September (AS) period (Northern Hemisphere "extended" summer). The analysis discusses the zonally averaged fields, the mean mid-latitude fields (of both Northern and Southern Hemispheres (NH and SH)) and some fields of particular relevance to tropical regions. For a more comprehensive discussion on the climate drift of the operational model during the first ten days of integration, the reader is referred to Arpe (1988) and Brankovic (1986).

3.1 The zonally averaged mean errors

Fig 1 shows latitude-height cross-sections of the zonal mean of zonal wind for the ensemble means of the last ten days (days 21-30) of model integrations at the four different resolutions (top to bottom), plus analysed values for comparison. The left panels refer to the OM period, while the right panels refer to the AS period. At first sight the model-simulated mean zonal flow field seems quite satisfactory. A closer examination of the difference fields for wind and temperature (Figs 2 and 3) shows, however, several important discrepancies. Although jet maxima and the general shape of the isolines of [U] and [T] are fairly well reproduced, their north-south positioning shows significant poleward shifts, with a fairly equivalent-barotropic structure, at least in extratropical regions.

Regarding the dependence of zonal mean errors upon model resolution, the T21 model stands out from the others. Its zonal climate is in some respect worse and in others better than those produced by the higher resolution models. The NH winter (OM) jet is better positioned and the tropospheric errors are smaller in the T21 runs, but the SH zonal flow and the tropical low-level easterlies are far too weak. Possible reasons for such discrepancies have

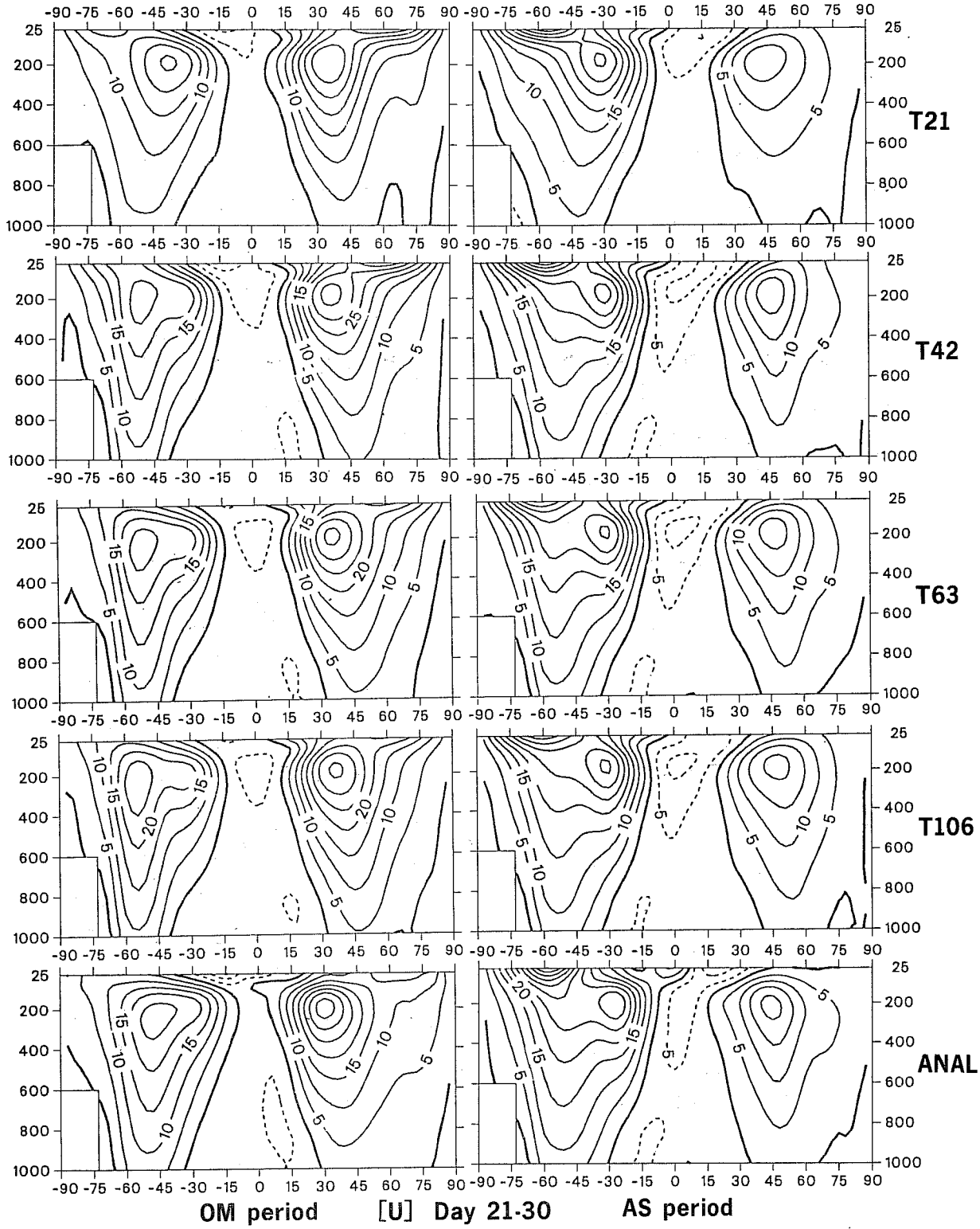

Fig 1 Latitude-height cross-sections of zonal mean of zonal wind [U]. Day 21-30 time mean. Top to bottom: T21, T42, T63, T106, analysis. Left: OM period; right: AS period. 12 cases in each period. Contours every 5 ms^{-1}.

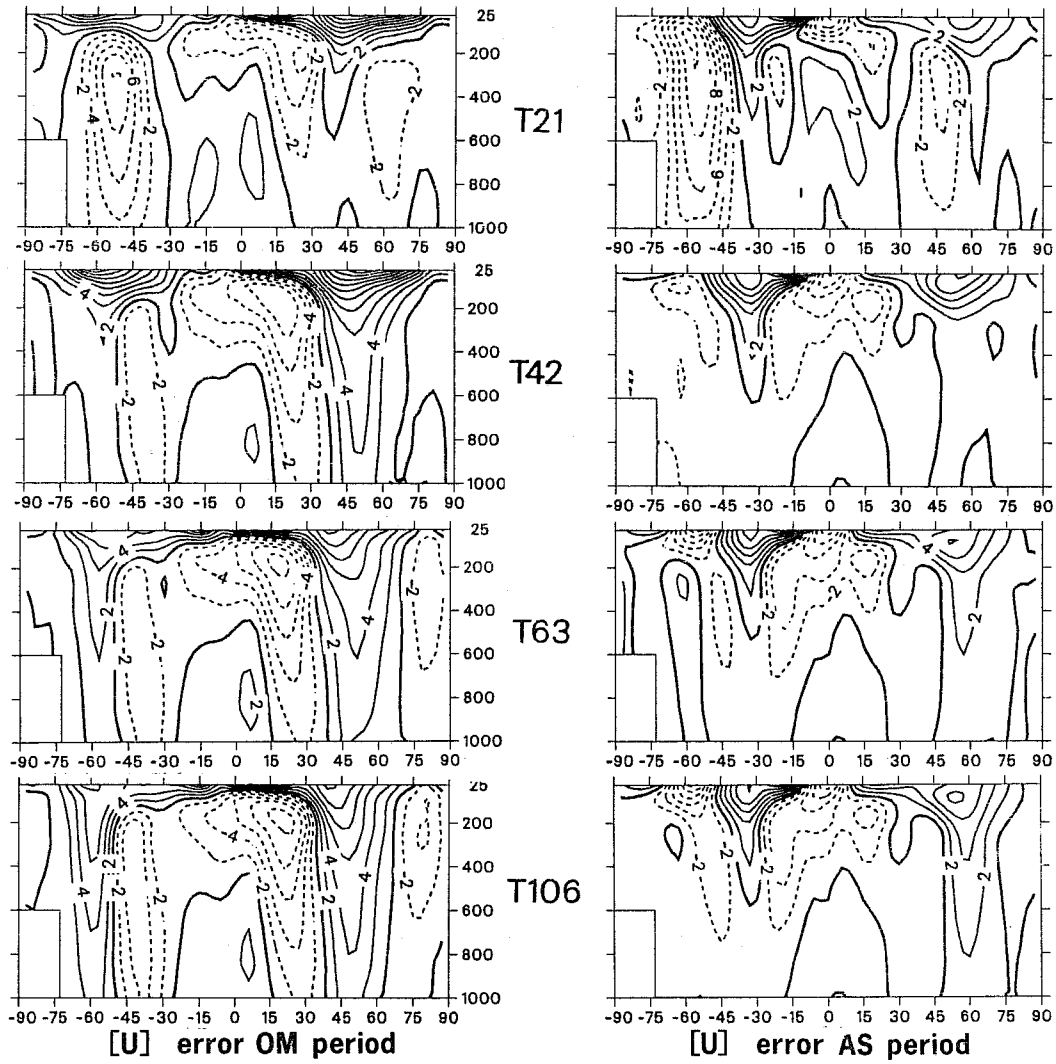

Fig. 2 Latitude-height cross-sections of zonal mean of zonal wind errors for days 21-30 time mean. Top to bottom: T21, T42, T63, T106. Left: OM period; Right: AS period. Contours every $2ms^{-1}$.

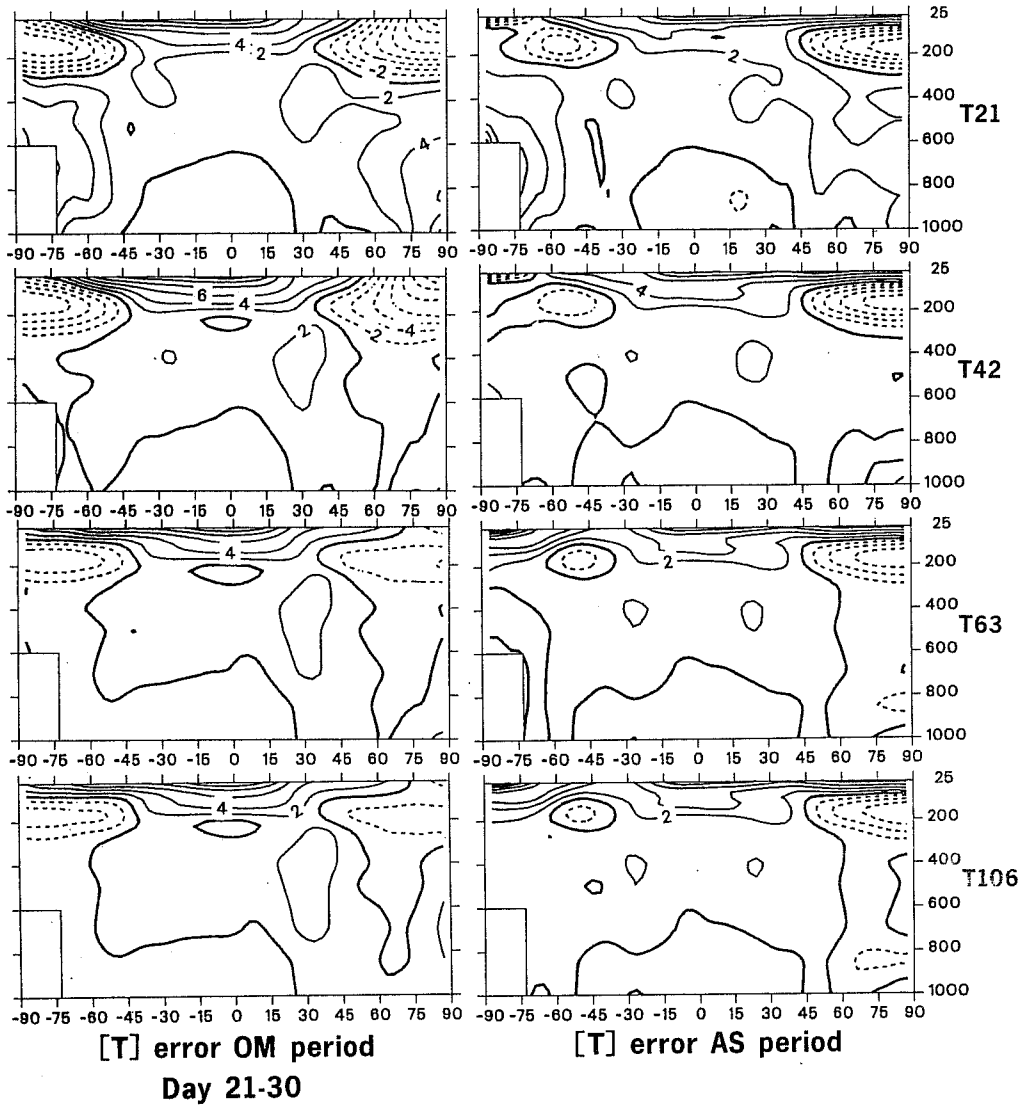

Fig. 3 Latitude-height cross-sections of zonal mean temperature errors for days 21-30. Top to bottom: T21, T42, T63, T106. Left: [T] OM period; right: [T] AS period.
Contours every 2 K.

been discussed by Miller et al. (1988) who suggest that they are probably connected with a compensation of resolution related errors between the eddy momentum flux convergence term and the orographically related zonal flow dissipation mechanisms (e.g. form drag). The T42 model climate has a somewhat "transitional" nature, with features resembling more the T63-T106 family, but also with features typical of the T21 behaviour, e.g. in the SH stratospheric [U] error during the OM period, and in the intensity of the extratropical stratospheric [T] errors, in both hemispheres and in both seasons. The [T] and [U] errors for the highest resolution models are very similar both in amplitude and pattern, with the T63 model showing marginally smaller amplitude mean [U] errors.

