TECHNICAL MEMORANDUM

152

Extended range predictions with ECMWF models. II. Influence of horizontal resolution on systematic error and forecast skill

S. Tibaldi, T.N. Palmer, C. Brankovic and U. Cubasch

Research Department

June 1989

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

ABSTRACT

The influence of horizontal model resolution on model systematic error and skill scores are examined for both winter and summer seasons from a set of extended-range integrations of the European Centre for Medium Range Weather Forecasts (ECMWF) numerical weather prediction model. The forecast data is composed of 24 30-day integrations, two per month from successive 12 UTC operational analyses, from April 1985 to March 1986 at four (triangular truncation) horizontal spectral resolutions of T21, T42, T63 and T106.

Sytematic errors in wind, temperature, heat and momentum fluxes are discussed, both in the tropics and extratropics. Dynamical reasons for differences between various model resolutions, and between simulations and observations, are proposed. Diagnosis of model systematic error is aided by comparing zonal-mean diagnostics of the observed flow in the NH winter and SH winter. The impact of resolution on regional simulations of rainfall during the Indian and African monsoon are also discussed.

In the extratropical troposphere, the behaviour of the T21 is quite distinct from the higher resolution models. The southern hemisphere flow is weak, and in both hemispheres horizontal momentum fluxes are severely underestimated. The systematic error of the T42, T63 and T106 models are quite similar to each other in the extratropical troposphere. Errors at these higher resolutions resemble the observed difference (in zonal mean wind, temperature and eddy fluxes) between diagnostics in the northern hemisphere winter and southern hemisphere winter, suggesting that orographic forcing in this version of the ECMWF model is inadequate, despite the inclusion of envelope orography.

In the extratropical stratosphere, the behaviour of the T21 and T42 models resemble each other, both having a cold polar bias, apparently associated with inadequate dynamical heating and therefore relaxation to radiative equilibrium. This in turn is related to insufficient Rossby wave focussing into the polar vortex at T21 and T42. The T63 and T106 resolution integrations have much smaller stratospheric cold pole biases.

In the tropics, there are serious systematic errors at all resolutions; however, some of these mean errors increase with increasing resolution. These include the global divergent and nondivergent wind errors, and more regional simulations such as the monsoon flow, where simulation of diabatic heating associated with convective activity is crucial. Since the local values of moisture flux convergence at T106 are effectively much noisier than the larger-scale values of such fluxes used to drive the convection scheme at T21, it is

speculated that diabatic heating fields are more likely to project onto the relevant tropical meteorological modes at low model resolution. On the other hand, it is shown that the ability to resolve local orographic features is important in accurately simulating local tropical precipitation maxima over land.

Despite having smaller NH systematic errors, 10-day mean T21 model anomaly correlation coefficient scores were worse than those of higher resolution models, at least up to day 20. However, there was no indication that the higher resolution models had significantly different extended-range skill characteristics. On the other hand, the increase of tropical systematic error with increasing horizontal resolution may indicate that the potential of higher resolution models in extended range prediction may be underestimated in this study.

The asymptotic spread between forecasts initialised 24 hours apart indicated that all models substantially underestimated low-frequency variability (by up to 50% in the extended summer period). There was some evidence that the lower resolution forecasts had significantly lower values of asymptotic spread than higher resolution forecasts.

Bearing in mind the computational burden of longer timescale integrations with complex numerical models of the atmosphere, the evidence presented in this paper suggests that the T106 resolution may be unnecessarily high for both extended-range forecasting, and for climate simulation studies. However, at present, the cost of integrating at T63 resolution may also be excessive for many climate studies. It would appear from our analysis, that in the troposphere the behaviour of the T42 model is comparable with the higher resolution models, and is a satisfactory compromise for many purposes, particularly bearing in mind the impact of resolution on tropical systematic errors.

However, it would appear that the extreme sensitivity of the 1985/86 ECMWF model climate drift to resolution in the range between T21 and T42, and the apparent inability of the ECMWF T21 model to simulate the correct internal nonlinear dynamics of the extratropics, makes questionable its use for climate studies at resolution lower than T42.

1. INTRODUCTION

There is considerable interest in atmospheric prediction beyond the limit of instantaneous deterministic predictability, normally assumed to be about 2 weeks. Because of considerable improvements in the formulation of numerical weather prediction (NWP) models, the systematic component of forecast error often does not reach large amplitude until well into the extended range. Diagnosis of such error can therefore be facilitated by the study of extended-range integrations.

In addition, interest has being growing in the potential of monthly timescale weather prediction as an operational forecasting product (e.g. the companion paper Palmer et al, 1989, hereafter referred to as I, and references therein), where even modestly skilful extended-range forecasts of time-mean fields may nevertheless be of considerable value. The development of probabilistic extended-range forecasts using ensembles of integrations (see, for example, the companion paper by Brankovic et al., 1989, hereafter referred to as III) may provide a more useful product and so greatly widen the spectrum of potential users. Furthermore, due to the availability of more powerful supercomputers, models designed primarily for weather prediction are increasingly being used for climate-timescale simulations. Diagnosis of extended-range systematic error in NWP models is an important indicator of the model's climate bias.

Ensembles of integrations of NWP models for extended-range prediction (as in III) can be computationally demanding, and for efficiency purposes, one would wish to integrate the model at the minimum resolution likely to give useful results. On the other hand, bearing in mind the objectives discussed above, it is important that systematic errors of the model used for extended-range prediction should be representative of the model used for operational shorter range forecasts. Moreover, since in many cases extended-range skill is small, it is unlikely that one would choose a resolution in which skill was seriously compromised for reasons of efficiency.

In 1985 an experimental programme was started at the European Centre for Medium Range Weather Forecasts (ECMWF) to produce a set of extended (30-day) integrations at different times of year. Some preliminary results are shown in Hollingsworth et al. (1987). In the first year of this programme, identical forecasts were run each month at four different horizontal resolutions (T21, T42, T63, T106). This paper analyses results from this first year of parallel integrations. Section 2 describes the model and forecast data; section 3 is devoted to the analysis of the systematic errors, while section 4 discusses the objective skill scores of the extended range integrations. Main results and conclusions are given in section 5.

The ECMWF NWP model is now used by a number of modelling groups in Europe, and elsewhere, as a general circulation model (GCM) for climate silmulation purposes. An assessment of the dependence of climate bias on horizontal resolution is crucial for studies where computational efficiency for multi-year integrations is of paramount importance. The results in this paper give an indication of the minimum horizontal resolution necessary for climate simulation with this model.

Some aspects of the effect of horizontal resolution in the ECMWF model have been already discussed. For example, Jarraud et al. (1988) found that, for medium-range winter forecasts, an increase from T42 to T106 resolution had an overall beneficial impact. Cubasch (1981) studied this impact on a very limited sample of extended-range integrations. In addition, Boer and Lazare (1988) have studied the dependence of horizontal resolution on climate simulations with the Canadian GCM in the range between T20 and T40.

2. THE FORECAST DATA AND THE ECMWF MODEL

The set of experiments considered in this paper consists of a monthly series of 30-day integration 'pairs', started from two consecutive initial conditions, separated by 24 hours, around the mid-month period. They are taken from the first year of the ECMWF extended-range programme from April 1985 to March 1986; Table 1 lists the initial dates of all the 24 forecasts. The choice of integrating from two consecutive days every month rather than from initial conditions separated by, say, a fortnight was motivated by the possibility of exploring the sensitivity of the integration evolution to relatively small details in the initial conditions. The difference between two such paired integrations can be viewed in terms of the evolution of the 24 hour forecast error field in the dynamical environment of a perfect model predictability experiment.

As mentioned above, all integrations were performed at four different model resolutions: T21, T42, T63 and T106 using the same version of the 16-level, operational ECMWF model, with the physical parametrization package as of March 1986. On the basis of an earlier study by Jarraud and Cubasch (1979) it was decided, at the time when experiments were made, that the (∇^4) horizontal diffusion coefficients would be unchanged for T21, T42 and T63 models, but they have been halved for T106. There is evidence that this choice of diffusion may not have been optimal for the T21 model, and that a more scale-selective diffusion (e.g. ∇^8) may lead to improved results at T21 (P. Valdes; personal communication). In Jarraud and Cubasch's study, resolutions lower than about T30 were not considered. The shallow convection and modified Kuo deep convection parametrizations (Tiedtke et al., 1988) are included as well as vertical diffusion in the free atmosphere, but gravity wave drag (Miller et al., 1989) and modified surface exchanges (Blondin and Böttger, 1987) are not. (It should be pointed out, however, that the introduction of the

gravity wave drag and removal of vertical diffusion above the PBL in later versions of ECMWF model had a major impact on the model mean; see I). Analysed sea surface temperatures (SSTs) were used as lower boundary condition and are held constant during an integration. The orography used in all the integrations is a 'one sigma' T106 envelope orography (Wallace et al., 1983). For integrations at resolution different to T106, the T106 orography was simply truncated at the appropriate model resolution. Furthermore, data assimilation was not run separately for all integrations at resolution lower than T106. Instead, the T106 resolution initial data were truncated to these lower resolutions.

Table 1

15 and 16	April	1985
15 and 16	May	1985
15 and 16	June	1985
16 and 17	July	1985
15 and 16	August	1985
15 and 16	September	1985
15 and 16	October	1985
15 and 16	November	1985
14 and 15	December	1985
18 and 19	January	1986
15 and 16	February	1986
15 and 16	March	1986

Table 1 List of initial dates for all experiments in the database. All integrations were started from 12 GMT analyses.

Integrations for the second, third and fourth year of the experimental extended-range forecasting programme were performed with revised versions of the model, both in vertical resolution and parameterization schemes, which makes it possible to address the problem of the impact of model changes in the context of interannual atmospheric variability (see I).

The forecast data is divided for all resolutions, into two 'extended' seasons, the October 1985 to March 1986 (OM) period, and the April 1986 to September 1986 (AS) period in both Northern and Southern Hemispheres, NH and SH respectively. In the present paper, the average over the 6 pairs of integrations within an extended season is referred to as an 'ensemble mean'(in contrast with the meaning in III).

3. DEPENDENCE OF SYSTEMATIC ERROR ON MODEL HORIZONTAL RESOLUTION

In sections 3.1 to 3.3 we study the influence of the model horizontal resolution on global scale systematic error within the two extended seasons. In section 3.4 we study the influence of resolution on more regional diagnostics, specifically rainfall during the Indian and African monsoon season.

3.1 Diagnosis of the zonally averaged systematic error

Fig.1 shows 30-day mean latitude-height cross-sections of the systematic error in zonally-averaged zonal wind for the ensemble means of model integrations at the four

different resolutions plus analysed values (bottom) of the zonal mean wind for comparison. For reasons of space, we show in this and other figures of zonal-mean cross-sections only values for the OM period.

Firstly, it can be seen that the dependence of systematic error on horizontal resolution is different in the stratosphere than in the troposphere. In the troposphere, the T21 model behaves substantially differently from the other resolutions. In the SH, in both seasons, the zonal mean wind of the T21 model is severely weakened throughout the depth of the troposphere in middle latitudes. This error is not evident at other resolutions. By contrast, in the NH, the T21 tropospheric errors are smaller than at other resolutions and of the opposite sign.

In the NH extratropical troposphere in both seasons, zonal mean wind errors for the T42, T63 and T106 models are quite similar to each other. There is a band of excessive westerlies extending down to the ground in middle latitudes, and a band of easterly errors at lower latitudes (indicating an erroneous poleward shift of the subtropical jet). In the southern hemisphere troposphere, the T42, T63 and T106 models again have broadly similar systematic errors, though during OM the T106 model errors appear to have the largest amplitude.

In the extratropical stratosphere, the influence of model resolution on zonal mean wind errors is somewhat different to that in the troposphere. In OM, both T21 and T42 versions of the model show comparably excessive westerlies in the extratropics of both hemispheres. In the T63 and T106 models, these OM westerly errors are reduced, and similar to each other.

In the tropics there are erroneous easterlies with maxima near the tropical tropopause during both seasons and at all resolutions. These extend down into the troposphere in the winter hemisphere. The tropical tropospheric easterly error is strongest in the NH during OM, where it is weakest at T21, and largest at T63 and T106 resolutions. Apart from equatorial region, the zonal mean wind errors are barotropic in character throughout the troposphere.

Latitude-height cross-sections of systematic error in zonally averaged temperature, are shown in Fig.2. These are manifestly consistent, through the thermal wind relations, with errors in zonal wind shown in Fig.1. The excessive stratospheric westerlies in the T21 and T42 models can be inferred from strong stratospheric high-latitude cooling and low-latitude warming shown in Fig.2. Similarly, the easterly error maximum in the SH tropospheric winds at T21 in Fig.1 is consistent with the erroneous high latitude warming in the

troposphere. In the tropical stratosphere, there is a clear warming at all resolutions. The tropical tropospheric easterly error in Fig.1, however, is related to the subtropical maximum in temperature error at about 400 mb, particularly in the NH OM period.

Fig.3(a-d) shows latitude-time diagrams of OM zonally averaged temperature error in the troposphere (average from 1000-300 mb). The high-latitude tropospheric warming in the T21 model is clearly illustrated in Fig.3a. In the SH, the evolution of the error is rapid, having almost reached saturation by day 10. (One has to exercise caution in interpreting the SH polar-latitude tropospheric temperature errors, because of possible interpolation of the post-processed data below the ground from the model to pressure levels.) In the NH, the error is smaller, and does not appear to have reached peak values until day 20. At T42 the high latitude warming has largely disappeared, except directly over Antarctica. The errors at T63 and T106 are similar to each other and to the T42 error, except that the Antarctic error has been reduced further at T63 and T106, and the NH subtropical error is a little larger at the higher resolutions (a point which is confirmed later in the paper). Note that this subtropical error may not have reached saturation by day 30. Broadly these results confirm those above, that in the troposphere, the T42, T63 and T106 models behave similarly, and quite differently to the T21 model.

In the stratosphere (Fig.3 e-h, averaged over 150-30 mb), errors are much larger. At T21 in high latitudes there is a large and rapid cooling in both hemispheres (completely the opposite of the high-latitude behaviour in the troposphere). In the tropics and subtropics there is a strong warming. The pattern is very similar at T42, though the intensity of the NH cooling is somewhat smaller. At T63, the high latitude stratospheric cooling is reduced dramatically, though, the tropical warming remains largely unchanged. The T106 model stratosphere is broadly similar to the T63 model. Again, these results confirm the discussion of Fig.2 that, in the stratosphere, the T21 and T42 models have little in common with the T63 and T106 models, as regards their behaviour in polar latitudes.

Forcing of the zonal-mean flow by the stationary and transient eddies is an essential feature of the observed general circulation. Systematic error in the zonal-mean flow necessarily implies systematic error in the zonally varying flow. Zonal cross-sections of the systematic errors in the total meridional eddy momentum and heat fluxes (stationary and transient waves) associated with such rectification are shown below.

The systematic error in tropospheric momentum flux for the four models is shown in Fig. 4. Compared with analysed values (bottom), the T21 model significantly underestimates the extratropical convergence of poleward momentum flux in both

hemispheres. For other resolutions, the pattern of systematic error is more complex with two separate maxima in the NH. In the SH, there is also a tendency to underestimate the momentum flux convergence at higher resolutions during OM (but to overestimate momentum flux convergence during AS, not shown).

In studying momentum flux errors, we have decomposed the contributions associated with zonal wavenumbers 1-3 (long waves), and zonal wavenumbers 4-9 (synoptic-scale waves). Fig.5 illustrates this error decomposition for the T106 resolution and verifying analysis during the OM period. In the NH, the analysed extratropical momentum flux convergence (bottom panels in Fig.5) is dominated by the contribution from the long waves; whilst in the SH, it is dominated by the contribution from the synoptic-scale waves. At T106 (and also T63 and T42, but not at T21), it can be seen that in the NH there is a clear overestimation of synoptic-scale flux convergence in mid-latitudes, and an underestimation of long wave flux convergence at higher latitudes, indicating some degree of compensation between errors in long-wave and synoptic-scale wave momentum flux. An interpretation of these results is given in the next section.

Errors in zonally averaged cross-section of horizontal heat flux are generally small in the troposphere, but increase into the stratosphere, where they are dominated by the long-wave contribution. Figs 6 and 7 show cross-sections of the long-wave sytematic error of horizontal heat and momentum flux (respectively) in the stratosphere. It can be seen that in the NH between T21 and T42, there is a decrease in heat flux error. Between T42 and the higher resolutions, the dependence of heat flux error on resolution is less marked than might be suggested by the resolution dependence on stratospheric zonal mean temperature error (see Fig.2). At all latitudes, and in all resolutions, the error corresponds to an underestimation of the poleward heat flux when compared with analysed values.

The fact that stratospheric heat flux errors do not decrease strongly between T42 and T63 may, at first sight, appear paradoxical, given the substantial decrease in zonal mean temperature error between these resolutions (see Fig. 2). However, as diagnostic studies in the real stratosphere have shown (O'Neill and Taylor, 1979; Palmer 1981) fluctuations in zonal mean temperature are often more strongly correlated with long-wave momentum fluxes than with long-wave heat fluxes. This can be readily understood in terms of transformed Eulerian mean dynamics (e.g. Palmer, 1981; see also below).

This also appears to be the case for the model simulations. Form Fig. 7, between 60-70 N there is a region of analysed equatorward momentum flux between 100 and 150 mb. It can be seen that these equatorward fluxes are considerably weaker at T21, with this error

reducing by about factor of 2 between T21 and T63. The dynamical significance of this is addressed in the next section.

In addition, note that near the top of the model, there is an overestimation of poleward momentum flux near 40N at all horizontal resolutions. We argue below that this is indicative of the influence of the reflecting upper boundary condition.

Despite the tendency to overestimate momentum fluxes at higher resolution, the eddy kinetic energy of the model is (except over Antarctica at T21 and T42) underestimated at all resolutions in both hemispheres at both times of year (Fig. 8). It is seen that the errors peak near the tropopause. For this particular diagnostic the dependence of error on horizontal resolution is not uniform. For example, in middle latitude NH, where errors are dominated by transient eddies, there is a decrease in error with increasing resolution. On the other hand, near the subtropical jet, errors are rather independent of resolution. It was shown in I, that these errors are substantially reduced by removing free atmosphere diffusion. The large positive errors over Antarctica at T21 are thought to be associated with an erroneous stationary anticyclone over Antarctica (see Fig 10). There is some evidence that this particular error is sensitive to the treatment of the pressure gradient terms in the vicinity of the Antarctic plateau (A. J. Simmons, personal communication).

3.2 Discussion of the zonally-averaged systematic error diagnostics

A possible interpretation of the extratropical systematic wind and temperature errors of the T21 version of the ECMWF model has already been given in Miller et al. (1989). The overall weakness of momentum flux convergence in middle latitudes in the SH is consistent with the weak zonal mean flow. However, as argued in Miller et al., the fact that systematic errors in the zonal wind at T21 are relatively small in the NH suggests that the weak middle latitude momentum flux convergence must be balanced by an erroneously weak momentum coupling to the land surface. It is hypothesised that this weak coupling is due to inadequate representation of sub-gridscale orography in the T21 model. There is evidence that the small values of momentum flux convergence in both hemispheres during both seasons is associated with dissipative effects specifically associated with running the model at that resolution (see above).

At resolutions higher than T21, the momentum fluxes increase. Consistent with this, the simulated zonal wind in the SH becomes more realistic, whilst the NH zonal wind during OM becomes quite unrealistic. It is reasonable to suppose that the erroneously weak momentum coupling to land surface persists at the higher resolutions.

One straightforward way of giving some insight into the role of orography on the general circulation is to study the difference between two sets of experiments, with and without orography (e.g. Held, 1983). In the present context, repetition of the extended range forecasts without orography would be impractical. A simple alternative would be to treat the analysed SH AS flow as a winter 'no mountain' flow, and the NH OM flow as a winter flow with mountains. Given the suggestion that the parametrization of subgridscale orography is deficient, it might be hypothesised that in winter the middle-latitude zonally averaged systematic error might resemble qualitatively the difference between the observed SH AS and NH OM zonal mean diagnostics ('no mountain' minus 'mountain' flow).

Fig.9 shows just such observed difference fields constructed from the verifying analyses shown in Figs.1,2 and 4-7 above. (For the sake of argument we ignore the 'amplitude' differences and we concentrate on discussing the pattern of these fields only). To a first approximation, the differences shown in Fig. 9 resemble to the model systematic errors at (mainly) higher resolutions. It can be seen that these differences broadly support the hypothesis as stated above. The pattern of the extratropical zonal mean zonal wind difference (Fig.9a) corresponds well with the systematic error of the T42, T63 and T106 models, with a band of westerly differences increasing in height, particularly in the stratosphere. It is interesting to note that there are no easterly differences in the tropics, suggesting that the model tropical easterly errors are not related to orography. Observed temperature difference fields are shown in Fig.9b. Consistent with the increase in westerlies with height, there is a strong cooling in the AS SH high latitudes, compared with the OM NH high latitudes. Again, this appears to be broadly qualitatively consistent with the OM zonal mean temperature errors in Fig.2 (particularly at T21 and T42).

Eddy momentum flux differences show an even more remarkable similarity to model systematic errors. (In order to form these 'difference' fields it was, of course, necessary to multiply the SH fluxes by -1 so that a SH poleward flux is transfomed into a NH poleward flux.) In Fig.9c,d we show the momentum flux differences split into the two wave bands: zonal waves 1-3, and zonal waves 4-9, respectively. It can be seen that in high-latitudes, the long-wave momentum flux difference is positive, and in lower latitudes the difference is negative (reducing the long-wave momentum flux convergence in the 'no mountain' relative to the 'mountain' flow). Exactly the opposite is the case for the synoptic-scale waves; the poleward momentum flux and its mid-latitude convergence is enhanced relative to the mountain flow.

These results suggest that there is a compensation between the momentum fluxes of the long and synoptic scale waves between a 'no mountain' and 'mountain' atmosphere, and this is exactly what was found for the momentum flux systematic errors in the higher resolution

models during the OM period (Fig.5). The reason for such compensation is probably straightforward. That is to say, NH momentum flux convergence is dominated by the long Rossby waves forced by large-scale topography and land-sea temperature contrasts. In the SH winter, with no strongly meandering jets, the zonally-averaged meridional temperature gradients are stronger in middle latitudes, and baroclinic wave activity is more pronounced.

Finally, in Fig.9e we show the observed long-wave heat-flux differences in the stratosphere. Once more, consistent with the model systematic error (Fig.6), there is a marked reduction in flux in high latitudes compared with OM NH values. It would appear therefore that the increase in the simulated long-wave heat fluxes in the lower stratosphere/upper troposphere with increasing horizontal resolution (or reduction of the model systematic error) is associated with the better resolved orography at higher resolutions. As Tibaldi (1986) has shown, enhancement of long-wave activity with envelope orography is primarily associated with the increments to the short-wave component of the orographic spectrum rather than the long-wave component. In a similar way we can expect the improved resolution of short-wave orographic components with increased resolution to enhance the vertical propagation of long-wave activity into the stratosphere.

We conclude the above discussion by suggesting that there is an inadequate representation of subgridscale orography in the 1985/86 ECMWF models.

Further interpretation of stratospheric errors is possible by combining the heat and momentum flux errors in terms of errors in the Eliassen-Palm (EP) flux, F=(-[u*v*], c[v*T*]), where [u*v*] and [v*T*] represent zonally averaged eddy momentum and heat fluxes respectively, $c=f/N^2$, where f denotes the Coriolis parameter and N is the Brunt-Väisälä frequency (see, for example, Edmon et al., 1981).

From the analysed fields in Figs. 6 and 7, and the discussion in Section 3.2, it can be seen that in the NH stratosphere during OM the EP flux vectors are directed upward and poleward between 100 - 150 mb and between 60 - 70 N. In the model simulations, the magnitude of this upward and poleward flux is reduced, with the most severe errors at T21. The direction of the EP flux can under suitable (WKBJ) conditions be interpreted in terms of the direction of propagation of Rossby wave activity (Palmer, 1982). Since convergence of the flux gives a measure of the local wave/mean-flow interaction (Edmon et al., 1981), it would appear that the stratospheric high-latitude zonal mean temperature error in Fig.2 is directly related to the direction of Rossby-wave activity into high

latitudes. At T21 and T42, the wave activity is erroneously directed away from polar latitudes, and the model mean state in high latitudes has relaxed towards radiative equilibrium and hence towards excessively cold temperatures.

A possible dynamical reason for insufficient wave/mean-flow interaction in polar latitudes is that the simulated mean flow at T21 and T42 cannot sustain sufficiently large meridional gradients in potential vorticity to focus wave activity poleward. The creation of such gradients is believed to be associated with irreversible Rossby wave breaking at the edge of the polar vortex (McIntyre and Palmer, 1984). Indeed, as Juckes and McIntyre (1987) have shown with a single layer T159 model, the structure of potential vorticity filaments stripped away from the polar vortex during wave-breaking episodes require considerable horizontal resolution to model accurately. This is also consistent with the GCM results of Mahlman and Umscheid (1987), who found that realistic simulation of middle atmosphere (stratosphere) circulations could not be achieved without increasing horizontal resolution to 1 x 1 degree (although these authors suggest additionally that an explicit gravity-wave dissipation is an important aspect of their simulations).

As noted in the previous section, the error in heat and momentum fluxes near the very top of the model were qualitatively similar at all resolution. In terms of the EP flux, these errors correspond to the suppression of vertical propagation of Rossby-wave activity and an erroneous enhancement of their equatorward propagation. These errors are not captured by the 'no mountain' minus 'mountain' analysis difference fields; that is to say they are not directly associated with errors due to representation of orography. It seems reasonable to speculate therefore that this error is directly associated with the reflective upper boundary condition which channels Rossby-wave activity further equatorward near the top of the model, leading to erroneous wave-phase characteristics.

It is likely that errors in eddy kinetic energy are not associated directly with inadequate orographic forcing, but with excessive internal damping processes acting across the tropopause in particular. As shown in I, substantial reduction in this error occurred when vertical diffusion in the model free atmosphere was removed.

3.3 The geographical distribution of large-scale systematic error

Fig.10 shows ensemble average 30-day mean 500 mb height error maps for forecasts from the OM period in both NH and SH for all four model resolutions (T21, T42, T63 and T106), together with the verifying analysis (bottom). As suggested by the zonal mean cross sections, 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. The T42, and particularly T63 and T106 models behave very similarly to each other.

Comparison of the higher resolution error in the NH fields with similar maps characteristic of earlier versions of the ECMWF operational model (e.g. Wallace et al., 1983) demonstrate that typical Day-10 error magnitudes of the 1981 operational model are not attained until around days 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 quasi-stationary planetary waves and strengthening of the westerly jet in mid-latitudes.

In the SH, the T21 climate suffers from particularly large deficiencies, as was mentioned above. The pole-to-equator geopotential height gradient is severely underestimated, together with the strength of the zonal flow. At higher resolution, the principal midlatitude error is a slight underestimation of the zonal mean flow, and the wavenumber 3 stationary-wave pattern.

We now focus our attention on the tropical region. Figs.11 and 12 show the 30-day mean 200 mb velocity potential and streamfunction errors, respectively (for the AS period only). It can be seen that the 200 mb divergent flow is underestimated at all resolutions. In particular, the maximum of upper divergence over the west Pacific is severely weakened, the magnitude of the error being smallest at T21 but larger and monotonically increasing at T42, T63 and T106. The error in upper level convergence at T21 is largely concentrated over east Africa, east of the region of maximum analysed convergence. At higher resolutions particularly at T63 and T106, this error is spread more uniformly across Africa and the Atlantic.

Maps of 200mb streamfunction errors are shown in Fig.12: these are consistent with the upper tropospheric tropical zonal mean easterly wind errors discussed in section 3.1. At all resolutions, significant tropical easterly errors extend across the Atlantic and Pacific; only across the Indian ocean are they small or westerly. The magnitude of the errors increases with resolution, being smallest at T21, and largest at T63 and T106.

The streamfunction and velocity potential error maps are qualitatively consistent assuming a linear Sverdrup-balance model of the tropical atmosphere (Gill, 1980). In particular, if a pattern of forcing is specified with a horizontal structure given by the error maps in Fig 11, the Gill model predicts anomalous easterlies to the east of the upper anomalous (in our case erroneous) convergence (west Pacific) and west of the anomalous upper divergence (Africa/Atlantic), and anomalous westerlies elsewhere. In Gill's model, the strength of the anomalous winds is linearly related to the strength of the anomalous forcing. Similarly, we note that the error in nondivergent wind increases (with resolution) as the error in the divergence increases over Indonesia.