We now discuss the T106 errors in some detail; the discussion applies equally to the T63 model and, to a lesser extent, to the T42 model as well.

Fig. 4 shows the day 21-30 mean cross sections (for the OM period) of both the observed mean fields (top) and the T106 mean errors (bottom) of [U] (left panels) and [T] (right panels). For purposes of discussion, the zonally averaged atmosphere has been divided into three main areas: area 1, the tropical regions; area 2, the NH extratropics and area 3, the SH extratropics. In the troposphere the subtropical regions belong to the "extratropics" areas, while in the stratosphere they belong to the "tropics" area. This is because in areas 2 and 3 the [U] mean errors are essentially equivalent-barotropic, as can be seen from the figures, while in area 1 they show marked changes with height. A comparison between the error fields and the observed fields shows that, in area 1 where the errors show a vertically varying structure, their sign is consistent with the hypothesis that their main effect is to smooth out the vertical shear of the zonal flow (positive errors where the mean flow is easterly and negative errors where the mean flow is mainly westerly). This is consistent with results of Miller and Viterbo (pers. comm.) that show that systematic errors (SEs) in tropical regions are sensitive to the intensity of the vertical diffusion used in the model, and that the tropical SEs are somewhat reduced if the vertical diffusion algorithm is not applied above the planetary boundary layer. The main characteristic of the [U] error in areas 2 and 3 is not so much an amplitude error as a poleward shift of the main jet maxima. To discuss the possible origin of this it is better to distinguish between the stratosphere, where both the radiative

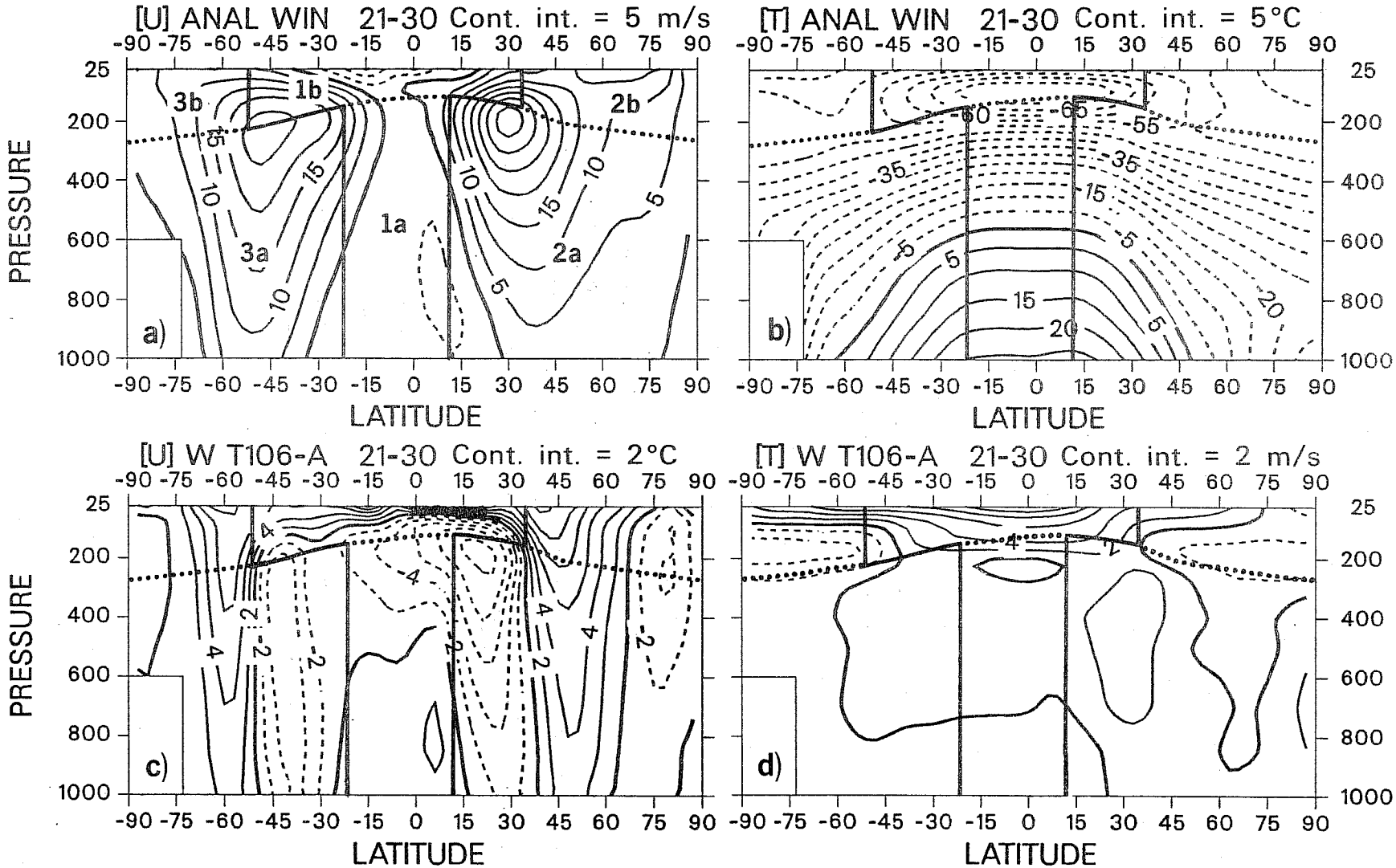

Fig 4 Latitude-height cross sections of day 21-30 zonal mean analysed fields (top) and forecast errors (bottom) of zonal wind (left) and temperature (right). For meaning of area division, see text. Contours every 5 ms^{-1} and 5°K for full fields and 2 ms^{-1} and 2°K for error fields.

equilibrium towards which the model tends, and the top boundary condition, are likely to play a predominant role, and the troposphere, where we have to look for a mechanism capable of shifting the jet cores poleward.

We already know that Gravity Wave Drag (GWD) and an improved stratospheric vertical resolution (19 levels compared to 16 levels) can alleviate part of the problem (they are both absent in the experiments discussed here) but are not able to change the essential pattern (and, therefore, probably nature) of the errors. The treatment of mid-latitude baroclinic eddies and their essentially barotropic feedback on the mean zonal flow (via convergence of eddy momentum flux) is a likely candidate for investigation. The eddy kinetic energy decreases progressively in midlatitudes during the model integration, in both seasons and for all wavelengths, while the zonal KE marginally increases in the NH winter for T63 and T106 models; Fig. 5 shows the kinetic energy as a function of forecast time for the NH mid-latitudes - SH is not shown. Indeed, the changes in the eddy momentum flux are consistent with these ideas, showing a poleward shift of the convergence zone in the OM period. Another interesting consequence of the elimination of the free-atmosphere vertical diffusion reported by Miller and Viterbo is to raise the level of eddy KE during the model integration, and to reduce (albeit to a much lesser extent than in tropical regions) the mid-latitude [U] mean errors as well. One possible link between vertical diffusion and the life cycle of baroclinic eddies is via its effect on atmospheric stability, which in turn affects the growth rate of the baroclinic waves.

Regarding the time evolution of the zonally averaged climate drift, Fig. 6 shows the time-latitude diagrams of [U] error, vertically integrated between 300 and 30 mb for the OM period. The error growth is nearly linear in time and takes place mainly during the first 20 days of integration. During the last ten days of the integrations the error growth seems to approach an asymptotic level. The T21 model behaves differently from all other three resolutions, with smaller errors in the NH and larger errors in the SH (growing much faster in the first ten days). The [U] errors for the NH troposphere show a marked seasonal cycle and are at a minimum in the AS season, with a faster growth of the error during the first 10 days, and more marked latitudinal shifts of the error features.

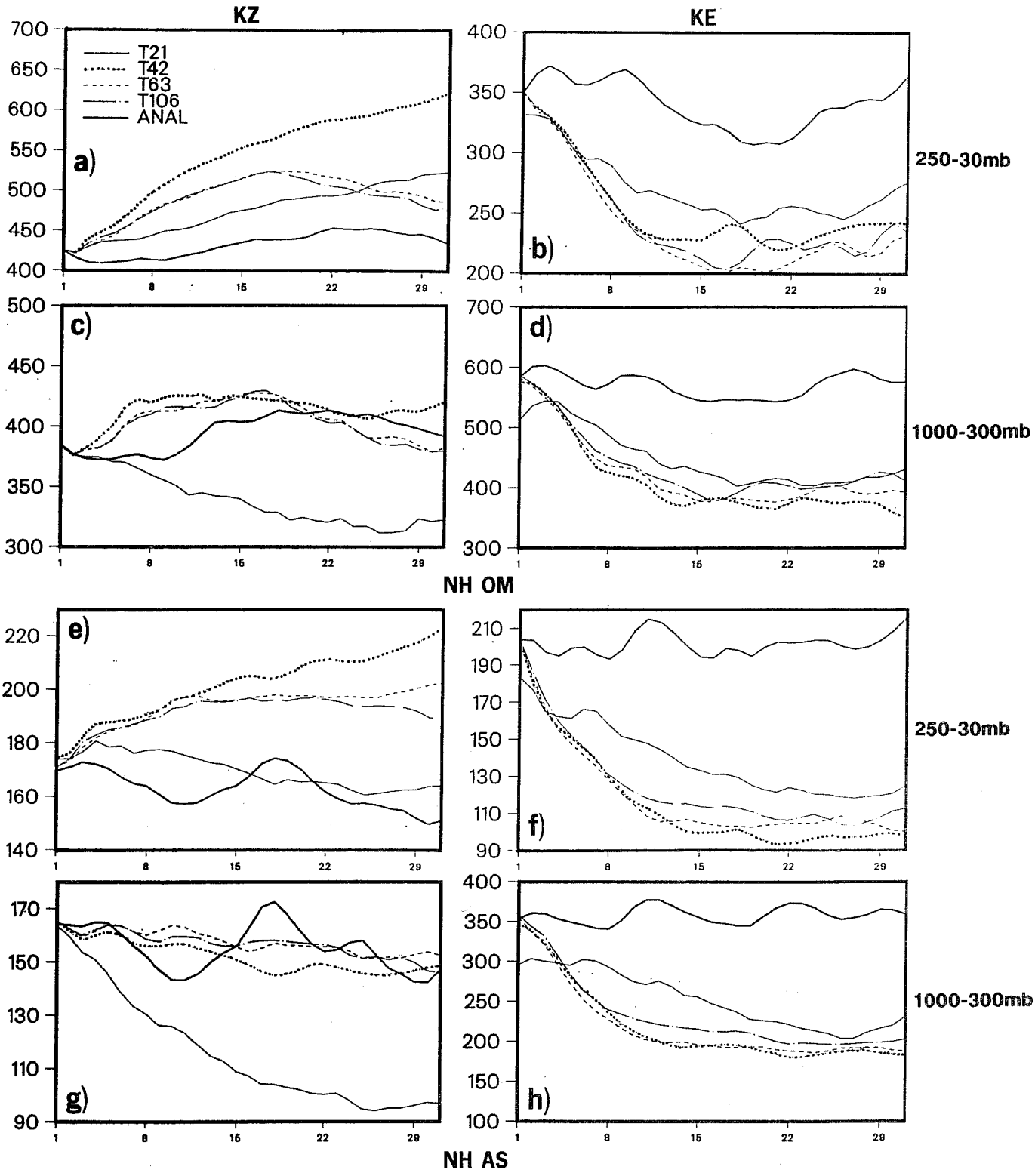

Fig. 5 Time evolution (day 1 to day 30) of zonal (left) and eddy (right) kinetic energy (kJ m^{-2})for northern hemispheric forecasts. T21, T42, T63 and T106 forecasts are shown, together with analysed values. a) to d) OM period; e) to h) AS period. a), b), e) and f) vertical integral from 250 mb to 30 mb; c), d) g) and h) from 1000 to 300 mb.