It is interesting to speculate why the tropical systematic error should get monotonically worse with increasing resolution. One possible explanation is that the physical parametrizations, notably the (Kuo) convection scheme, only give realistic estimates of diabatic heating when driven by large-scale moisture flux convergences. Another possibility, related to this is that in the high resolution models, the relatively small-scale (and high frequency) moisture flux convergences are quite "noisy" and unrealistic (B.J. Hoskins, personal communication). This raises the question of whether diabatic heating fields in high resolution models should be determined by very small scale features of the flow.

Errors in the global 850 mb wind fields (not shown) are broadly the opposite of those at 200 mb, with erroneous easterlies over Indonesia, and erroneous westerlies elsewhere, corresponding to a weakening of the trade winds. These erroneous easterlies are notably stronger than the erroneous westerlies, and the zonal mean wind errors at 850 mb are consequently relatively small.

As an example of 850mb flow, relevant to the discussion in section 3.4, Fig.13 shows the 30-day mean 850 mb monsoon flow at the four horizontal resolutions from 17 July 1985. Compared with the analysed values (Fig.13, bottom), the strength of the easterlies over the south Indian ocean and the low level jet off the Somali peninsula are all underestimated. Notice that, as with the global 200mb streamfunction and velocity potential fields, systematic error increases with increasing resolution.

3.4 Influence of resolution on simulation of monsoon rainfall

In paper I we studied the extended-range predictability of tropical precipitation during the summer monsoon. In this section we show, as an example of the impact of resolution on regional simulation, monthly mean precipitation over the Indian and African regions for forecasts initialised on 17 July 1985. These are shown in Figs.14 and 15 together with the corresponding orography over the regions. Also shown is the monthly-mean value of 24hr accumulated rain from the operational ECMWF Day-1 forecast integrations during the 30-day period of the extended-range integration; we use this latter field with caution to 'verify' the 30-day forecast fields.

The T21 model is unable to simulate the rain shadow effect of the Western Ghats, which is a well-observed feature of the Indian monsoon, and is clearly shown in the 'verification' field. At T21 there is a single rainfall maximum positioned over the Indian sub-contintent with a minimum in rainfall over the Bay of Bengal. A major difference occurs at T42 with rainfall maxima positioned over the Bay of Bengal and to the north-east of the Western Ghats. Notice the strong gradients in rain amounts to the north of this maximum near 25N,

approximately correctly simulated. If the precipitation field from the T42 simulation is truncated itself at T21 (not shown), the principal features remain, thus confirming the benefit of using the higher resolutions for rainfall prediction. The simulations at T63 and T106 clearly show more small-scale features than the T42 resolution: for example, the values of maxima are increased and the rain-shadow effect near the southern tip of India is strongly enhanced at T106. However, the rainfall features on the scale of the subcontinent are not significantly changed at higher resolution. The erroneous northward displacement of the rainfall maxima over the west coast of India is consistent with the error in low level flow shown in Fig.13.

The sensitivity to resolution of the positioning of rainfall maxima, seen in Fig.14, is not apparent over Africa (Fig.15). Apart from some sharpening of gradients at higher resolution, the major difference between T21 and T42 prediction is in the simulation of the coastal west African maximum. Note that the increase of this maximum with resolution occurs as an orographic feature "moves" into the ITCZ region! The strength of this rainfall maximum, and the maximum over the Ethiopian highlands, increases fairly monotonically with resolution through to T106.

These results broadly confirm those found from studying diagnostics of global scale flows; the simulation of regional rainfall can change markedly between T21 and T42 resolution; however, between T42 and higher resolution the major differences are ones of detail, relating to strength of maxima, and sharpness of gradients. The ability to resolve orographic features improves locally the realism of the model rainfall simulations.

4. DEPENDENCE OF OBJECTIVE SKILL SCORES ON MODEL HORIZONTAL RESOLUTION

It might be supposed that in the medium and extended range, the skill of an NWP model is directly correlated with its systematic error. As was noted above, the magnitude of the tropospheric NH OM height and wind errors were smallest at the lowest resolution T21 model. This raises the question as to whether the T21 model is the most skilful of the four resolutions during the NH OM period. In the following, skill scores (anomaly correlation coefficient, ACC, and RMS error) are defined over the area enclosed between 22.50 and 86.25 degrees, and are computed for time-mean 500 mb height and 850 mb temperature in the NH and SH.

Fig.16 shows the time series of the 500 mb ACC height error throughout the year of the experiments, for both the mean of days 1-10 and days 11-20. For each month the curves are drawn through the mean of the two 30-day forecasts initiated 24 hours apart.

In the NH, it can be seen that for days 1-10, T21 has consistently the poorest forecast skill. The T42 is the next poorest, though on occasion (e.g. October 1985) it can 'outperform' the higher resolution models. The relative skill of T63 and T106 is hard to distinguish visually.

For days 11-20, the T21 model shows somewhat different behaviour to the other models, and the ACC is lower than in the other models in almost half of the cases. There is one case (January 1986 NH) when both of the T21 forecasts are much superior to all the others. This case is discussed at length in III, though in a different context. During the second pentad of the forecast period, an intense block developed over the Euro/Atlantic region, and geopotential height anomalies were generally positive over high latitudes in the NH. During the first 10 days, the T21 forecast was the poorest of the set (Fig.16). However, subsequent to this, all of the forecasts exhibited their characteristic climate drift, with the T42, T63 and T106 forecasting anomalously strong hemispheric westerlies, and the T21 model forecasting weak anomalous easterlies in mid-latitudes. Because of its relative failure in the early period, the relative success of the T21 model in days 11-20 must be viewed as a chance occurence.

On the other hand during February 1986, the T21 model was clearly the poorest at days 11-20. However, in this case the success of the higher resolution models cannot be ranked as a chance occurrence, again because they were very skilful in the first ten days. These higher resolution integrations predicted the correct evolution of anomaly fields during a period characterised by significant blocking anomalies (c.f. paper III).

Similar conclusions can be made for the SH skill scores. From Fig.16, it can be seen that the T21 model is poorest for days 1-10 of the forecast period, with some indication that T42 is marginally poorer than the higher resolutions. For days 11-20, the three high resolution models do not show obvious systematic differences. As in the NH, there is an interesting period during February 1986 when all models show comparable high skill. The 500 mb height anomaly maps (not shown) for this forecast period indicate anomalous positive geopotential height in SH high latitudes. Unlike the January 1986 NH case, these anomalies were already well established during the first ten days. Again the T21 model developed its characteristic systematic error which in this instance was 'skilful', whilst the higher resolution models were able to maintain the anomalous flow during the 30-day period. Indeed, these two cases (January 1986 NH and February 1986 SH) typify and confirm results of earlier medium range forecast studies (Grönaas, 1982; Palmer and Tibaldi, 1988; Tibaldi and Molteni, 1989) that the operational model is poorest when blocking develops more than a few days into the forecast period, but is most skilful when blocking is already established.

Fig.17 shows time series of 30-day mean 850 mb temperature for the northern and southern hemispheres. It is interesting to note the apparently high degree of skill in forecasting this variable, particularly in the SH. This level of skill is maintained throughout the forecast range (not shown). This would appear to indicate the direct influence of sea surface temperature (whose autocorrelation time exceeds one month) on the model's boundary layer thermal structure (Molteni et al., 1988, and III). In contrast to the 500 mb height field ACC, the systematic error of the T21 model in the SH has a fairly uniformly deleterious effect on 850 mb temperature skill scores.

In Fig 18, we show 10- and 30-day mean RMS and ACC scores for OM for the 4 resolutions. It is interesting to note that the T21 model is poorer than the others up to day 11-20 at least in terms of ACC. This suggests that apparent forecast skill at this range with the higher resolution models is not purely the result of chance sampling. We believe that this gives some support to the conclusions of I and III, that with operational NWP models, there is evidence of forecast skill in the range between days 11-20.

In order to assess more quantitatively the extent to which differences in skill between model resolutions are statistically significant, a Student t-statistic has been calculated to determine whether the skill/spread at a given resolution is significantly higher than the skill at the next higher resolution. There is, however, considerable ambiguity in the calculation and interpretation of t, not least associated with uncertainty in the number of degrees of freedom chosen. However, for comparison purposes, we quote the standard statistical test that with 22 degrees of freedom (12 pairs of forecasts), a one sided t-test would be accepted with 90% confidence if $t>1.32$.

On this basis, the t value associated with the difference in skill between the (z transform of) ACC between the T21 resolution and the T106 resolution at days 11-20 during OM is 1.67. By comparison, the t value associated with the the difference in skill between the T63 resolution and T106 resolution is 0.44. Hence, whilst Fig 18 indicates that T63 is more skilful than T106 in the extended range, it is likely that this is due to sampling. (Note also that for daily scores within the first ten days, Simmons et al., 1989, have shown with a much larger sample that the difference in skill between the T63 and T106 model is measurable, and in favour of the T106 model.) On the other hand, whilst sampling cannot be ruled out, it is less likely that T21 and T106 have comparable ACC skill in the extended range between days 11-20. (In terms of RMS, there is no significant difference between T21 and T106 during OM at days 11-20; however, there is during AS, where the t value for days 11-20 is 1.49.)

In Fig 19 we show the 10- and 30- day mean RMS spread of hemispheric 500mb height between the forecast twins for OM. This gives information about internal error growth at each model resolution. It can be seen that, on average over the four models, the asymptotic 10-day mean RMS spread between adjacent forecasts is about 80 m for the OM season. For a model with realistic low-frequency variability, the asymptotic spread should approach the asymptotic persistence error, i.e. the difference between two randomly chosen fields. However, it can be seen that the asymptotic model spread is about 70% of persistence during OM, (and about 50% of persistence during AS). This indicates a significant underestimation of atmospheric low-frequency variability at all resolutions (consistent with the loss of eddy kinetic energy disussed above and also in I).

It is worth noting that there does appear to be some evidence that the asymptotic value of spread between the higher and lower resolution models is significantly different. For example, the t value associated with the difference between the day 21-30 spread of the T42 and T63 models is 1.52.

One important caveat should be made concerning these results. It is possible that the potential of the higher resolution models is not being fully realised in the extended range, due to the increase of tropical systematic error with horzontal resolution. In a companion paper, Ferranti et al. (1989) showed that extended-range extratropical skill can be improved considerably with an accurate simulation of the tropical flow. In view of our earlier comments, and the results of Ferranti et al., it is possible that extratropical skill of the T63 and T106 models in the range days 11-20 could be enhanced if some of the physical parametrization schemes were driven by spatially truncated fields.

5. CONCLUSIONS

In recent years, the difference between numerical weather prediction (NWP) models and general circulation models has become less distinct, just as the boundaries between extended range prediction and climate forecasting are themselves ill-defined. As such, the present study of the influence of horizontal resolution on the extended-range systematic error and skill of the ECMWF spectral NWP model is of relevance to both weather prediction and climate simulation studies. We have studied the behaviour of a set of 30-day forecasts during one annual cycle, at four resolutions: T21, T42, T63, T106. The low resolution model, whilst not a serious candidate for NWP, is used in many centres for multi-year climate integrations.

Zonal mean error diagnostics from these integrations clearly indicate that in the extratropical troposphere, the climate drift of the T21 model is qualitatively different to the higher resolution models. The impact of horizontal resolution between T21 and T42 is

broadly in agrement with Boer and Lazare (1988). The T63 and T106 models have very similar mean errors, and the T42 model errors closely resemble these. In the extratropical stratosphere, however, the T42 model systematic errors are comparable to that of the T21 model, whilst again the T63 and T106 model climates appear very similar to one another. These results indicate that the T42 model is in some sense transitional between the T21 low resolution and T63, T106 high resolution models. However, for many tropospheric purposes the T42 model could be viewed as having systematic errors quite representative of higher resolution models.

The behaviour of the T42 model as a 'low' resolution model in the stratosphere is rather interesting, and suggests that the dynamical warming of the stratosphere through irreversible Rossby wave breaking cannot be adequately simulated at T42 resolution, but can at higher resolutions. This appears to be consistent with the emerging view of the stratosphere as a region of the atmosphere with abundant fine-scale structure. Results from the T42 model in this study also raise important questions concerning the coupling of the stratosphere with the troposphere. Specifically, the fact that the T42 model can mimic the T21 model in the stratosphere, yet resemble the T106 model in the troposphere, suggests that systematic errors in the two regions of the atmosphere can vary quite independently.

Differences between zonal-mean diagnostics of the observed winter flow in the NH and the SH are well correlated with the winter zonal mean systematic errors of the NH higher resolution models - including excessive westerlies throughout the depth of the atmosphere, an overestimation of transient meridional eddy momentum flux, and an underestimation of meridional long wave heat flux. It is suggested that this indicates that topographic forcing in the model was underestimated, and it is noted that the version of the ECMWF model used in this study had no parametrization of orographic gravity wave drag (although it did have an envelope orography).

On the other hand, other systematic errors have mainly weak dependence on horizontal resolution, and have no clear relation with orography. These included an underestimation of subtropical eddy kinetic energy and erroneous momentum and heat fluxes near the very top of the model. It is shown in I that the error in eddy kinetic energy is associated with excessive dissipation near the tropopause level. In this study, it was argued that errors in the momentum and heat fluxes near the top of the model were most likely associated with the model's reflective boundary condition.

Some aspects of the large-scale tropical flow appear to deteriorate with increasing horizontal resolution. For example, upper tropospheric divergence over Indonesia is

underestimated at all resolutions, though is particularly poor at T106. The low-level monsoon flow is most accurately simulated at T21 resolution, and is generally weak at T106. It was suggested that, for example, the Kuo parametrization scheme may only give an accurate estimation of large-scale diabatic heating when driven by large scale flow fields.

The regional simulation of rainfall during the Indian and African monsoon was also studied. Consistent with results above, significant changes were apparent between the T21 and T42 resolution models. However, apart from matters concerning details in the rainfall field, like strength of maxima and sharpness of gradients, the T42 model behaves qualitatively similarly to the others. The ability to resolve local orographic features is important in accurately simulating local precipitation maxima or minima.

On the other hand, bearing in mind the deterioration of some aspects of the large-scale tropical flow at high horizontal resolution, and results from Ferranti et al (1989) showing significant impact of tropical forecast errors on extended range extratropical forecast skill, it is possible that the potential of the higher resolution models for extended range forecasting may be underestimated in this study.

The evidence presented in this paper suggests that, bearing in mind the computational cost, the T106 resolution may be unecessarily fine for both extended-range forecasting, and for climate simulation studies. However, at present, the computational burden of T63 resolution may also be excessive for many climate studies as well. It would appear from our analysis, that with some caveats (particularly in the stratosphere) the behaviour of the T42 model is comparable with the higher resolution models, and is a satisfactory compromise for many purposes, bearing in mind the impact of resolution on tropical systematic error. However, it would appear that the extreme sensitivity of the ECMWF model climate drift to resolution in the range between T21 and T42, and the apparent inability of the ECMWF T21 model to simulate the correct internal nonlinear dynamics of the extratropics, makes integration at resolution lower than T42 questionable.

REFERENCES

Blondin, C. and H.Böttger, 1987: The surface and sub-surface parametrization scheme in the ECMWF forecasting system: Revision and operational assesment of weather elements. ECMWF Tech.Memo. No.135, 20 pp. ECMWF, Reading, UK.

Boer, G.J. and M. Lazare, 1988: Some results concerning the effect of horizontal resolution and gravity-wave drag on simulated climate. J.Clim., 1, 789-806.

Brankovic C., T.N.Palmer, F.Molteni, S.Tibaldi, and U.Cubasch, 1989 : Extended range predictions with ECMWF models: III Time-lagged ensemble forecasting. (submitted to Quart. J. R. Meteor. Soc.)

Cubasch, U., 1981: The performance of the ECMWF model in 50 day integrations, ECMWF Tech. Memo. NO. 32, 74 pp. ECMWF, Reading, UK.

Edmon, H.J., B.J.Hoskins and M.E.McIntyre, 1980: Eliassen-Palm cross sections for the troposphere. J. Atmos. Sci.,37, 2600-2616.

Ferranti, L., T.N. Palmer, F. Molteni and E. Klinker, 1989: Tropical, extratropical interaction associated with the 30-60 day oscillation, and its impact on medium and extended range predictability. Submitted to J.Atmos.Sci.

Gill, A.E., 1980: Some simple solutions for heat-induced tropical circulation. Quart. J. R. Meteor. Soc., 106, 447-462.

Grönaas, S., 1985: A pilot study on the prediction of medium range forecast quality. ECMWF Tech.Memo. No.119, 23 pp. ECMWF, Reading, UK.

Held, I.M., 1983: Stationary and quasi-stationary eddies in the extratropical troposphere: theory. In: Large-scale dynamical processes in the atmosphere, edited by B.J.Hoskins and R.P.Pearce. Academic Press, New York, 127-168.

Hollingsworth, A., U. Cubasch, S. Tibaldi, C. Brankovic, T.N. Palmer and L. Campbell: Mid-latitude atmospheric prediction on time scales of 10-30 days. In: Atmospheric and oceanic variability, ed. H. Cattle, Royal Meteorological Society Monograph, Bracknell, U.K., 117-151.

Jararud, M. and U. Cubasch, 1979: Horizontal diffusion experiments with the ECMWF spectral model. ECMWF Tech. Memo. No. 8, 17 pp. ECMWF, Reading, U.K.

Jarraud, M., A.J. Simmons and M. Kanamitsu, 1988: Sensitivity of medium-range weather forecasts to the use of an envelope orography. Q.J.R.Meteorol.Soc., 114, 989-1025.

Juckes, M.N. and M.E.McIntyre, 1987: A high resolution one-layer model of breaking planetary waves in the stratosphere. Nature, 328, 590-596.

Mahlman, J.D. and L.J. Umscheid, 1987: Comprehensive modeling of the middle atmosphere: the influence of horizontal resolution. In: Transport processes in the middle atmosphere, edited by G.Visconti and R.R.Garcia, Riedel Publishing Company, Dordrecht, 251-266.

McIntyre, M.E. and T.N.Palmer, 1984: The 'surf zone' in the stratosphere. J. Atmos. Terr. Phys., 46, 825-849.

Miller, M., T.N.Palmer and R.Swinbank, 1989: Parametrization and influence of subgridscale orography in general circulation and numerical weather prediction models. Meteor. and Atmos. Phys., 40, 84-109.

O'Neill, A. and B.F.Taylor, 1979: A study of the major stratospheric warming of 1976-77. Quart. J. R. Meteor. Soc, 105, 71-92.

Palmer, T.N., 1981: Diagnostic study of a wavenumber-2 stratospheric sudden warming in a transformed Eulerian mean viewpoint. J. Atmos. Sci., 38, 844-855.

Palmer, T.N., 1982: Properties of the Eliassen-Palm flux for planetary-scale motions. J. Atmos. Sci., 39, 992-997.

Palmer, T.N., 1987: Modelling atmospheric low-frequency variability. In: Atmospheric and Oceanic Variability, Roy. Met. Soc. Monograph, 75-103.

Palmer, T.N. and S.Tibaldi, 1988: On the prediction of forecast skill. Mon. Wea. Rev., 116, 2453-2480.

Palmer, T.N., C.Brankovic, F.Molteni and S.Tibaldi, 1989: Extended range predictions with ECMWF models: I Interannual variability in operational model integrations. (submitted to Quart. J. R. Meteor. Soc)

Simmons, A.J., D.M. Burridge, M.Jarraud, C.Girard and W. Wergen, 1989: The ECMWF medium range prediction models. Development of the numerical formulations and the impact of increased resolution. Meteorol. Atmos. Phys., 40, 28-60.

Tibaldi, S., 1986: Envelope orography and maintenance of the quasi-stationary circulation in the ECMWF global models. Adv. in Geophys., 29, 339-372.

Tibaldi, S. and F.Molteni, 1989: On the operational predictability of blocking. Tellus xxA (in press).

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

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

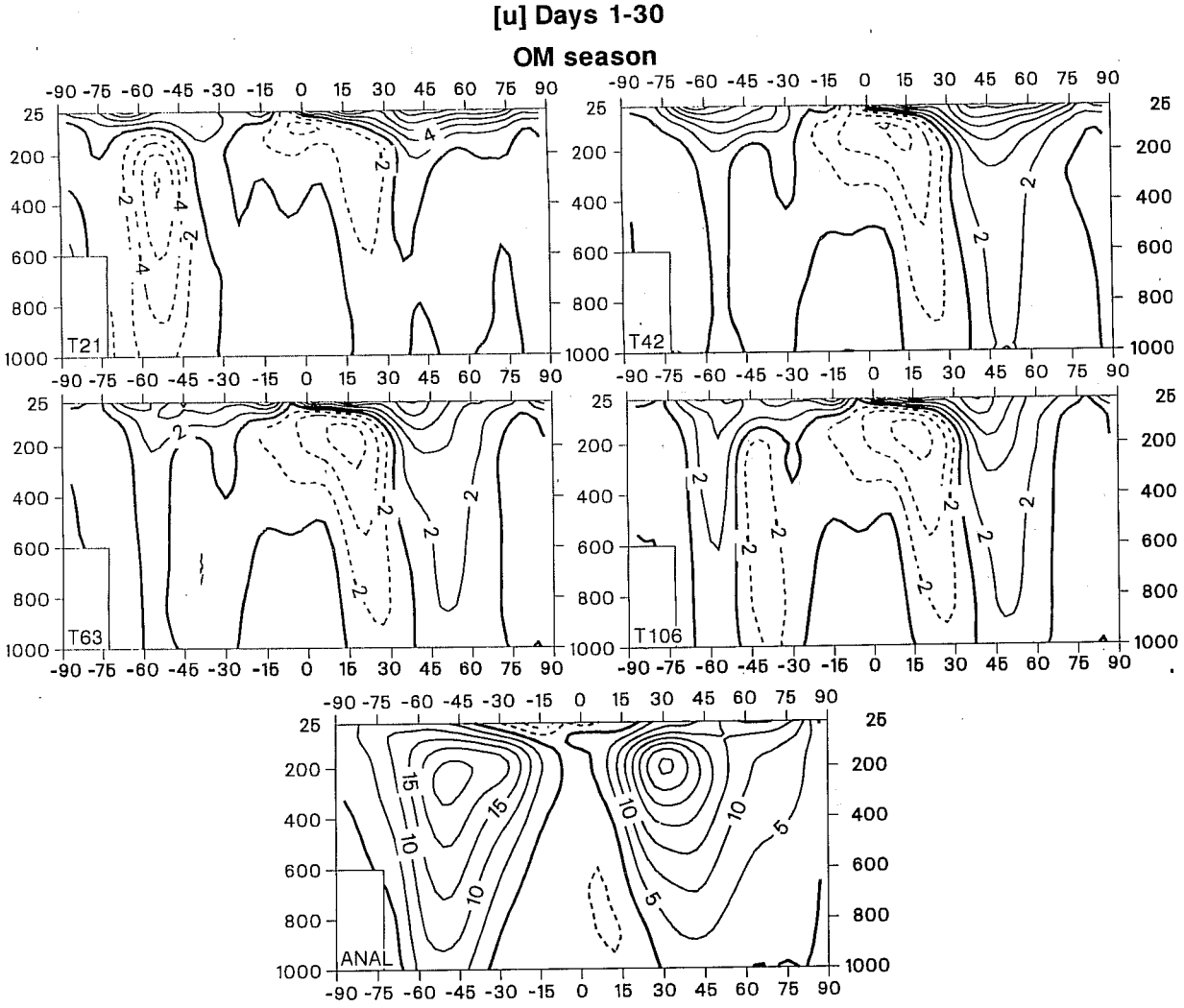

Fig.1 Zonally averaged latitude-height cross sections of 30-day mean zonal wind error at all four resolutions together with the verifying analysis: October 1985 to March 1986. Contour interval 2 ms^{-1} for errors and 5 ms^{-1} for the analysis. Negative errors and easterlies dashed.

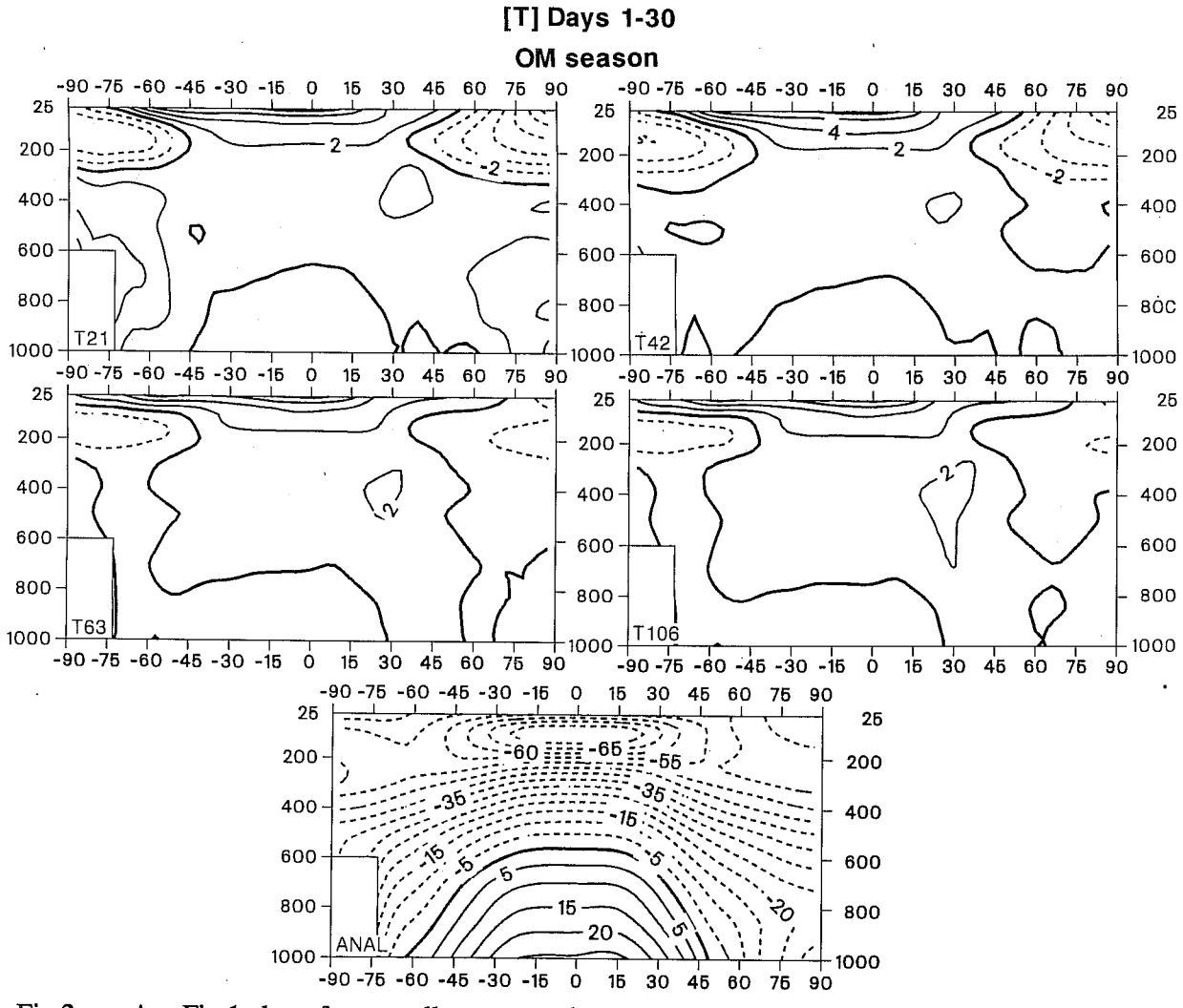

Fig.2 As Fig.1 but for zonally averaged temperature. Contour interval 2 K for errors and 5 K for the analysis. Negative errors and negative temperatures dashed.