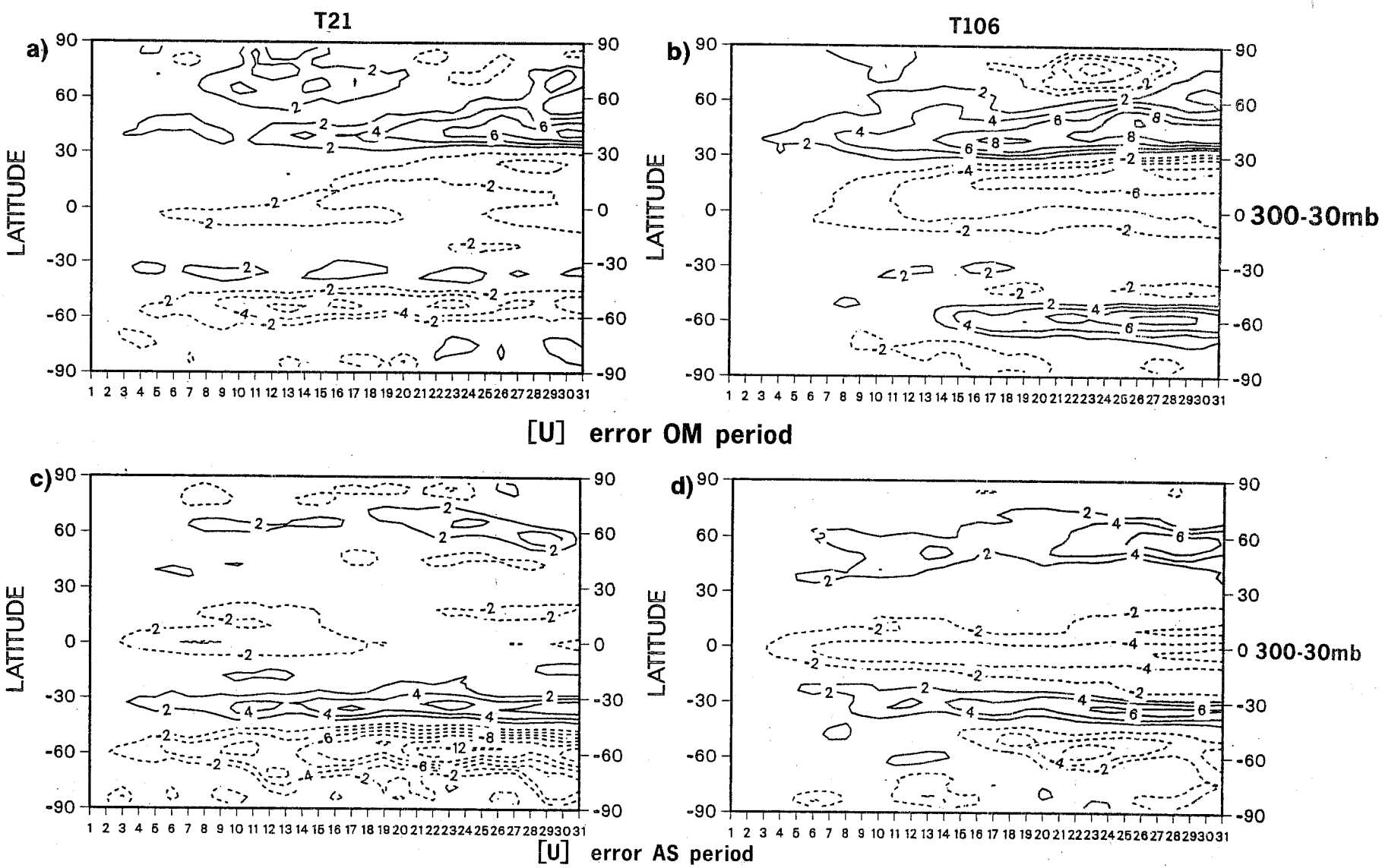

Fig. 6 Latitude-time diagrams of vertically integrated [U] error between 300 and 30 mb for the OM period (top) and AS period (bottom). Left: T21; right: T106.

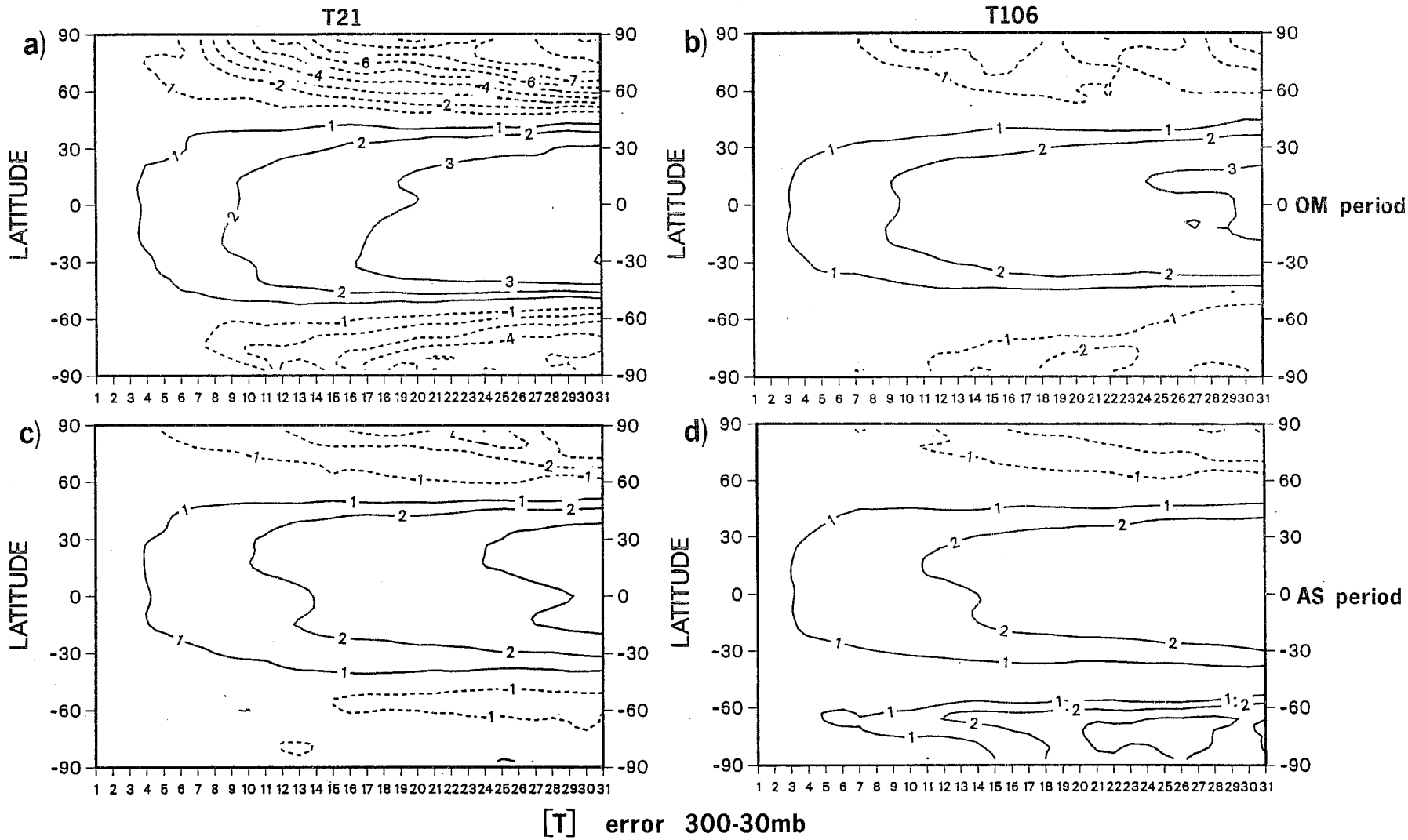

Fig. 7 Latitude-time diagram of vertically integrated [T] error between 300 and 30 mb.
Left: T21; right: T106. Top: OM period; bottom AS period.

A similar diagram (Fig. 7) highlights a large seasonal cycle in the SH [T] stratospheric errors (positive in the AS period and negative in the OM period) and an almost complete absence of seasonal cycle in the errors in the NH and in the tropical stratosphere. In the troposphere the [T] errors are small (tropospheric diagrams are not shown, but see Fig. 3). Also worthy of note are the similarities between the T21 and T42 [T] errors in both seasons, with the absence of the SH positive error area during the AS period that characterizes the T63 and T106 integrations. This is probably due to an interaction between the radiation scheme and the model's resolution.

3.2 The mean errors in the extratropical eddy fields

This section will be devoted to the extratropical mean errors of the model integrations and their dependence upon resolution and season. For the sake of brevity, the discussion will be limited to 500 mb height fields.

We first consider full fields. Fig. 8 shows mean 500 mb height maps for the last 10 days of integration (day 21-30), for the OM period, for all four model resolutions (T21, T42, T63 and T106) together with the observed maps. The overall picture that emerges is a familiar one: the T21 model shows a better simulation of the NH winter climate than the other (higher) resolutions but, conversely, heavily damps the strength of the westerlies in the SH during the same period, and also during the AS period (not shown). A possible explanation for this, connected to the representation of orography (and associated drag) and the convergence of eddy momentum flux has already been extensively discussed in Miller et al. (1988).

The T42 model shows, again, a "transitional" behaviour, while the T63 and T106 models behave very similarly. We therefore concentrate the discussion mainly on the T106 model maps from here on. The main characteristic is the loss of amplitude of the large-scale planetary waves, especially in the NH and during the OM period, when their observed amplitude is larger. Fig. 9 shows the actual Northern Hemisphere errors for all resolutions and for the averaging period day 21-30. Errors for both OM and AS seasons are shown. Comparisons with similar maps characteristic of older versions of the ECMWF operational model (e.g. Wallace et al., 1983) demonstrate that the rate of growth of such errors has been reduced considerably, so much so that typical 10-day SE amplitudes of the 1981-type operational model are now to be found around days

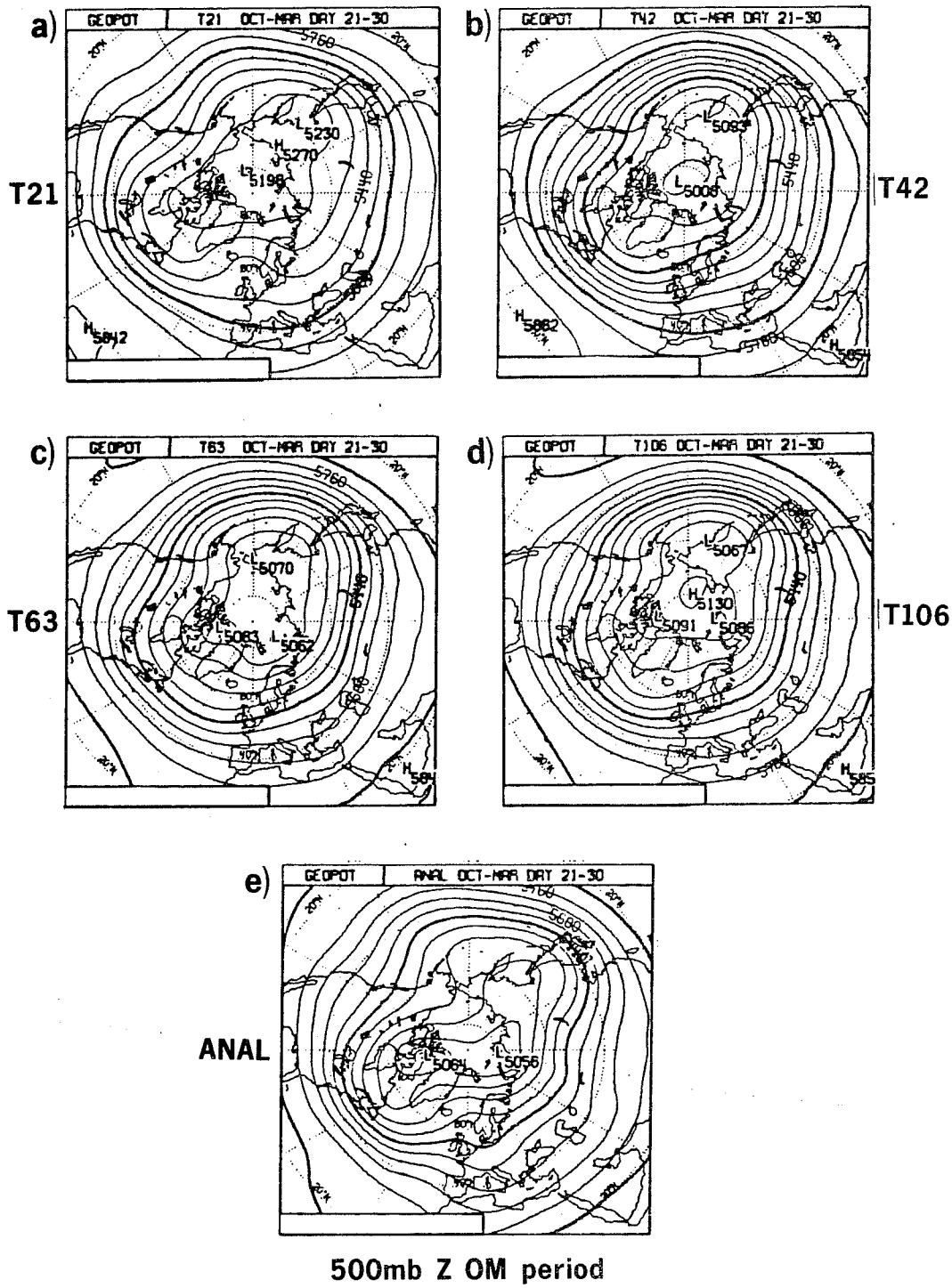

Fig. 8 500 mb mean geopotential field for the 12 integrations, days 21-30, in the OM period. a) T21; b) T42; c) T63; d) T106; e) analysis.

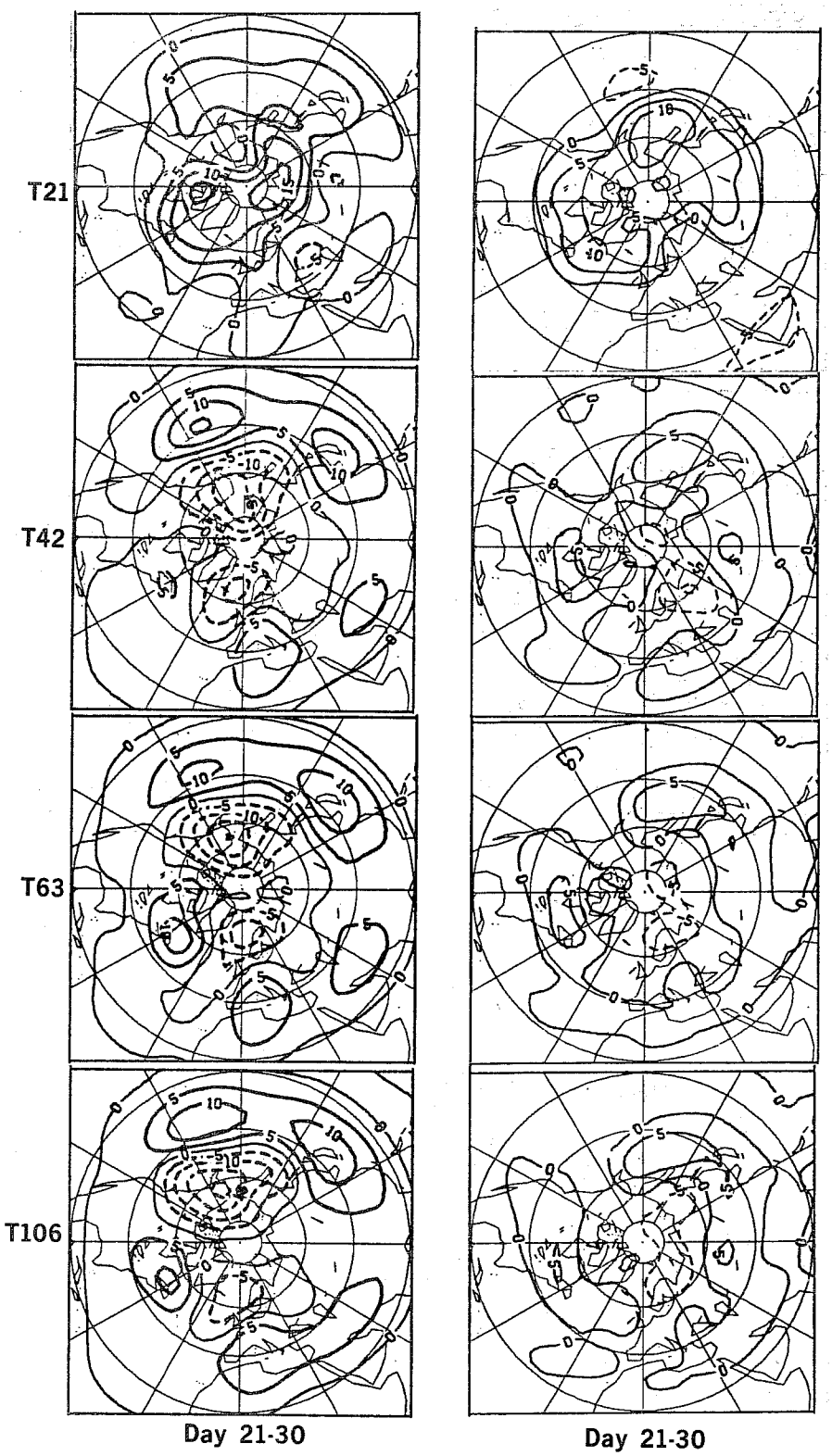

Fig. 9 500 mb geopotential height Northern Hemispheric mean errors for days 21-30 .
Top to bottom: T21, T42, T63 and T106.
Left: OM period, right: AS period.
Contour intervals every 5 dam.

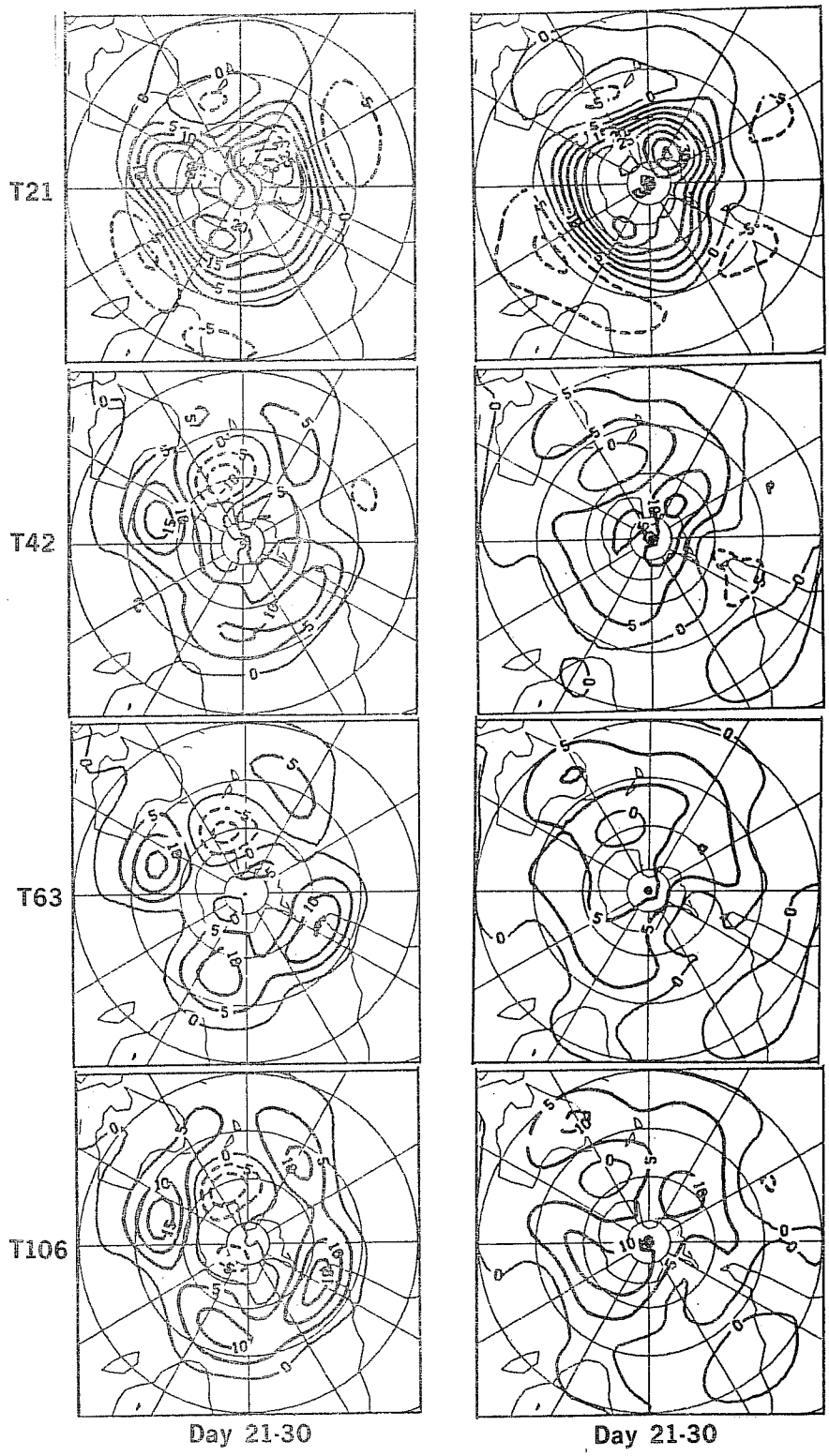

Fig. 10 As Fig. 9, but for the Southern Hemisphere.

20 to 30. The general character and structure of the errors, however, has changed less. We still find a large north-south negative-positive dipole over the Central Pacific and a similar, weaker one, over the E-Atlantic and W-Europe, all corresponding to weakening of the planetary waves and strengthening of the westerly jet in mid-latitudes.

While, the (mainly negative) error maxima at high latitudes in the earlier operational integrations were dominated by a combination of wavenumbers 2 and 3, now wavenumber 1 and 2 seem to dominate. The low-latitude positive maximum over the Central Pacific (see Fig. 1 of Wallace et al., 1983) is now broken into two weaker features. Since one still cannot discount the representation of orography as a possible contributor to such errors, it will be interesting to compare these results with those of the second year of experimentation, carried out with a model that includes parametrization of GWD and with 19 levels. Results of diagnostic work on operational 10-day medium-range forecasts indicate that the general character of SEs is left unchanged by increased vertical resolution (Simmons et al., 1987).

The NH extratropical 500 mb mean errors show a very large dependence upon season, which is larger than the dependence upon model resolution (excluding T21). For the SH (Fig. 10), although the seasonal dependence is much less marked, it is still larger than the effect of resolution, if we exclude again T21. The T21 climate of the SH suffers from particularly large deficiencies, as has been indicated above. The pole-to-equator geopotential height gradient is severely underestimated, together with the strength of the zonal flow. For a discussion of the possible causes, see again Miller et al. (1988). It is, however, interesting to notice how the mid-latitude SH time-mean errors for the last 10 day period of the AS season seem to have somewhat decreased with respect to the preceding 10-day interval, indicating a much faster levelling-off time. This could be related to the much lower planetary scale wave activity of the SH and to the consequently lower large-scale eddy mean flow interactions.

Such differences can be put into perspective by comparing the error maps with the full field maps which shows that the error centres move with the seasonal (and hemispherical) displacements of the main troughs and ridges of the stationary eddies. This large dependence of the mean errors upon flow regime

is reflected in the comparison between the mean errors and the RMS errors shown in Fig. 11 (for the T106 and T63 model, day 21-30 only), together with the ratio of mean error variance to RMS error variance. The contribution of the mean error variance to the total error variance is quite modest in mid-latitudes during the earlier forecast periods (not shown) but it is fairly large during the last 10-day period. This shows that, even within the same season, the time-mean errors depend considerably on the single forecasts realization and therefore, presumably, on the particular large-scale synoptic situation. This lends support to the idea that prognostic information about the forecast skill can be derived from the Großwetterlage of both initial conditions and forecast evolution (Palmer and Tibaldi, 1987).

The similarity of the mean (and, to a lesser extent RMS) errors at T63 and T106 model resolutions is an important result, but it is not yet fully understood. Previous experience suggests that extrapolation of such behaviour to higher model resolutions is quite unjustified. On the other hand, if the speculations on the role of baroclinic eddies on the climate drift have any basis, this probably suggests that either further resolution increases or a fundamental change in physical parametrization (further to envelope orography, GWD and even a possible reduction in free-atmospheric vertical diffusion) are needed to reduce the patterns shown in Figs. 9 and 10.

3.3 The mean errors in the tropical regions

This subsection will be devoted to a diagnosis of the mean model errors in the tropical regions. Since the tropical dynamics are dominated by moist processes with approximate balance between heating and vertical velocity, we will start by looking at the field of velocity potential.

Fig. 12 shows the velocity potential (χ) for the AS period for the last ten days of integration at 850 mb and 200 mb levels. The full χ field for T106 and T21 model resolutions and for the observed analysis are shown. The T63 and T42 fields are very similar to T106. It is known from the operational medium-range forecasts that the model represents the low-level convergence pattern fairly satisfactorily during the first ten days. The same is not true for the divergence at upper levels, where the general pattern of the Walker circulation dipole is satisfactory but its strength is already severely underestimated during the first ten days of integration. During the last ten