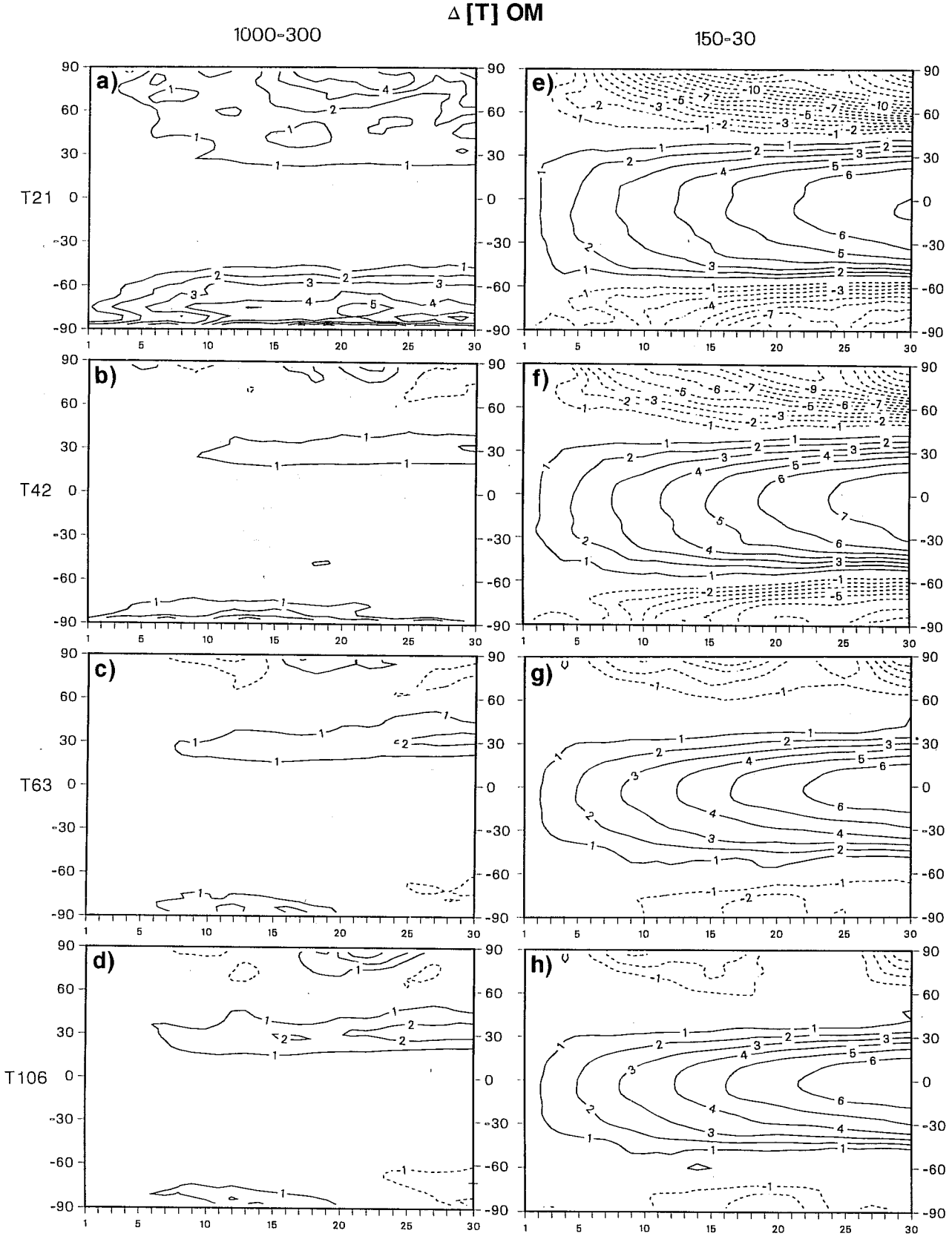

Fig.3 Time evolution of zonally averaged temperature errors at all four resolutions in the season October 1985 to March 1986 averaged for the slab 1000-300 mb (left) and 150-30 mb (right). Contour interval 1 ms^{-1}, negative errors dashed.

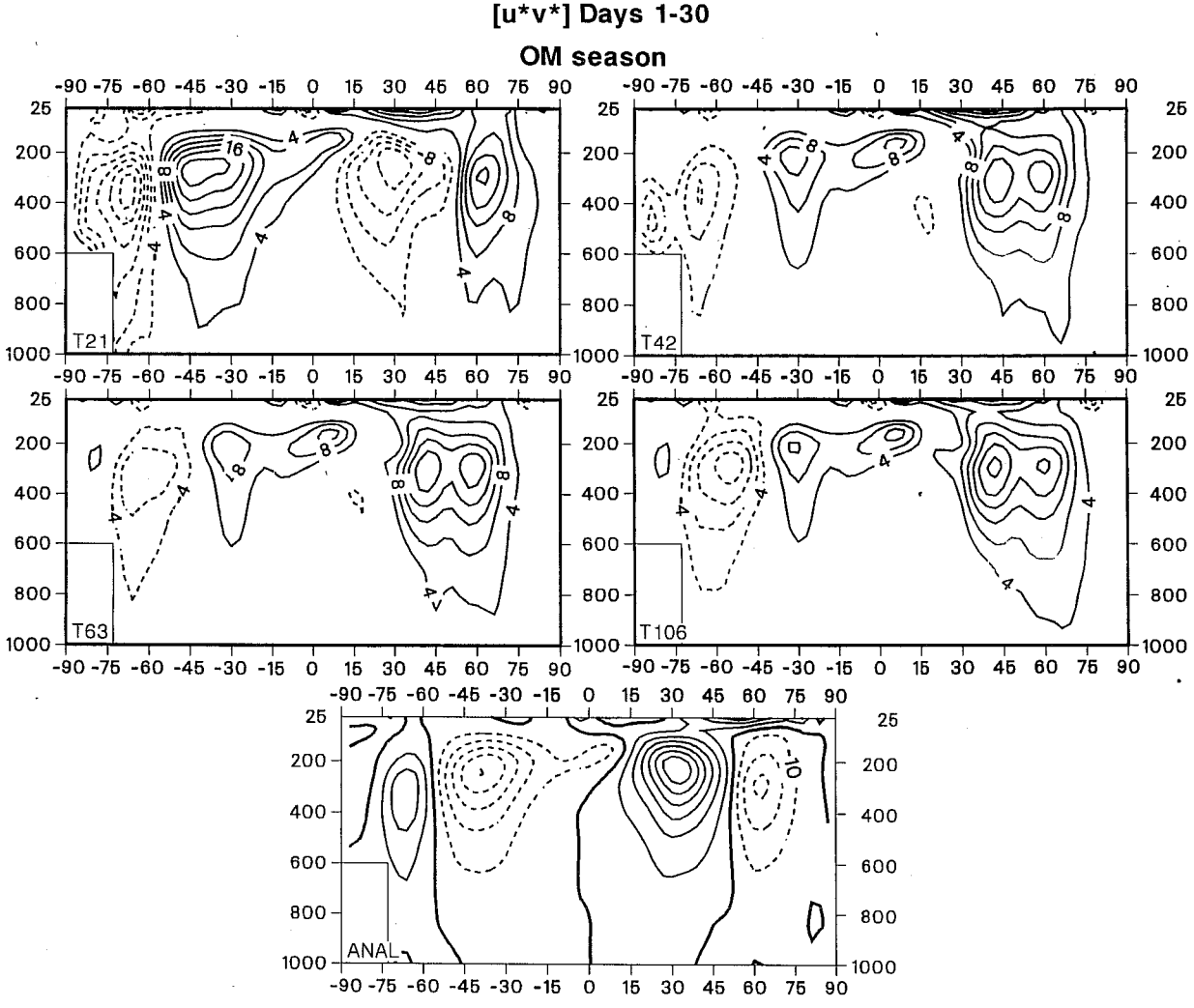

Fig.4 As Fig.1 but for zonally averaged meridional eddy momentum flux (all zonal wavenumbers). Contour interval 4 m^2s^{-2} for errors and 10 m^2s^{-2} for the analysis. Negative errors and northerly fluxes dashed.

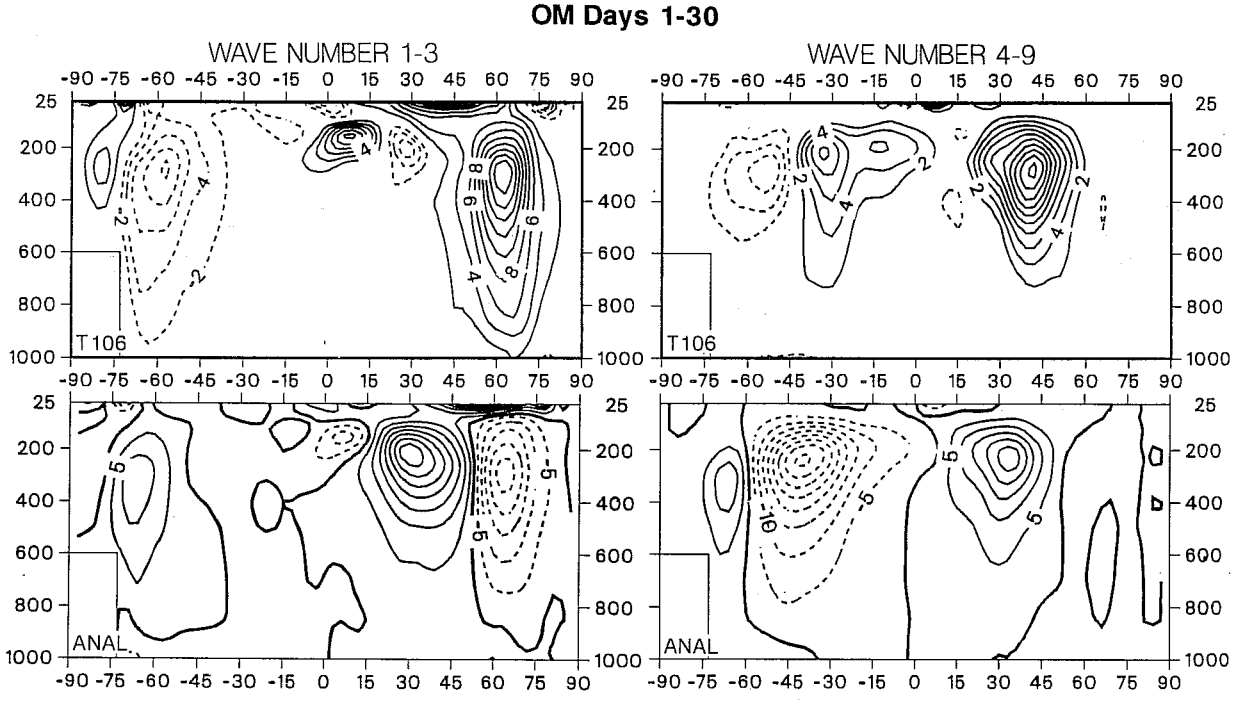

Fig.5 Zonally averaged latitude-height cross sections of 30-day mean errors for T106 model meridional eddy momentum flux (top) and verifying analysis field (bottom) in the season October 1985 to March 1986. Left: wavenumber 1-3, right: wavenumber 4-9. Contour interval 2 m^2s^{-2} for errors and 5 m^2s^{-2} for the analysis. Negative errors and northerly fluxes dashed.

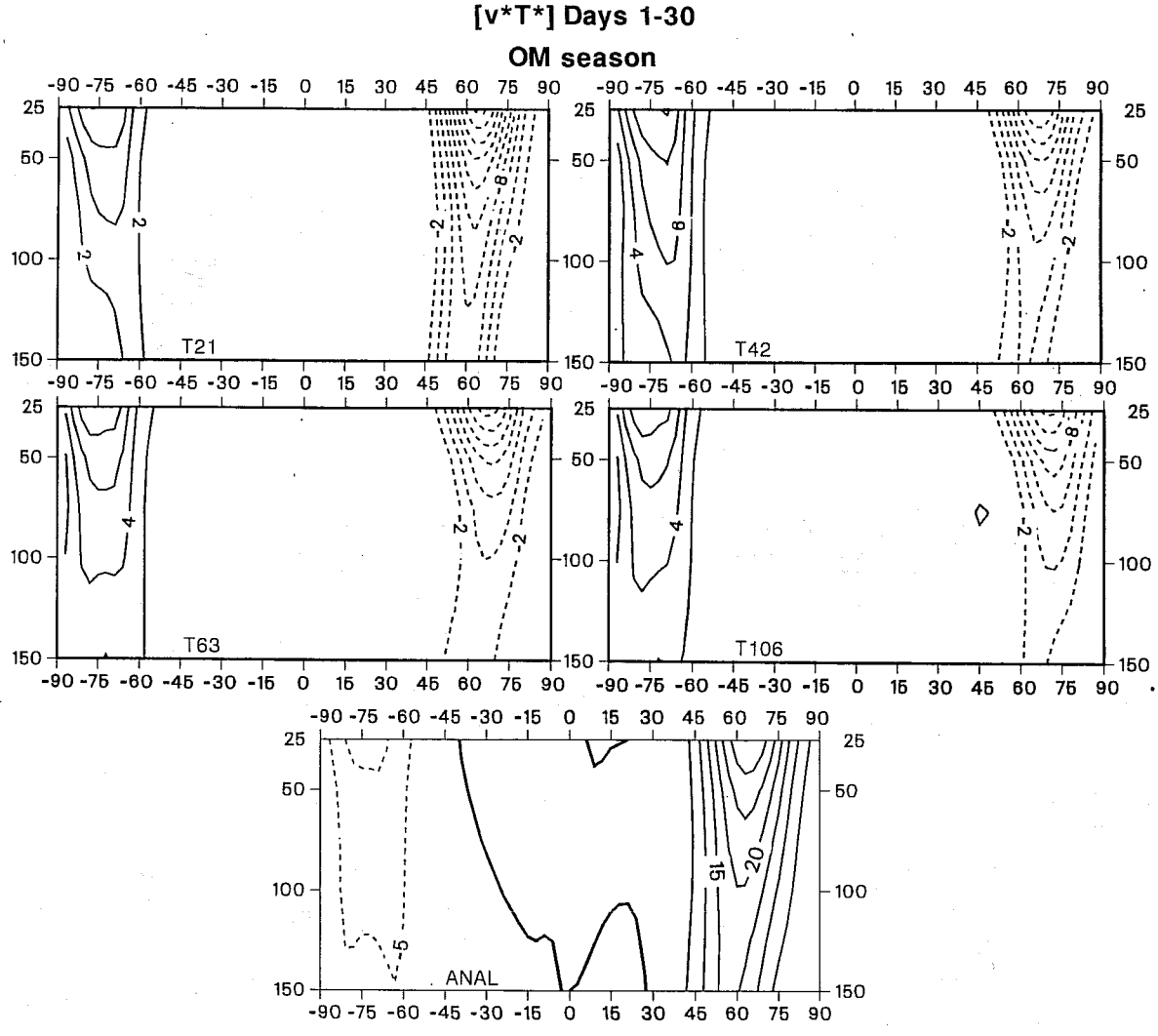

Fig.6 As Fig.1 but for wavenumber 1-3 zonally averaged eddy heat flux in the stratosphere (slab between 150 and 25 mb) and the season October 1985 to March 1986. Contour interval 2 Kms^{-1} and 5 Kms^{-1} for errors and analysis respectively.