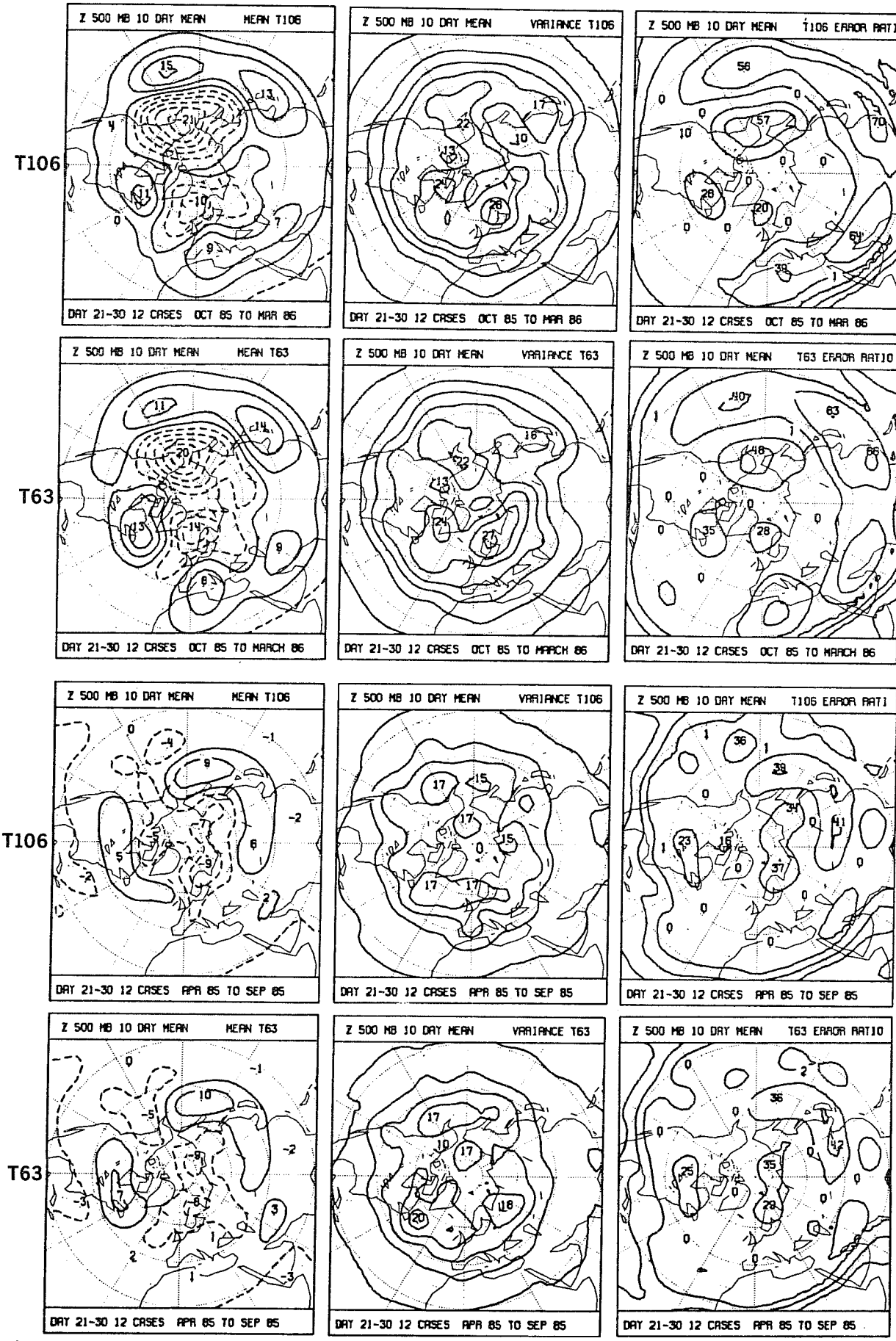

Fig. 11 Day 21-30 500 mb geopotential height errors for the T106 and T63 models. Top six panels: OM period. Bottom six panels: AS period. Left: mean errors. Centre: RMS errors. Right: percent ratio of mean error variance to total error variance. Contour interval is 5 dam for left and centre panels and 20% for right panels.

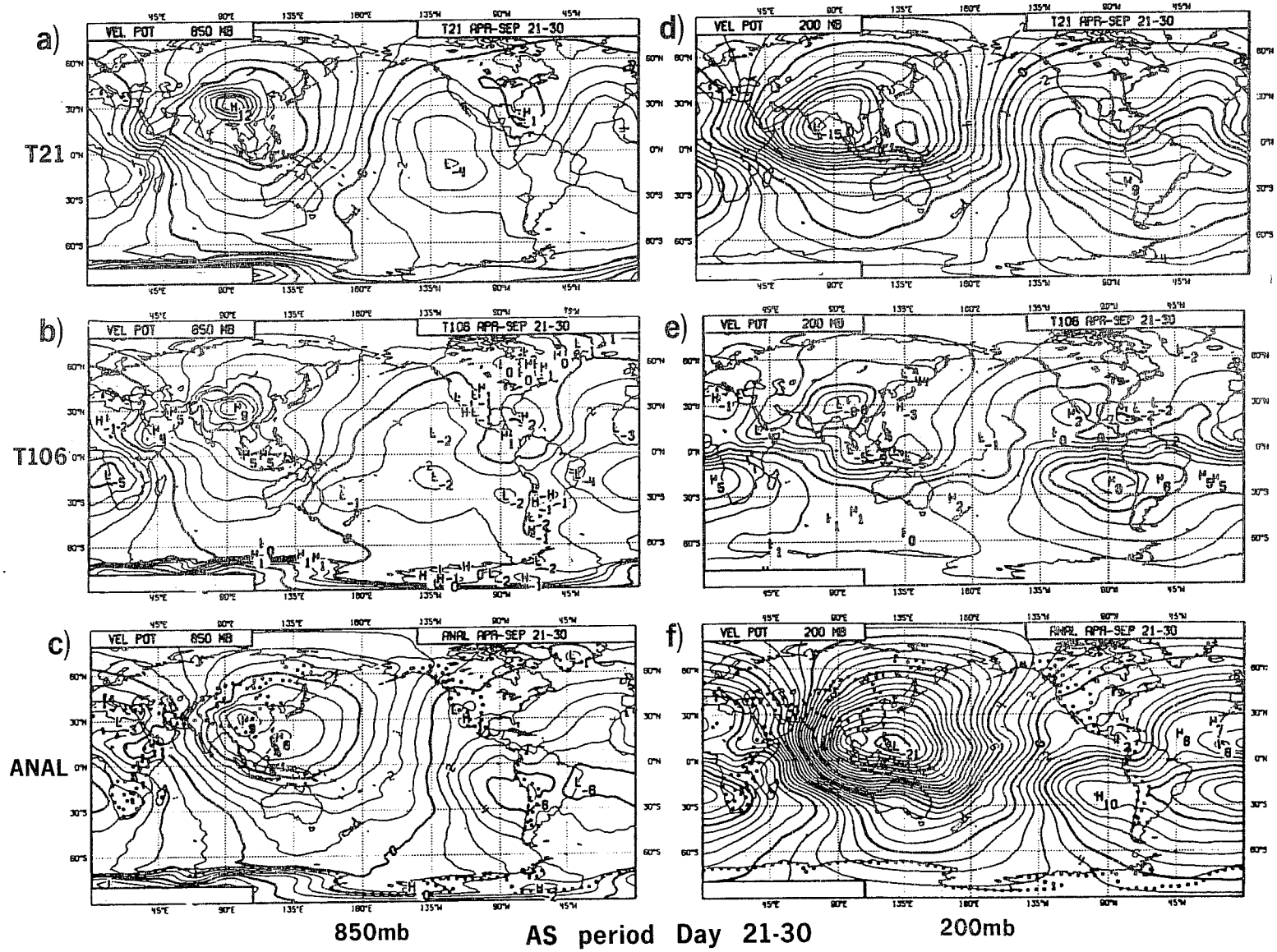

Fig. 12 Mean dd 21-30 fields of velocity potential (χ) for the AS period.
Top: T21. Centre: T106.
Bottom: analysis. Left: 850mb. Right: 200 mb.

days the tendency seen is the first ten days has been maintained and intensified, showing a low-level convergence pattern with an underestimated intensity (the excessively concentrated and intense maxima are mostly due to values extrapolated under high ground). The upper level divergence, furthermore, is very severely under-estimated, with maxima of the velocity potential less than 40% of the observed values (see also Heckley, 1985).

Since the low-level mass convergence in the model is not as poorly represented as the upper-level divergence, while mass is conserved, we must deduce that the upper level model divergence is not only underrepresented, but must have an erroneous vertical structure.

Fig. 13 shows an example of a vertical profile of divergence integrated over the Indonesian region, as shown in the inset. This figure was derived from the extended-range forecasts for the October 1986-March 1987 period and shows 'instantaneous' profiles for the forecast days 10, 20 and 30. It is clear that the upper-level divergence is severely underestimated by day 30, and has an erroneous vertical profile. It seems that the error at 200 mb is saturated between days 20 and 30.

The general weakening of the convectively driven Hadley (and Walker) circulations shown by the model is confirmed by the latitude-height cross sections of Fig. 14 which show the $[\bar{\omega}]$ and $[\bar{v}]$ full fields at day 21-30 mean for the T106 model and for the verifying analyses. The main characteristics of the model mean errors in these fields are: the circulation is substantially weaker, the cells are displaced polewards in both hemispheres and the vertical structure of the upper-level Hadley return flow is spread out, instead of being concentrated around the 200 mb level (see above). The poleward shifts of the meridional circulation cells (consistent with the poleward shift of the jetstreams) is much larger during the OM period (for both hemispheres) and weaker in the AS period (and NH only).

4. OBJECTIVE SKILL SCORES AT THE EXTENDED RANGE

We now turn to the objective evaluation of the skill of the model integrations, considered as deterministic extended range forecasts. We start by briefly reviewing the skill scores used. We then show the daily skill

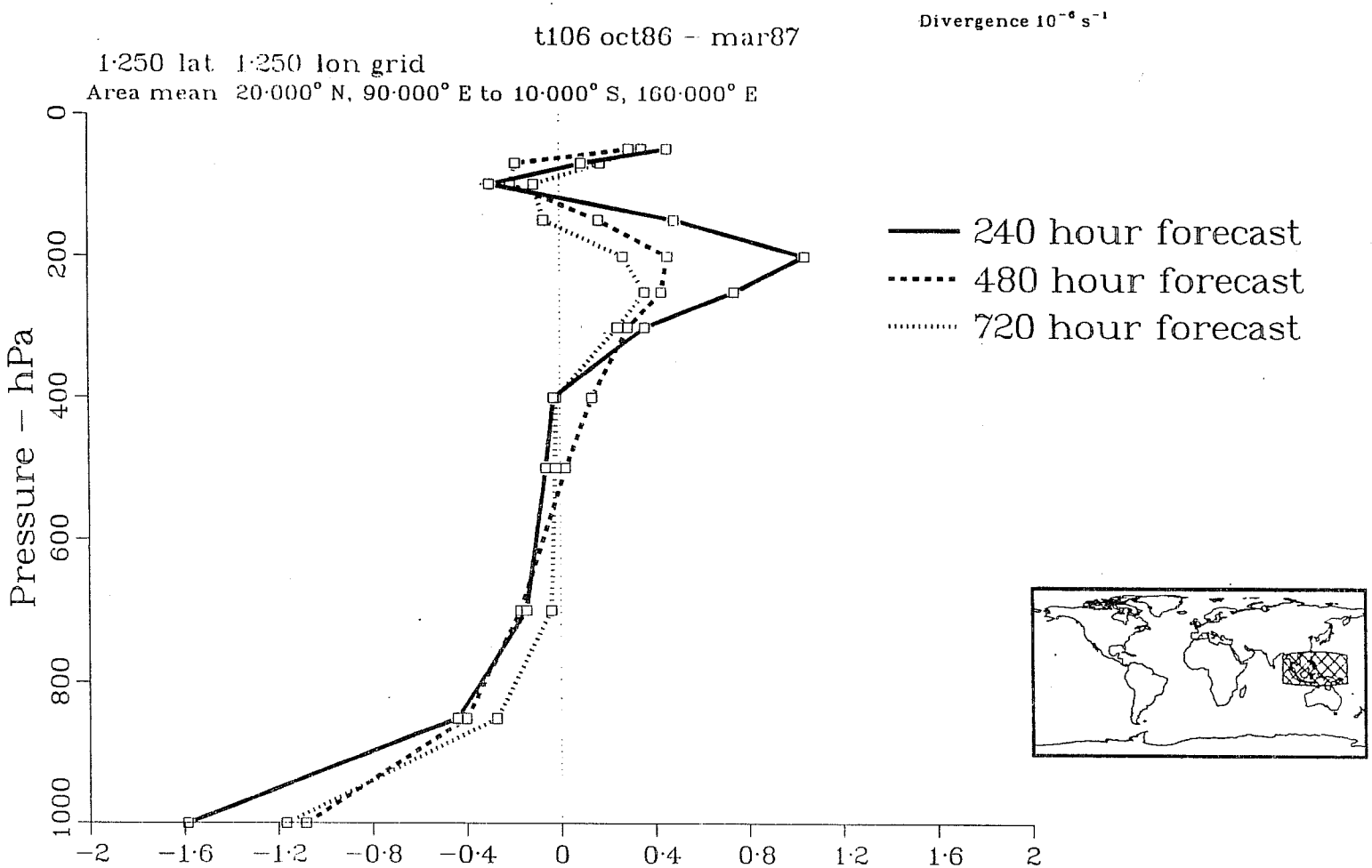

Fig. 13 Vertical profile of divergence, integrated in the hatched area shown in the small panel.
Full line: day 10; dashed line: day 20; dotted line: day 30.
Averaged over 12 cases of extended range forecasts, October 1986 - March 1987.

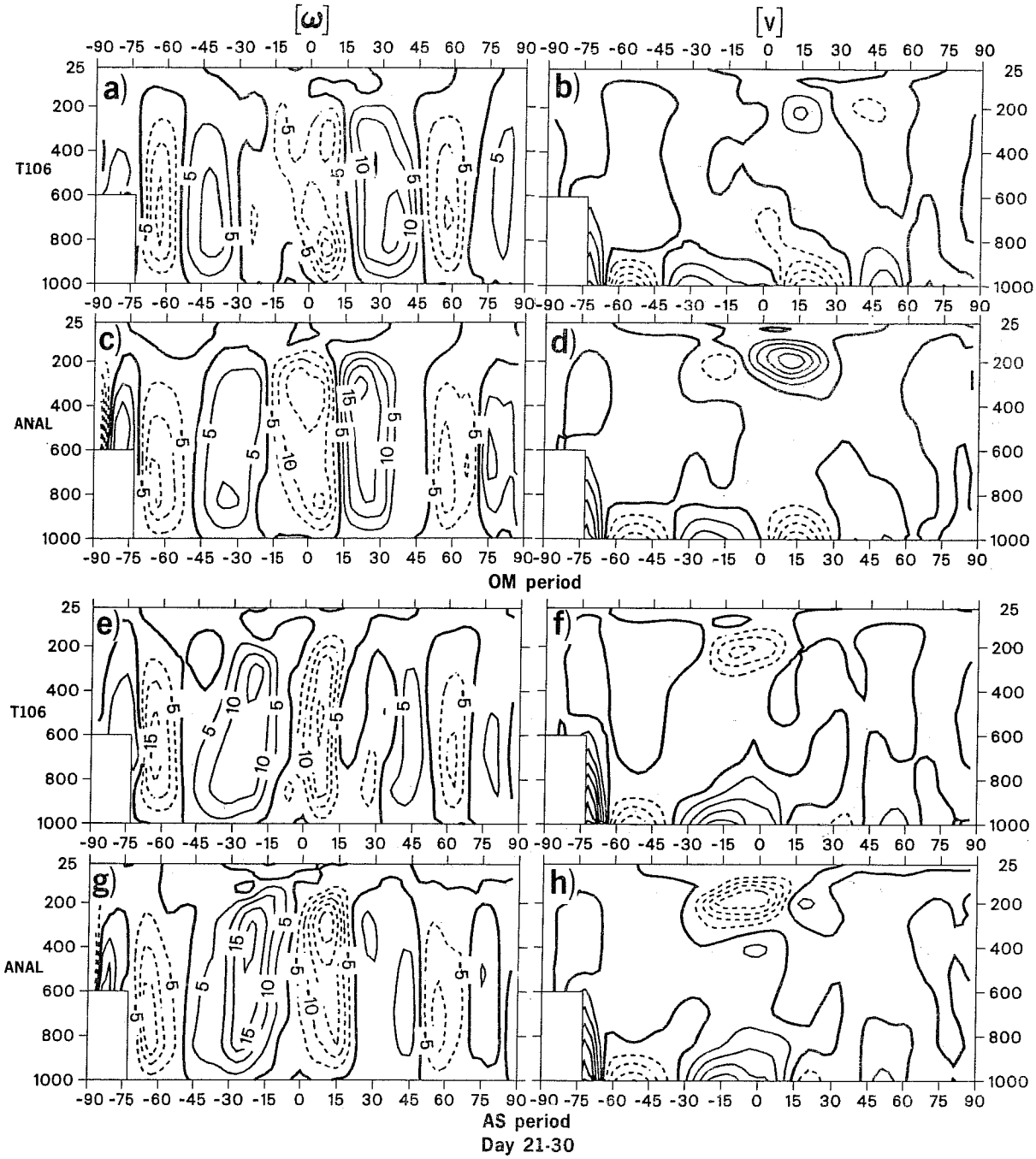

Fig. 14 Latitude-height cross sections of $[\bar{\omega}]$ (left panels) and $[\bar{v}]$ (right panels). Top to bottom: T106 OM period; analysis OM period; T106 AS period; analysis AS period. Units: $mPas^{-1}$ and ms^{-1}

scores and the 10-day mean scores. Only hemispheric scores will be discussed here. For a more comprehensive comparison of these results with similar sets of experiments (e.g. Miyakoda et al., 1986 and Mansfield, 1986) the reader is referred to Hollingsworth et al. (1987).

4.1 Definition of the skill scores

Two quantities, anomaly correlation coefficient (ACC) and RMS error, are usually employed to evaluate objectively the success of numerical forecasts. We have calculated both such quantities for all resolutions considered in our 30-day forecast set and for the following meteorological parameters: geopotential height at 1000, 500 and 300 mb, and temperature at 850, 500 and 300 mb. We will, however, limit our discussion to 500 mb height and 850 mb temperature.

Both skill scores have been computed on T21 spectrally truncated fields evaluated on a 3.75° regular latitude/longitude grid. Hemispheric scores are defined over the area enclosed between 22.50° and 86.25°N. In the following subsections the scores are computed on daily fields, 10-day mean fields or 30-day mean fields and the performance of the model will be discussed in terms of NH skill scores and their dependence on resolution and season.

4.2 The daily scores

Fig. 15 shows RMS error and ACC daily scores for 500 mb height and for both OM and AS periods. In addition to the model scores, the diagrams also show scores for persistence of full fields. The model scores, representing the average of 12 model integrations for each period, are computed over the Northern Hemisphere and are averaged within the sample of 12; the Fisher z-transform (e.g. Morrison, 1983) is used to average the anomaly correlations. We shall consider here Fig. 16 as well, where dotted curves show the departures between successive T106 or T21 forecasts, with initial conditions separated by 24 hours (the "forecast spread"). These curves are, therefore, the mean over 6 pairs and are computed between forecast verifying at the same time. In the RMS error diagrams, the lower horizontal dotted line represents the climate norm for the corresponding field. In the ACC error diagrams, the 60% level is indicated by the dashed line. It should be noted that, with a perfect model, not only persistence but also forecast RMS error (and RMS forecast spread) should asymptote to $\sqrt{2}$ times the climate norm (e.g.

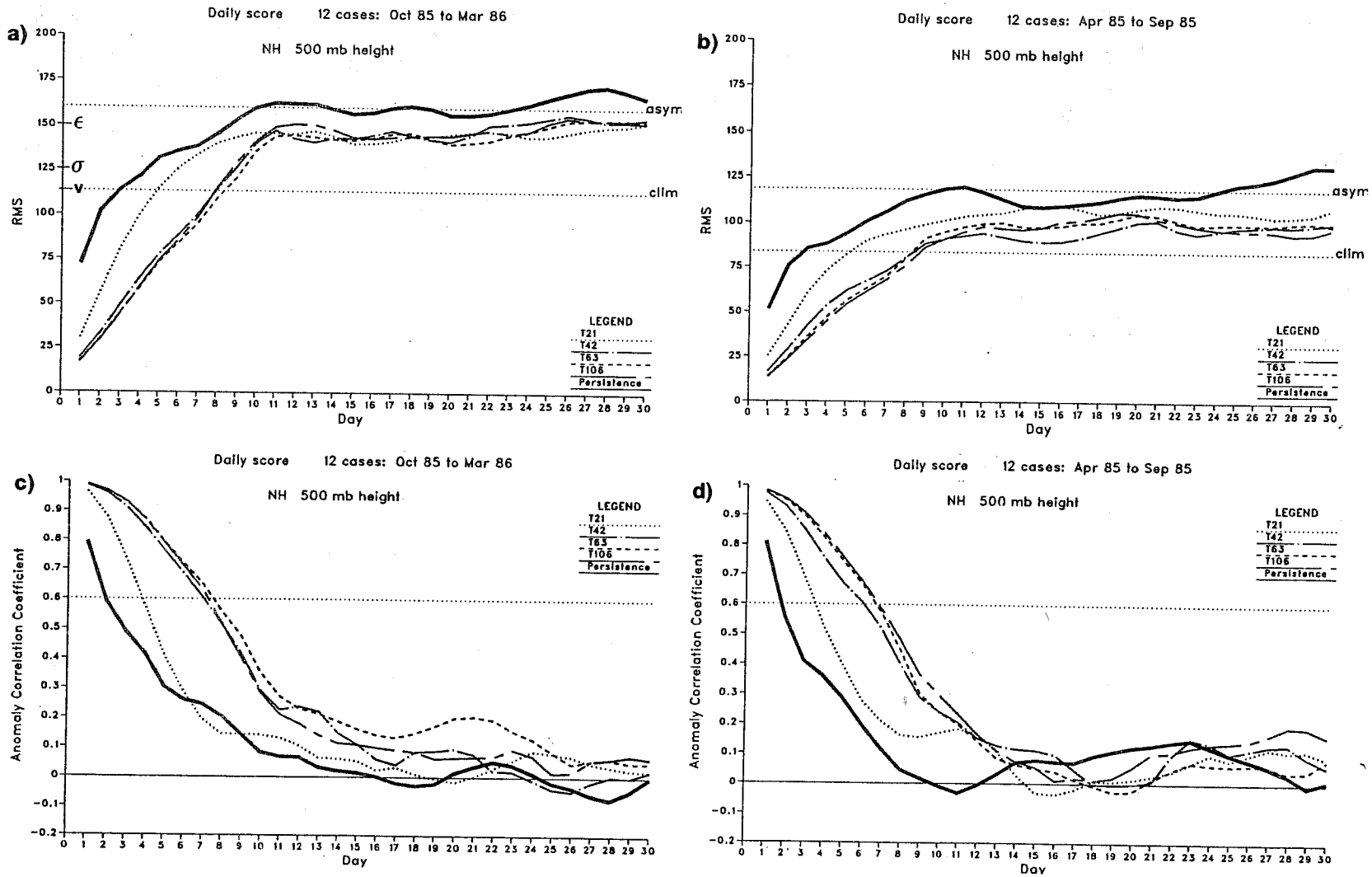

Fig. 15 RMS error (top) and ACC (bottom) of daily northern hemispheric 500 mb height forecast fields.
Left: OM period. Right: AS period.
(For the meaning and use of ε, σ and v, see the Appendix.)

Miyakoda et al., 1972). Indeed persistence does asymptote to approximately the expected value (upper dotted line), but forecast error does asymptote to a lower level. The fact that the model has systematic errors will tend to increase the RMS forecast error asymptote but will have no effect on the RMS forecast spread. Conversely, the fact that the model has less time variability than the real atmosphere will affect both RMS error and RMS spread, by lowering both their asymptotic values (e.g. Hollingsworth et al., 1987). Simple considerations about the asymptotic levels reached by the various quantities plotted in the diagram of Fig. 15a (see Appendix) suggests that, for daily fields, the variance of the "winter" 500 mb height mean model error is about $\|\bar{F}\|^2 = (50m)^2$ and that the transient variance of the higher resolution models is approximately 35% lower than observed (and for T21 is more than 40% lower). By and large, this means that the lack of transience is the dominating effect (over the mean model error) in determining the asymptotic level of the model RMS error, for both daily and 10-day mean fields. We will see later how such considerations carry through to 10-day mean RMS errors.

An analysis of Fig. 15 reveals that the complete loss of predictability of 500 mb geopotential height (measured as the time it takes to the RMS error to reach its asymptotic value) is reached by higher resolution models around day 12, with little or no difference due to either model resolution or season. For T21 model asymptotic value is reached 2-3 days earlier. The T106 model appears marginally worse than the T63 in the OM period and marginally better during the AS period. There is also some evidence that the T106 model has slightly higher mean errors (see also Section 3.2) and slightly lower loss of transience than the T63. The time at which the higher resolution model RMS error reaches the climate error (around day 7 to 8) also coincides fairly well with the time it takes for the ACC to reach .6 (around day 6 to 7) and is consistent with the medium-range forecasting experience.

Since we know from other, more extensive, independent comparison studies (e.g. Simmons et al., 1987) that T106 is measurably more skillfull than T63 in the medium range, irrespective of season, we must conclude that our sampling is too small. An indirect confirmation that there is a sampling problem is to be found when comparing the T106 RMS spread curve of Fig. 16a with a similar curve evaluated on a much larger sample of 100 medium range forecasts