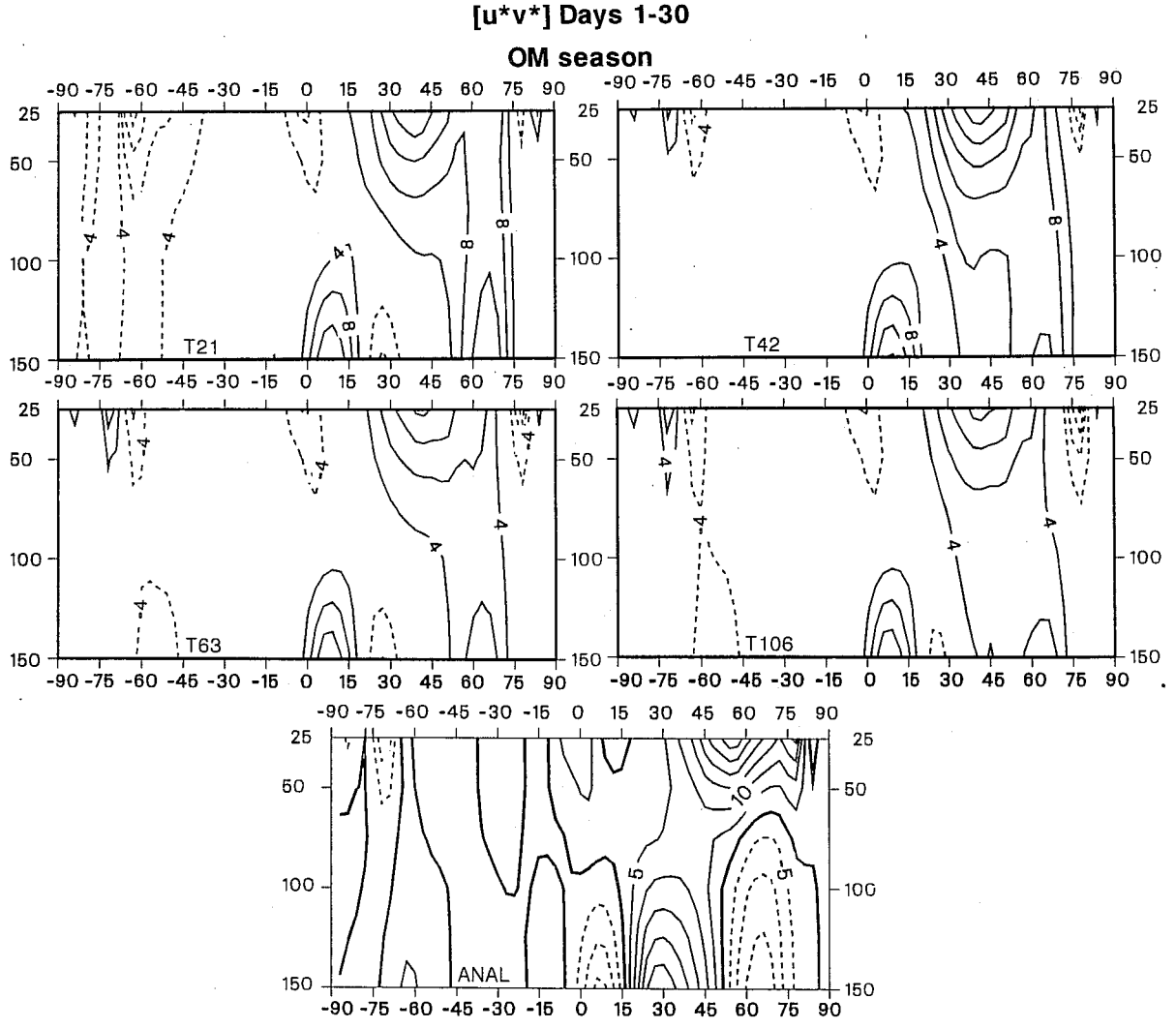

Fig 7. As Fig 6 but for momentum flux. Contour interval 4 m^2s^{-2} and 5 m^2s^{-2} for errors and analysis respectively.

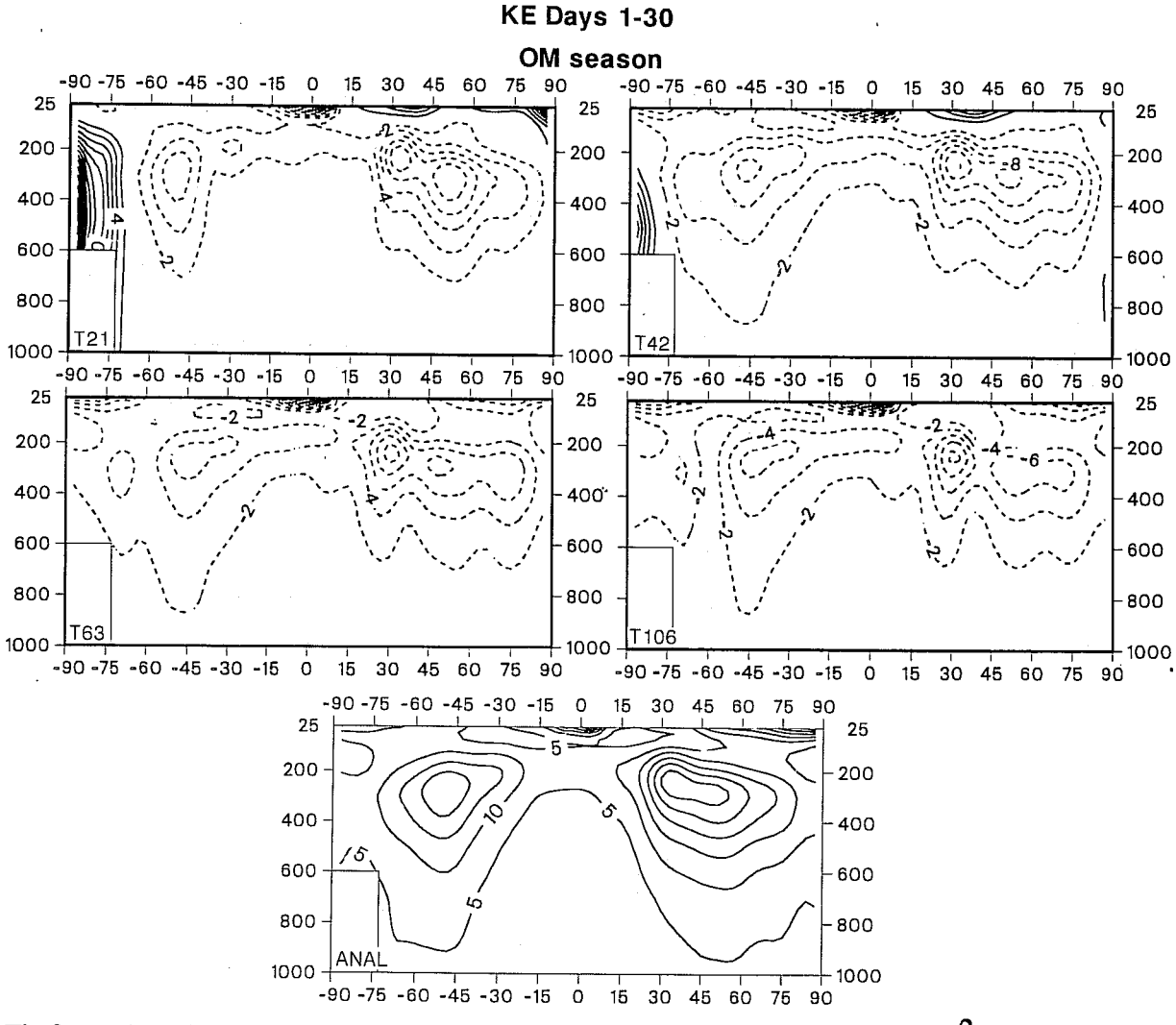

Fig.8 As Fig 1 but for eddy kinetic energy. Contour interval 2 kJm^{-2}; negative error dashed.

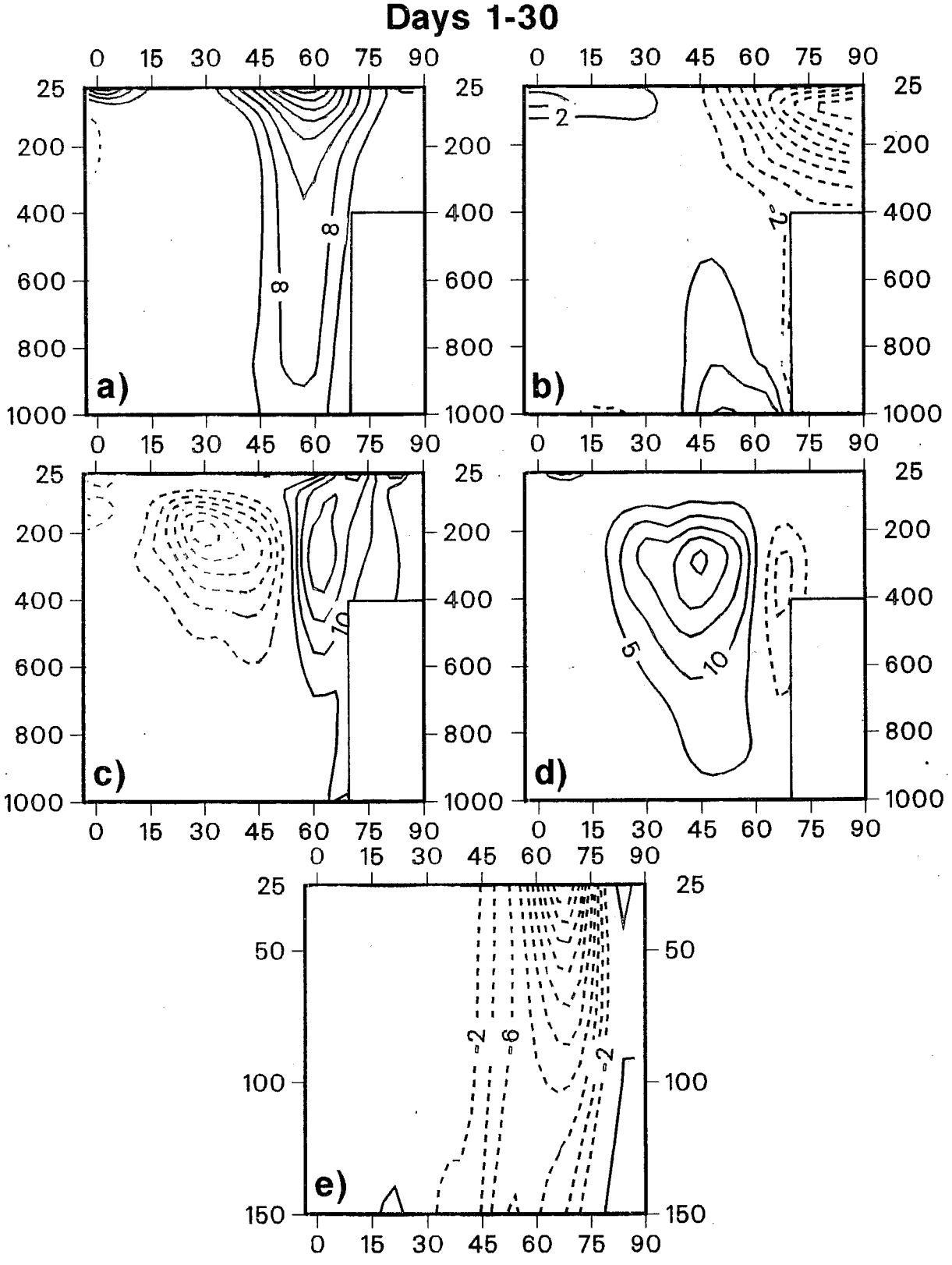

Fig.9 Difference between 30-day mean analysed zonally averaged southern hemisphere April-September 1985 diagnostics and analysed zonally averaged northern hemisphere October 1985-March 1986 diagnostics (wintertime 'no mountain' minus 'mountain' flow). a) zonal wind; b) temperature; c) contribution to meridional eddy momentum flux from zonal wavenumbers 1-3; d) contribution to meridional eddy momentum flux from zonal wavenumbers 4-9; e) contribution to stratospheric meridional eddy heat flux from zonal wavenumbers 1-3 .

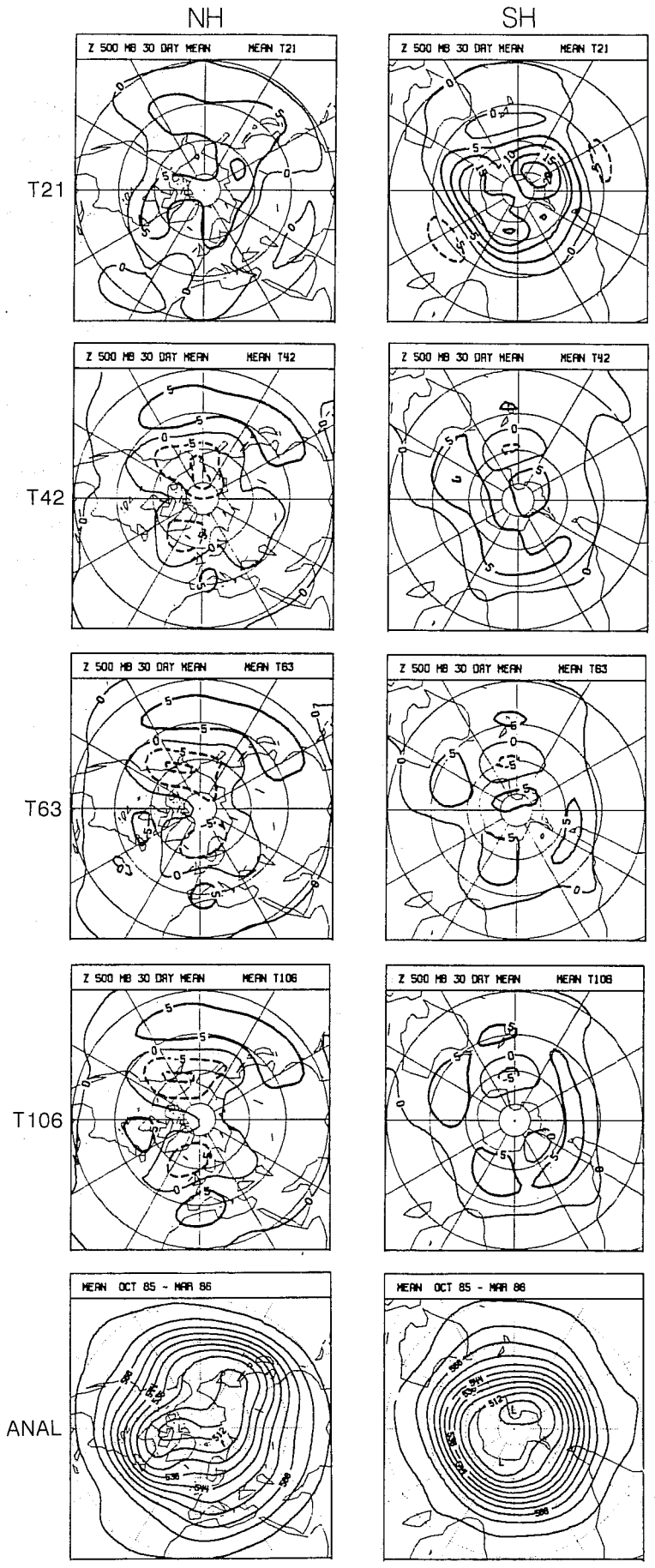

Fig.10 30-day mean 500 mb geopotential height errors for the four resolutions and verifying analysis (bottom) in the season Octobar 1985-March 1986. Left: northern hemisphere; right: southern hemisphere. Contour interval 5 dam for errors and 10 dam for analysis; negative errors dashed.