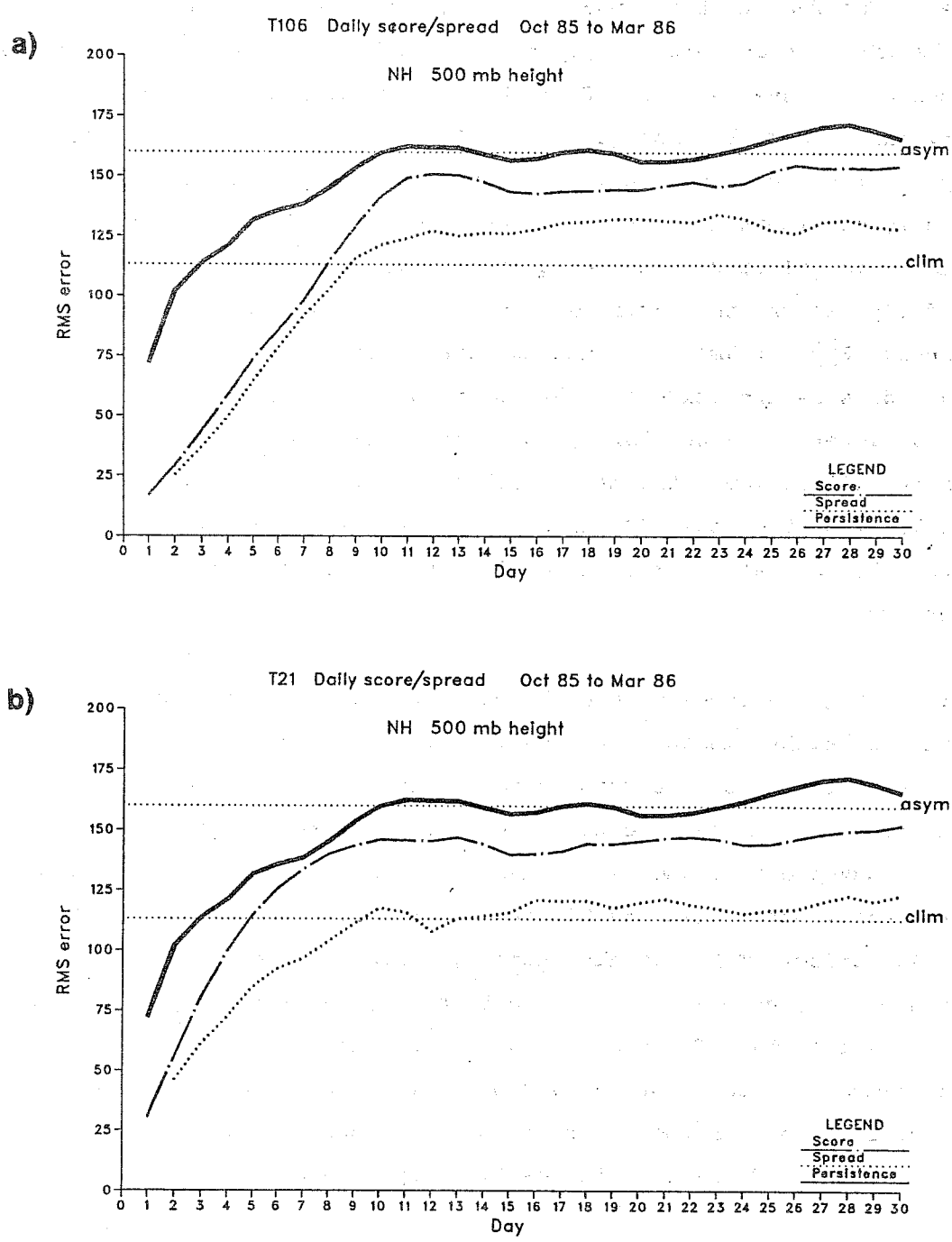

Fig. 16 RMS error of daily forecast spread for a) T106 and b) T21 in the season OM. Spreads are computed between successive forecasts, initiated 24 hours apart (six pairs per six-month period).

performed in the same period and using the same model, see Hollingsworth et al. (1987). The two curves are quite different, the one shown here growing much faster.

4.3 The ten-day mean scores

Fig. 17 is the exact counterpart of Fig. 15, but for 10-day mean fields and with addition of persistence of anomalies. Some interesting aspects of model behaviour may be discerned. Firstly, the predictability or, more precisely, the "forecastability" of the low-frequency part of the atmospheric behaviour is only marginally higher than that of the total fields. The 500 mb height RMS error crosses the climate norm only 1 to 2 days later than for instantaneous fields. What this implies for atmospheric predictability is open to speculation. The substantial equivalence of the two T63 and T106 models in terms of extended-range objective skill scores is also confirmed, together with their relative opposite behaviour in the two "seasons".

An obvious change of behaviour can be seen in the relative differences between the T106 and T21 forecast spread curves (Fig. 18). The T106 now shows a considerably larger variability than the T21 model, indicating that most of the difference in variability shown in the total curves of Fig. 15 was due to low frequency rather than day-to-day variability. Moreover the effect is reduced towards the end of the integration period. It is, therefore, likely that sampling again has to be held responsible for such a difference (in this case the sample has only six members).

Interestingly, the 850 mb temperature ten day means (Fig. 19) appear to be measurably more predictable than 500 mb height (ten days to cross the climate norm, as opposed to eight for 500 mb height). This effect is only present in the OM period (extended NH winter) and does not occur during the AS extended NH summer. The mid-latitude ocean surface thermal forcing is much larger in winter than in summer and so could be a possible reason for the increased predictability of the low-frequency variability of 850 mb temperature. Forcing from the time-evolving land surface temperature could also contribute to this effect since it is a model prognostic variable. One might expect, however, such effect to be larger in summer than in winter. It is worth mentioning that the asymptotic values for 850 mb temperature ACC (for both OM and AS periods) is considerably higher than zero, unlike, for example for

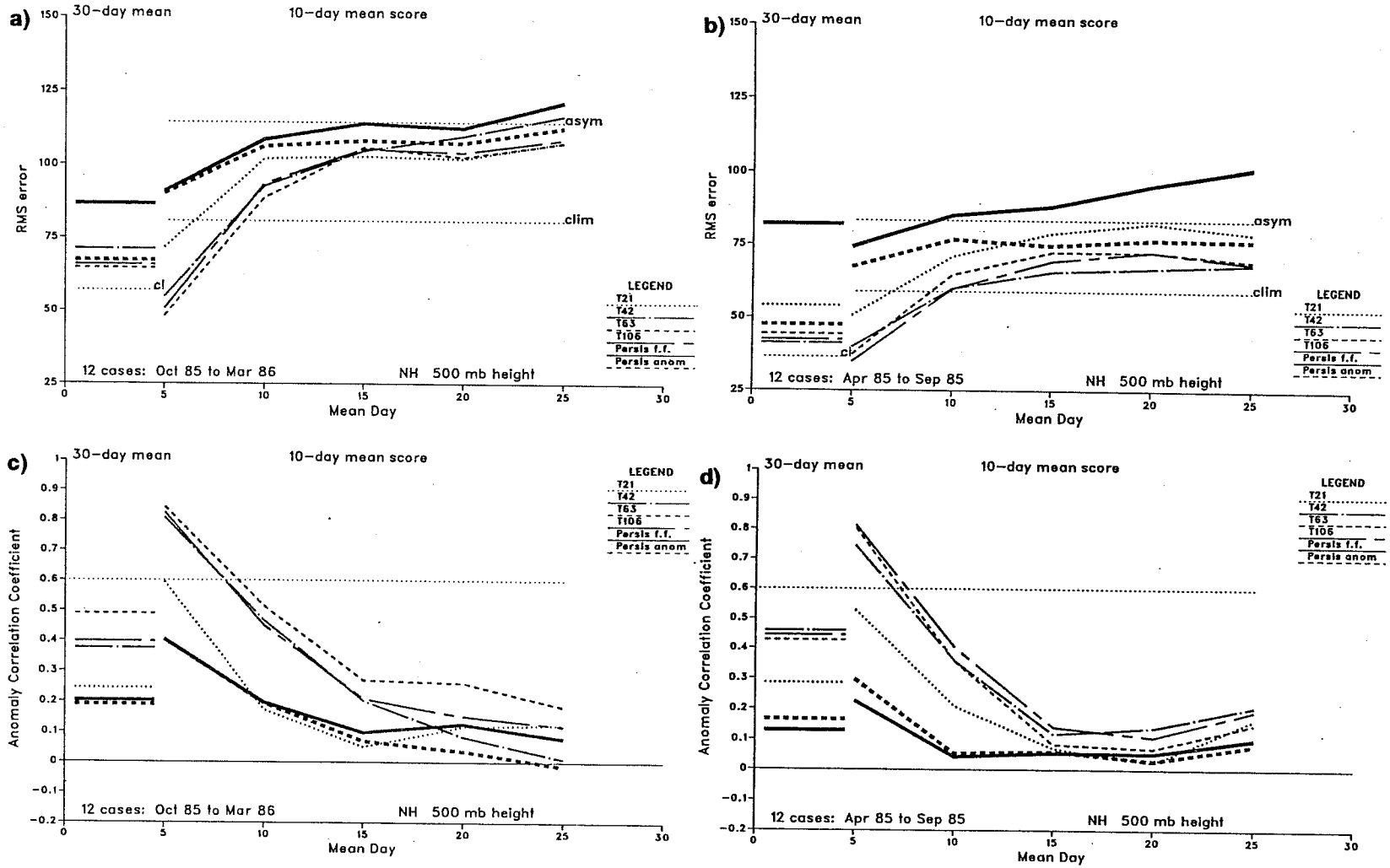

Fig. 17 As Fig. 15 but for 10-day running mean fields. Thick dashed line is persistence of anomalies.

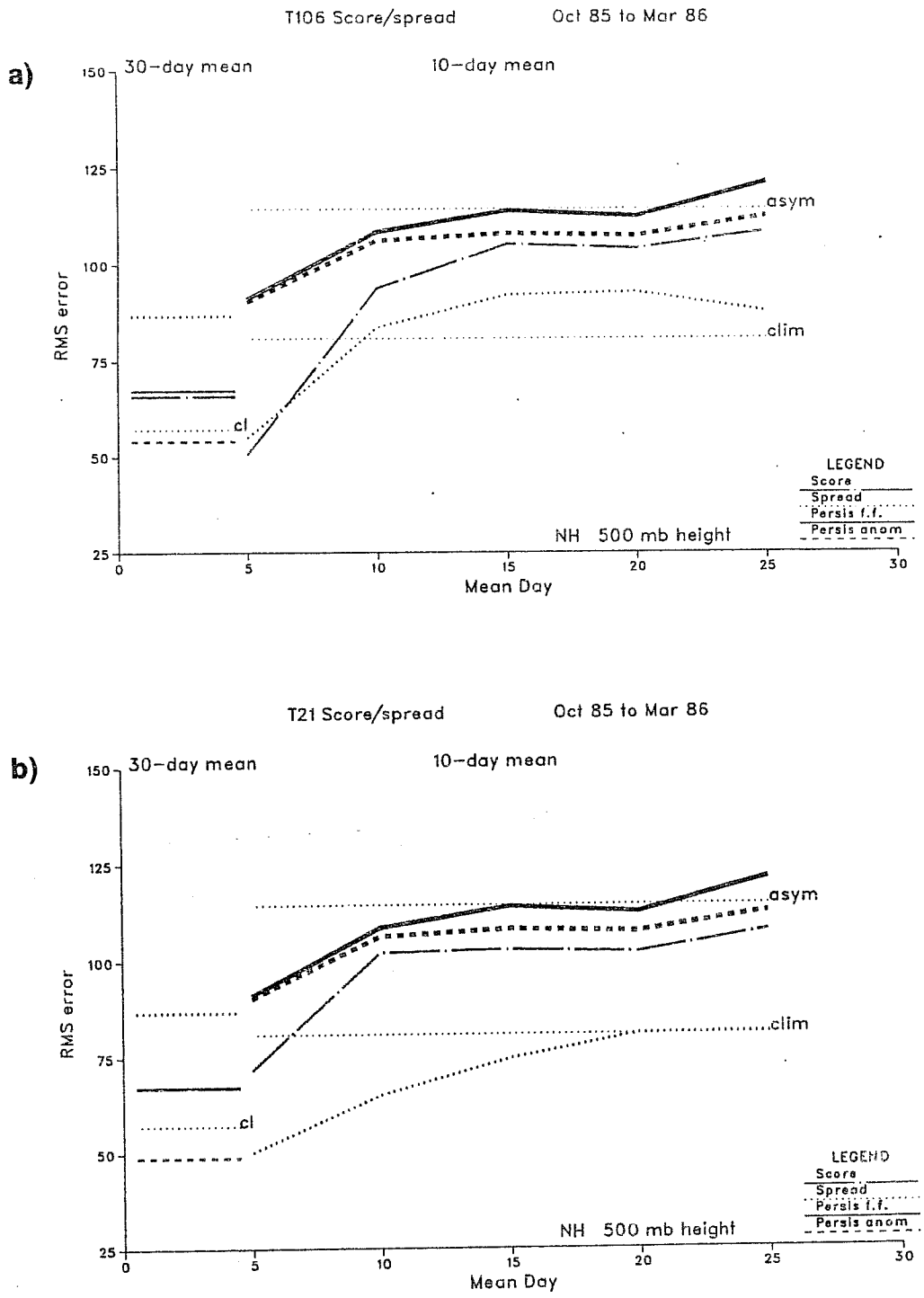

Fig. 18 As Fig. 16 but for 10-day means.

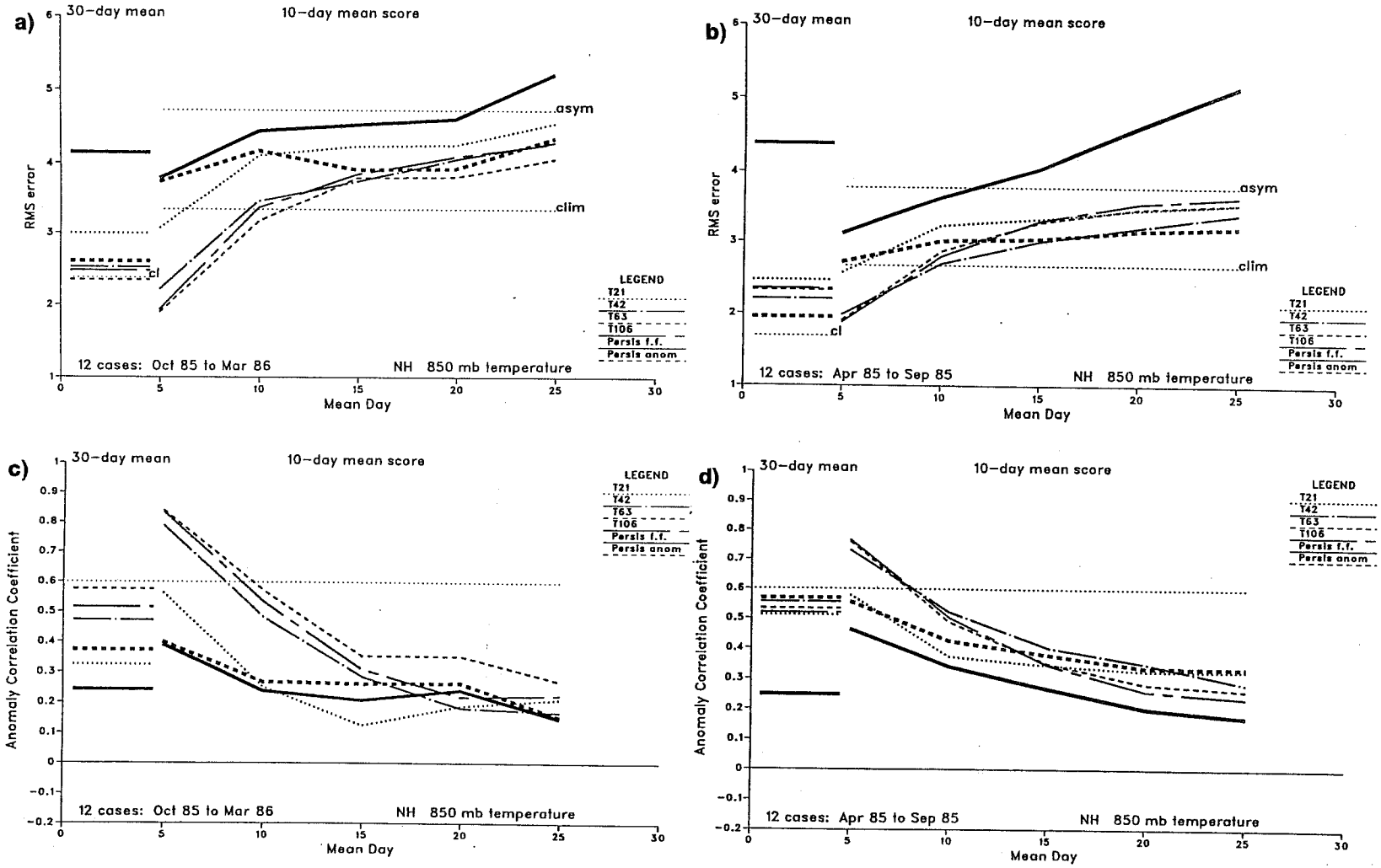

Fig. 19 As Fig. 17 but for 850 mb temperature.

500 mb height. This confirms similar results reported in Molteni et al. (1987). It should, however, be noticed that the same advantage of 850 mb temperature over 500 mb height ACC is also noticeable for persistence.

Using the results of the Appendix, and the asymptotic values of Fig. 17a, it can be seen that the estimated SE variance is $(38m)^2$ and the low-frequency model transient variance is 44% lower than observed. This indicates that the loss of transient eddy activity experienced by the model affects low-frequency variability even more than high-frequency variability.

5. CONCLUSIONS

The dataset of 30-day integrations at various model resolutions being accumulated at the Centre has allowed us to evaluate the climate drift of the operational model as a function of resolution and season. Furthermore, the same database can be used to investigate the forecast skill of the model beyond the currently accepted average limit for deterministic predictability.

The results can be summarized as follows:

- The zonal mean and mid-latitude eddy systematic errors show a strong dependence upon resolution at the lower end of the resolution spectrum (T21-T42). At the upper end (T63-T106) the dependence upon resolution is weak and shows indications of levelling off.

- Mid-latitude SEs (and time mean errors of a single realization) are largely dependent upon season and large scale Großwetterlage, indicating that some information about the initial conditions is retained almost throughout the integration-time, although not necessarily of a nature that can be exploited in practical forecasting. This shows promise for predictability (forecast skill) studies.

- The zonal mean tropospheric SEs in mid-latitudes can be mainly described as a poleward displacement of the main jet maxima, with an equivalent barotropic structure.

- The tropical SEs (and the mid-latitude stratospheric SEs) can be mainly described as a progressive loss of the observed vertical shear of the zonal wind (and thermal wind balance related temperature errors). Although the main contributions to such errors are probably to be found in the moist convection and radiation schemes, the structure of the errors is also consistent with the hypothesis that too much free-atmosphere vertical diffusion of momentum (and heat) is taking place.

- The mid-latitude SEs of the eddy flow can be mainly described as a progressive loss of stationary and transient eddy activity, with the model 500 mb height time mean errors having a winter asymptotic variance of about $(50\ m)^2$ and a transient variance more than 35% lower than observed for daily fields and 44% lower than observed for 10-day mean fields.

- Both the time-mean circulation and the eddy activity are much weaker than observed in tropical regions and the vertical structure of the (mainly diabatically forced) divergence field is erroneously represented by the model. Because of the strong feedbacks between the mean flow and the eddies it is very difficult to partition, even qualitatively, the responsibility for such loss between the moist convection parametrization schemes and the other model characteristics.

- The T106 daily 500 mb height RMS error during NH "extended winter" reaches it asymptotic value around day 12 ("complete" loss of instantaneous deterministic skill). Very similar values (relative to the respective climatic variances and, therefore, asymptotic values) are attained by both 500 mb geopotential height and 850 mb temperature RMS error, for both seasons. The effect of taking 10-day means of forecast field is to increase very marginally (1-2 days) the perceived "forecastability" of 500 mb height but of 2-3 day that of 850 mb temperature, possibly indicating a measurable beneficial effect of surface thermal forcing on atmospheric predictability.

- RMS forecast error of both daily 500 mb height and 850 mb temperature reaches the climate norm around day 7-8, by which time the anomaly

correlation coefficient has dropped to about 60%. These values are only slightly increased by 10-day mean averaging the fields with a somewhat better impact on 850 mb temperature. This, together with the comparisons with persistence forecasts (persistence, however, of full fields and not of the anomaly, which is a better zero-cost forecast) indicates obviously useful information on average until day 7-10, with some potentially useful information in ten-day mean fields until the day 6 to 15 averaging period.

- The 30-day mean forecast fields are almost consistently better than persistence, but how much this depends on the first 10 to 15 days of integration remains to be more accurately evaluated.

- The variability of extended-range forecast skill is large; larger even than medium-range skill. A major difficulty is undoubtedly the forecasting of blocking onset beyond the first few days of model integration time. The possibility of estimating a priori the skill of a forecast might be vital if extended range forecasts are to move to the forefront of operational weather prediction.

APPENDIX

A quantitative estimate of the model's transient variance

An estimate of the asymptotic values of RMS forecast error and RMS forecast spread (as can, for example, be obtained from Figs 24a and 26a for daily fields and 10-day mean fields respectively) together with the knowledge of the climate norm, makes it possible to estimate the mean hemispheric variance of the model's systematic error and the proportion of observed transient variance represented by the model (e.g. Hollingsworth et al., 1987).

If we define

A: analysed anomaly field

F: forecast anomaly field

$E \equiv F-A$: forecast error field

$S \equiv F_1 - F_2$: forecast spread field

$\overline{} \equiv$ expectation operator (ensemble mean)

$v^2 \equiv \overline{\|A\|^2}$: total climatic variance

and consider a large enough ensemble, for which

$\bar{A} = 0.$

If the model has a systematic error,

$\bar{F} \neq 0,$

and we can define

$F' = F - \bar{F}.$

Let us also consider a forecast time at which there is no more predictability (i.e. the RMS has reached its asymptotic value). Then

$$\overline{A \cdot F'} = 0 \; ; \quad \overline{F'_1 \cdot F'_2} = 0$$

We then define forecast error and forecast spread variances respectively as:

$$\varepsilon^2 = \overline{\|E\|^2} = \overline{(F-A) \cdot (F-A)} = \overline{\|F\|^2} + \overline{\|A\|^2} + \overline{2F \cdot A}$$

$$\sigma^2 = \overline{\|S\|^2} = \overline{\|F_1 - F_2\|^2} = \overline{\|F'_1 - F'_2\|^2} =$$

$$= \overline{\|F'_1\|^2} + \overline{\|F'_2\|^2} - 2\,\overline{F'_1 \cdot F'_2} = \overline{2\,\|F'\|^2}$$

But

$$\overline{\|F\|^2} = \overline{(\bar{F} + F') \cdot (\bar{F} + F')} = \overline{\|\bar{F}\|^2} + \overline{\|F'\|^2} + \overline{2\bar{F} \cdot F'} =$$

$$= \|\bar{F}\|^2 + \overline{\|F'\|^2}$$

that is, the perceived model variance is the sum of a "true" part, $\overline{\|F'\|^2}$, that represents real time-transience, and a "false" part, $\|\bar{F}\|^2$, that only reflects the existence of a model mean error. Of course, we also know that the "true" model variance is likely to be smaller than the observed one. It is, therefore, interesting to evaluate their ratio.

We can see that

$$\overline{F \cdot A} = \overline{(\bar{F} + F') \cdot A} = \bar{F} \cdot \bar{A} + \overline{F' \cdot A} = 0$$

So, if we define

$$\alpha^2 = \overline{\|F'\|^2} \,/\, \overline{\|A\|^2}$$ ratio between "true" model variance and observed variance

and

$\beta^2 = \|\bar{F}\|^2 / \overline{\|A\|^2}$ ratio between mean square amplitude of systematic error ("false" model variance) and observed variance,

we have

$$\varepsilon^2 = v^2 (1 + \alpha^2 + \beta^2)$$

$$\sigma^2 = 2v^2\alpha^2$$

Concluding:

$$\alpha^2 = \frac{\sigma^2}{2v^2} \qquad \text{and} \qquad \beta^2 = \frac{\varepsilon^2}{v^2} - 1 - \alpha^2$$

If we now evaluate approximately ε and σ from Fig. 15 (500 mb height RMS NH "winter" errors for daily maps) and from Fig. 17 (500 mb height RMS NH "winter" errors for 10 day mean maps) we have, for daily values:

$v = 110$m		$\alpha^2 = .65$
$\varepsilon = 150$m	and therefore:	$\beta^2 = .21$
$\sigma = 125$m		$\|\bar{F}\| = \beta v = 50$m

and, for 10-day mean values

$v = 80$m		$\alpha^2 = .56$
$\varepsilon = 107$m	and therefore:	$\beta^2 = .23$
$\sigma = 90$m		$\|\bar{F}\| = \beta v = 38$m

β^2 is statistically significant at the 95% confidence level if it exceeds $2/n$, where n is the sample size, in our case 12. β^2 has, therefore, to exceed .17 to be highly significant.

References

Arpe, K., 1988: Planetary-scale diabatic forcing errors in the ECMWF model. Proceedings of the ECMWF Workshop on Diabatic Forcing, 30 November - 2 December 1987.

Blondin, C. and H. Böttger, 1987: The surface and sub-surface parametrization scheme in the ECMWF forecasting system. ECMWF Tech. Memo. No. 135.

Brankovic, C., 1986: Zonal diagnostic of the ECMWF 1984-85 operational analyses and forecasts. ECMWF Tech. Rep. No. 57.

Heckley, W.A., 1985: Systematic errors of the ECMWF operational forecasting model in tropical regions. Quart.J.Roy.Meteor.Soc., 111, 709-738.

Hollingsworth, A., U. Cubasch, S. Tibaldi, C. Brankovic, T.N. Palmer and L. Campbell, 1987: Mid-latitude atmospheric prediction on time scales of 10-30 days. In "Variability in the atmosphere and oceans", Ed. H. Cattle, 117-151.

Mansfield, D.A., 1986: The skill of dynamical long-range forecasts, including the effect of sea surface temperature anomalies, Quart.J.Roy.Meteor.Soc., 112, 1145-1176.

Miller, M., T.N. Palmer and R. Swinbank, 1988: Orographic gravity wave drag: its parametrization and influence in general circulation and numerical weather prediction models. To appear in Met. and Atmos. Phys.

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

Miyakoda, K., J. Sirutis and J. Ploshay, 1986: One-Month Forecast Experiments. Mon.Wea.Rev., 114, 2363-2401.

Molteni, F., U. Cubasch and S. Tibaldi, 1987: Monthly forecast experiments with the ECMWF spectral models. Proceedings of the ECMWF Workshop on predictability in the medium and extended range. ECMWF, Shinfield Park, Reading, U.K.

Morrison, D.F., 1983: Applied linear statistical methods. Prentice-Hall, Englewood Cliffs, N.J., 562 pp.

Palmer, T.N. and S. Tibaldi, 1987: Predictability studies in the medium and extended range. ECMWF Tech Memo No. 139.

Simmons, A.J., L. Dell'Osso, M. Jarraud, J.M. Hoyer, G. Sakellarides, 1987: Some results from studies of increased horizontal and vertical resolution. ECMWF Tech. Memo. No. 136.

Tiedtke, M., W.A. Heckley, J. Slingo, 1988: Tropical forecasting at ECMWF: The influence of physical parametrisation on the mean structure of forecast and analyses. Quart.J.Roy.Meteor.Soc., 114, 639-664.

Wallace, J.M., S. Tibaldi and A.J. Simmons, 1983: Prediction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. Quart.J.R.Meteor.Soc., 109, 683-717.