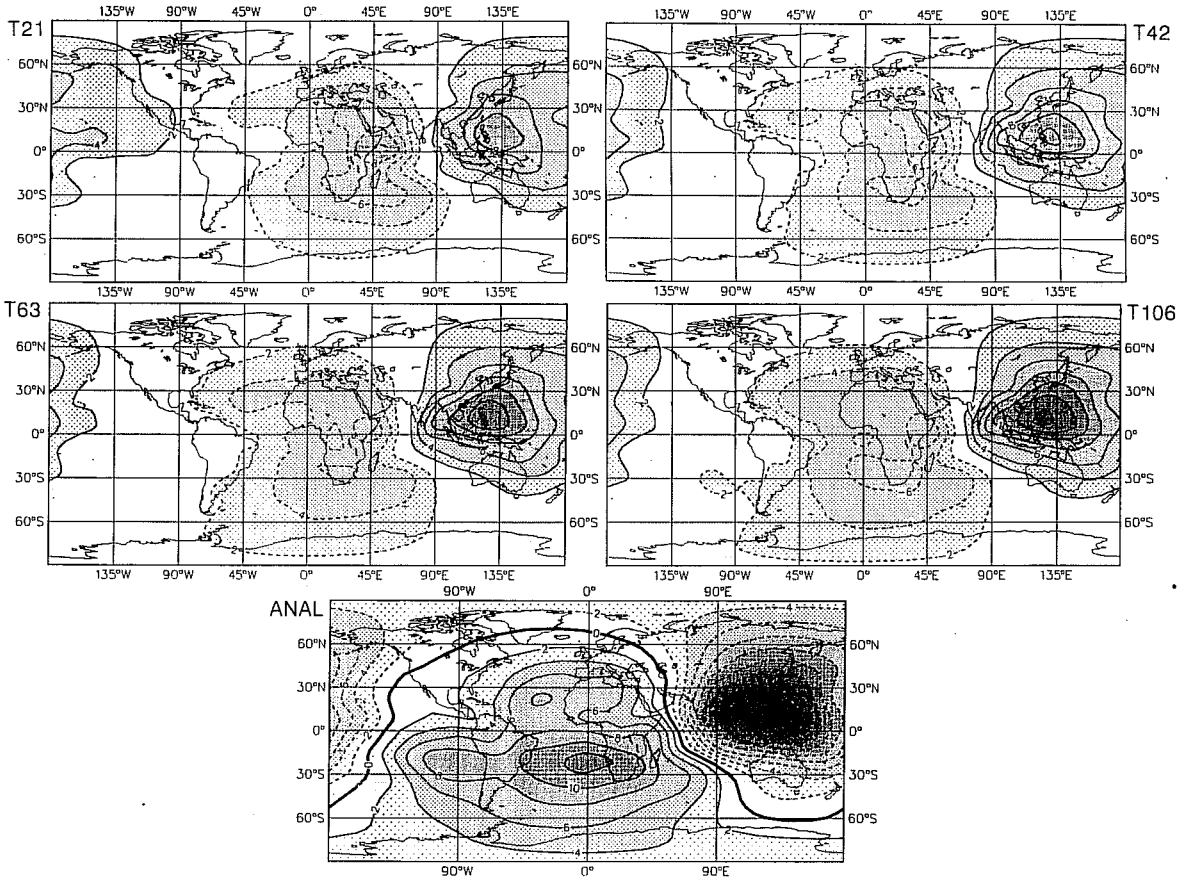

Fig.11 30-day mean 200 mb velocity potential errors for the four model resolutions and verifying analysis (bottom) in the season April-September 1985. Contour interval 2×10^6 m^2s^{-1} for errors and 4×10^6 m^2s^{-1} for the analysis; negative errors and analysed divergence dashed.

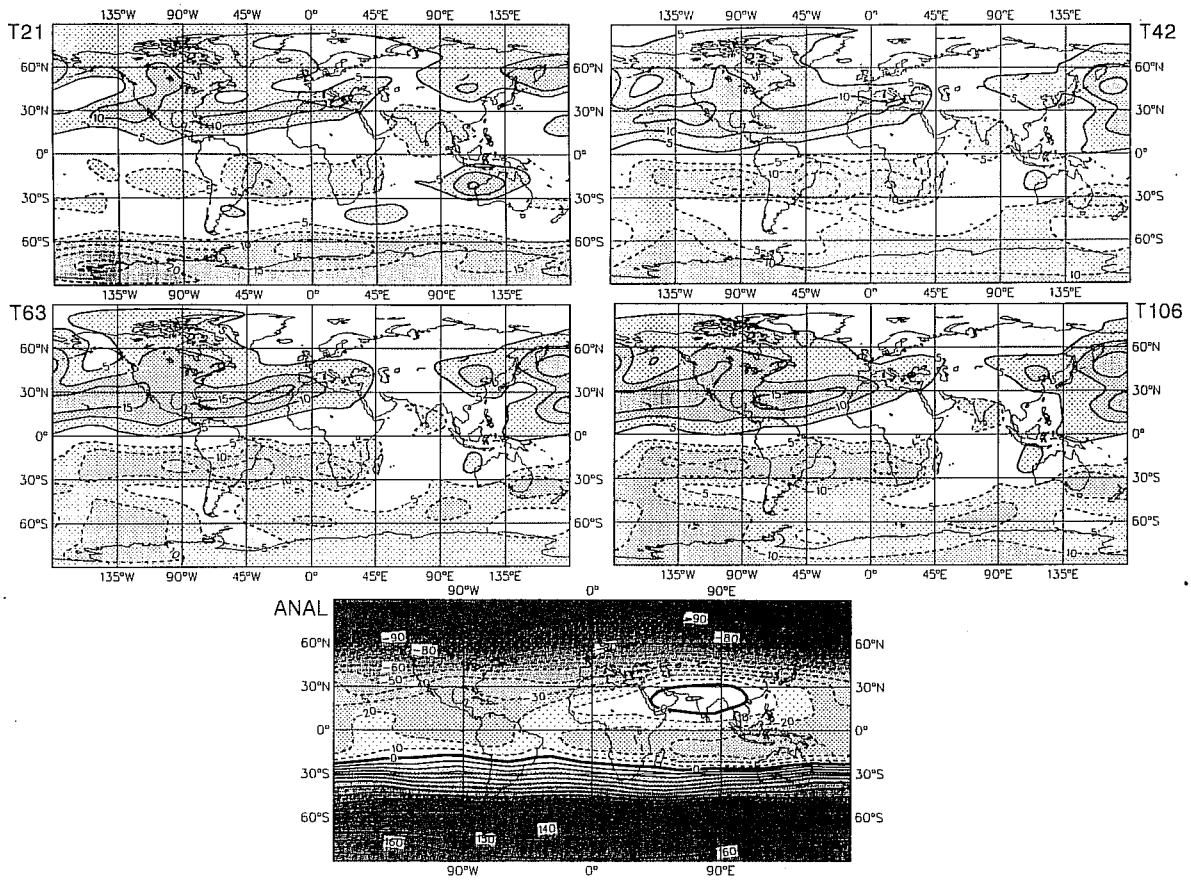

Fig.12 30-day mean 200 mb stream function errors for the four model resolutions and verifying analysis (bottom) in the season April-September 1985. Contour interval 5×10^6 m^2s^{-1} for errors and 10×10^6 m^2s^{-1} for the analysis.

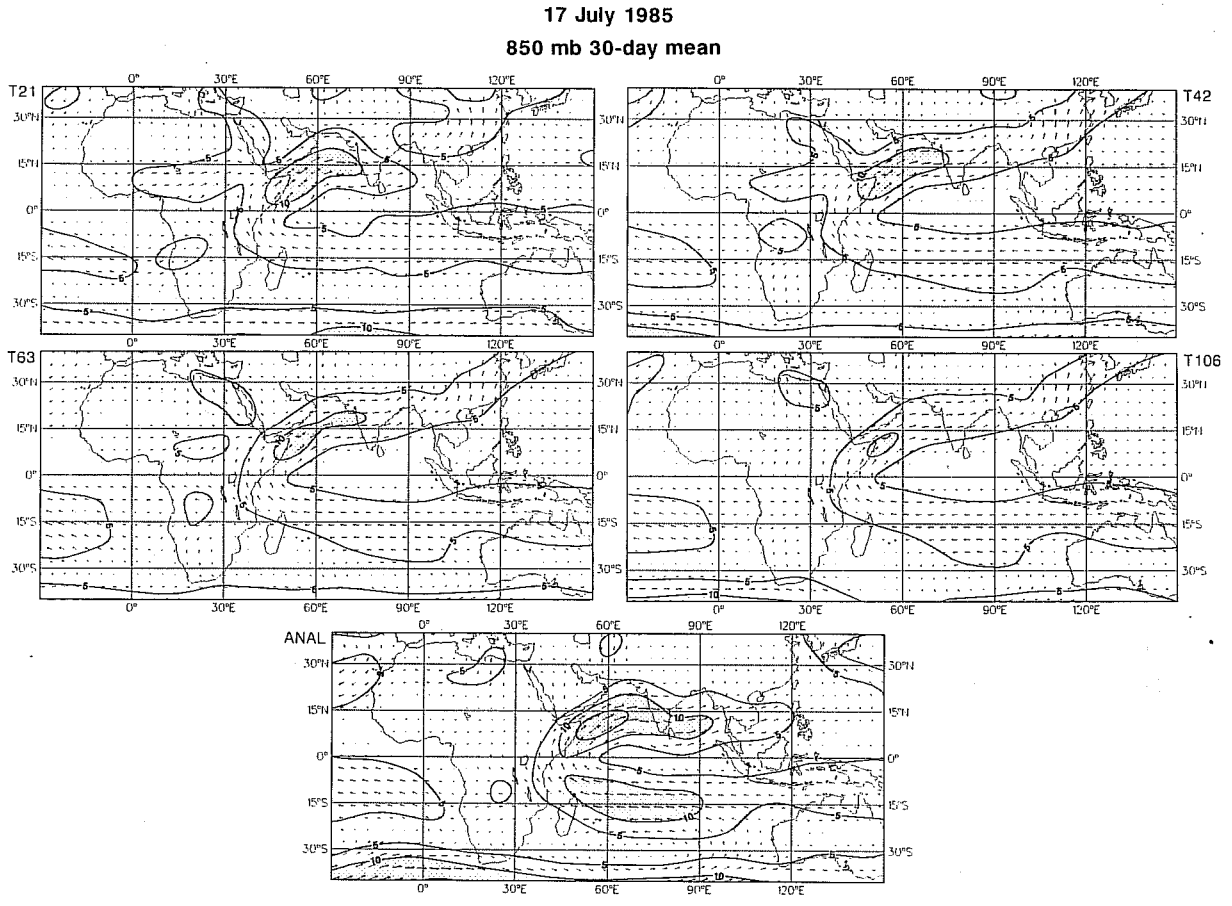

Fig.13 30-day mean wind field at 850 mb for 17 July 1985 for all four model resolutions and verifying analysis (bottom). Isotachs every 5 ms^{-1}, and shaded above 10 ms^{-1}.

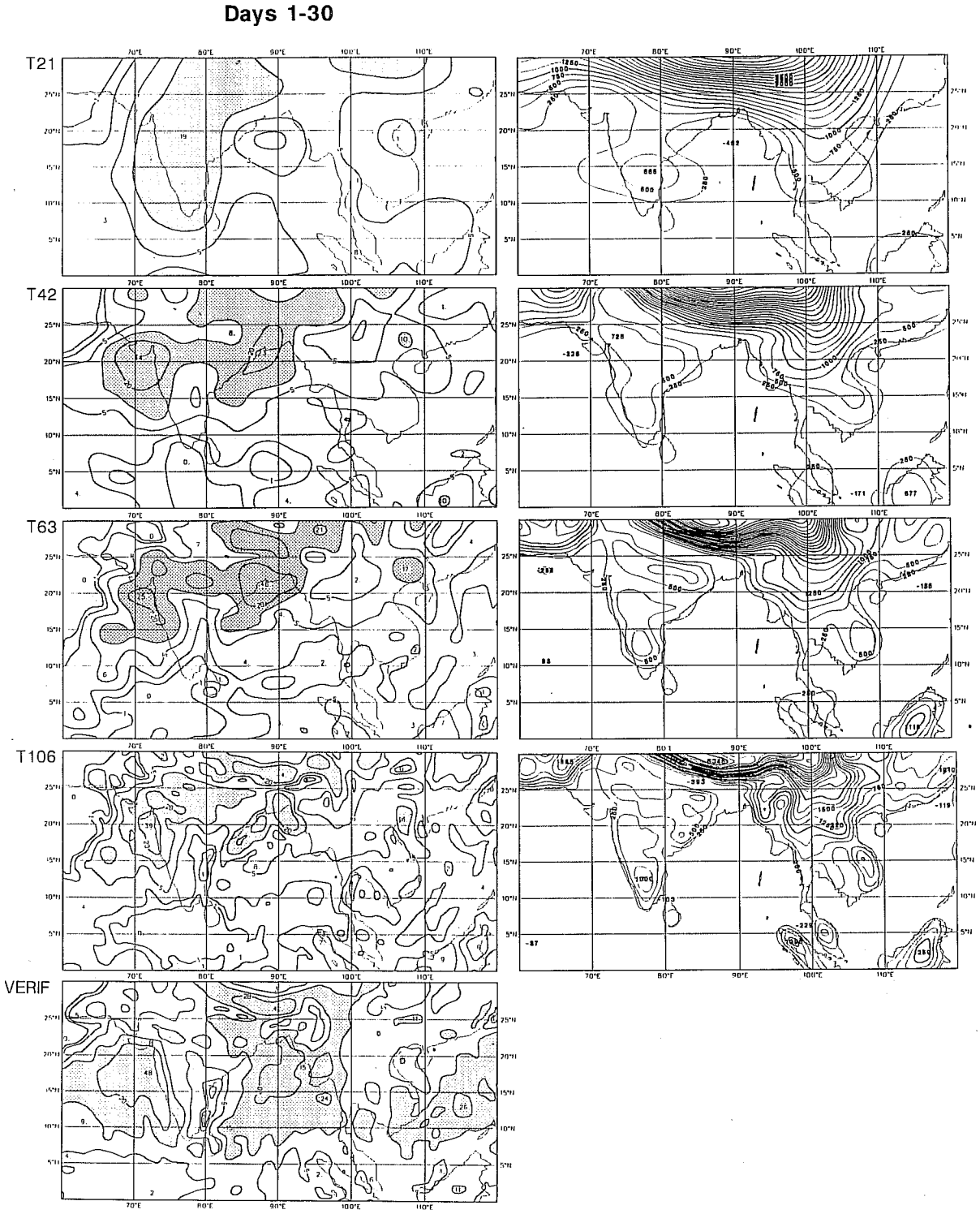

Fig.14 Left: 30-day mean precipitation over the Indian region for all four model resolutions from integrations initialised on 17 July 1985 and 'verification' derived from the operational 24-hour accumulated rainfall over the period which covers 30-day integration (bottom left). Right: orography over the same region. Contour interval 1, 2, 5, 10, 20 mmday^{-1} for the rainfall and 250 m for the orography.

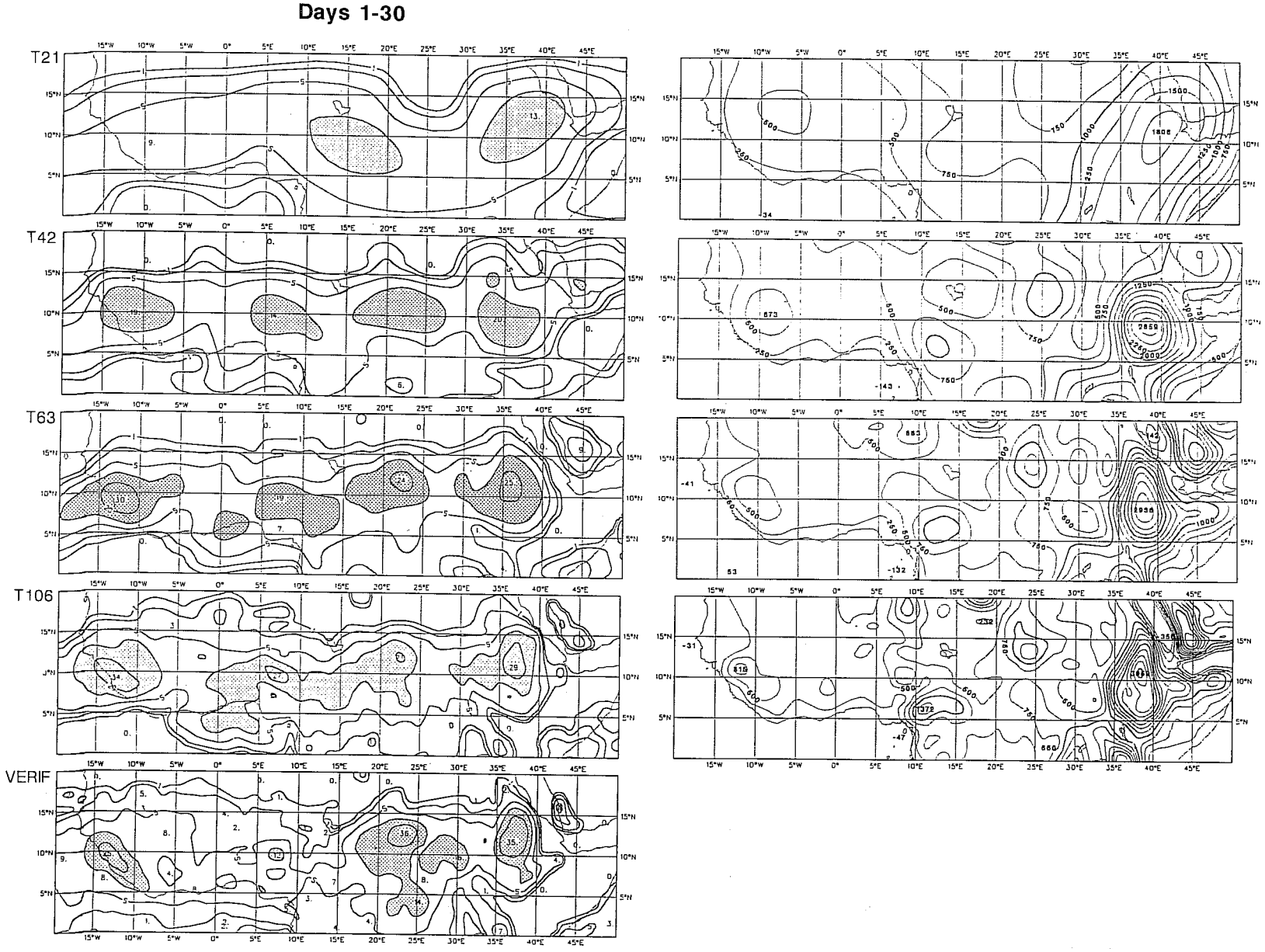

Fig.15 As Fig.14, but for African Sahel region.

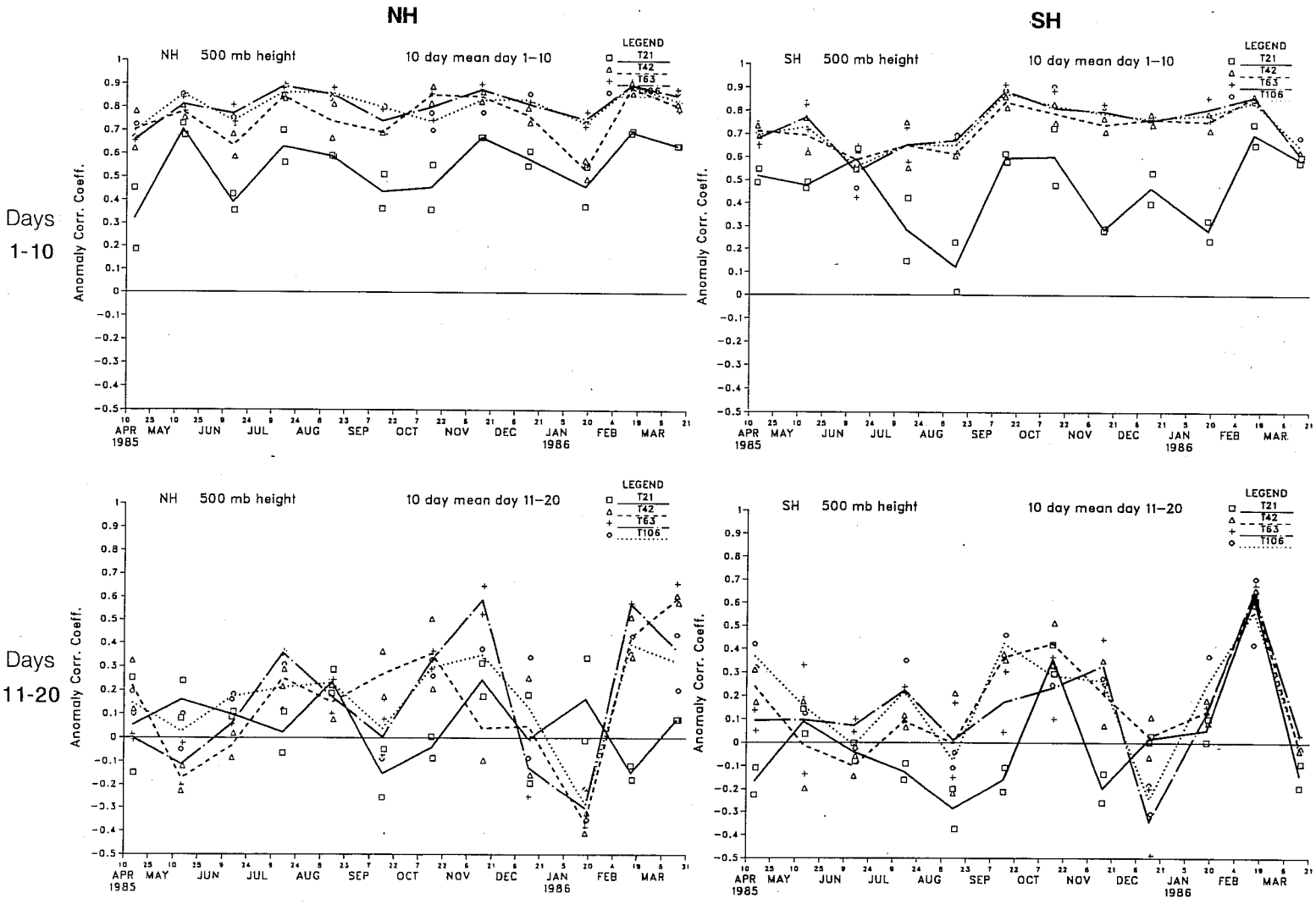

Fig.16 Northern hemisphere 10-day mean (top) and 30-day mean (bottom) anomaly correlation coefficient of 500 mb height field for forecasts throughout the year from April 1985 to March 1986. Left: northern hemisphere; right: southern hemisphere. Solid line T21; dashed line T42; dot-dashed line T63; dotted line T106.

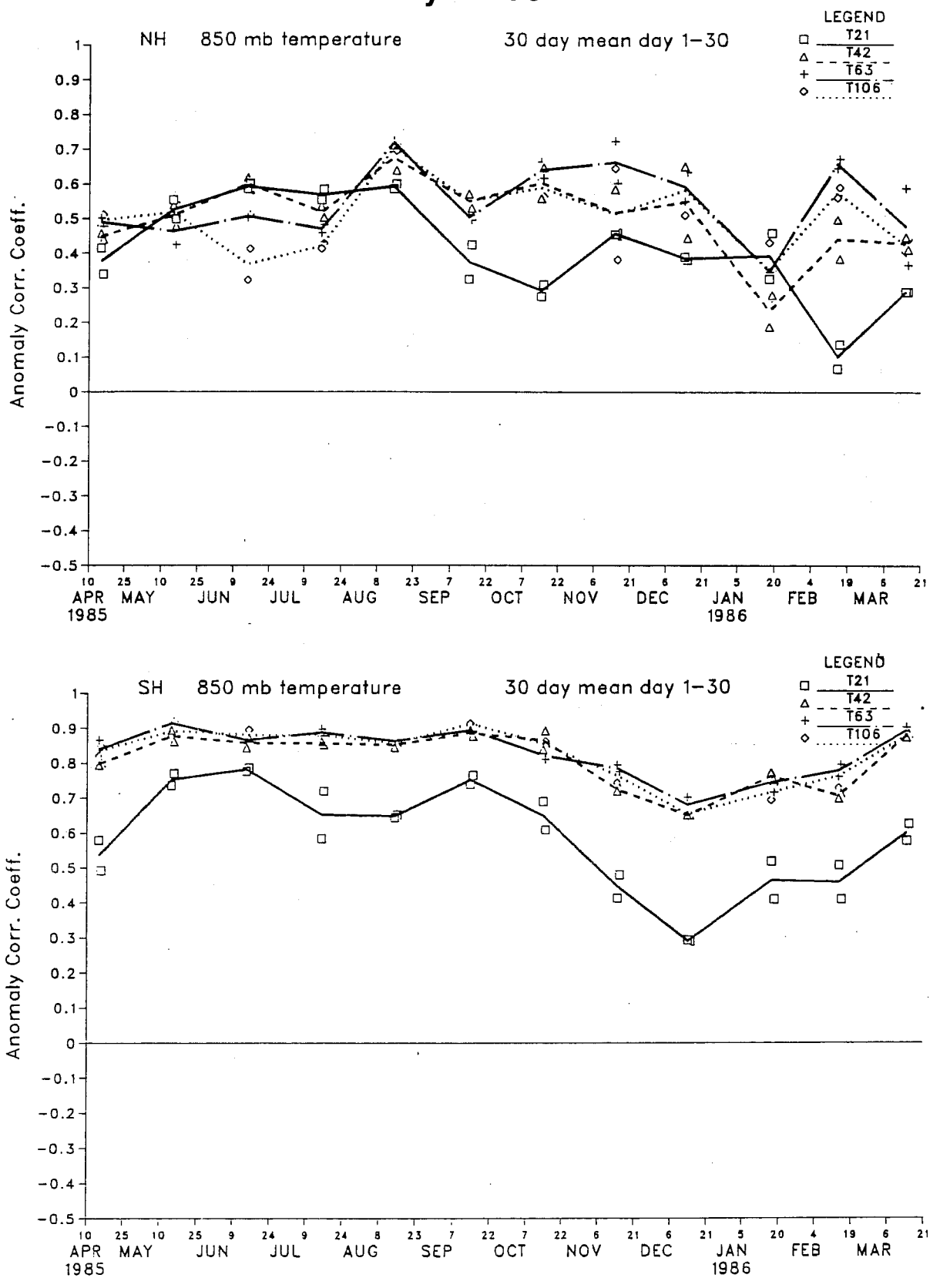

Fig.17 As Fig.16 but for 30-day mean 850 mb temperature. Top: northern hemisphere; bottom: southern hemisphere.

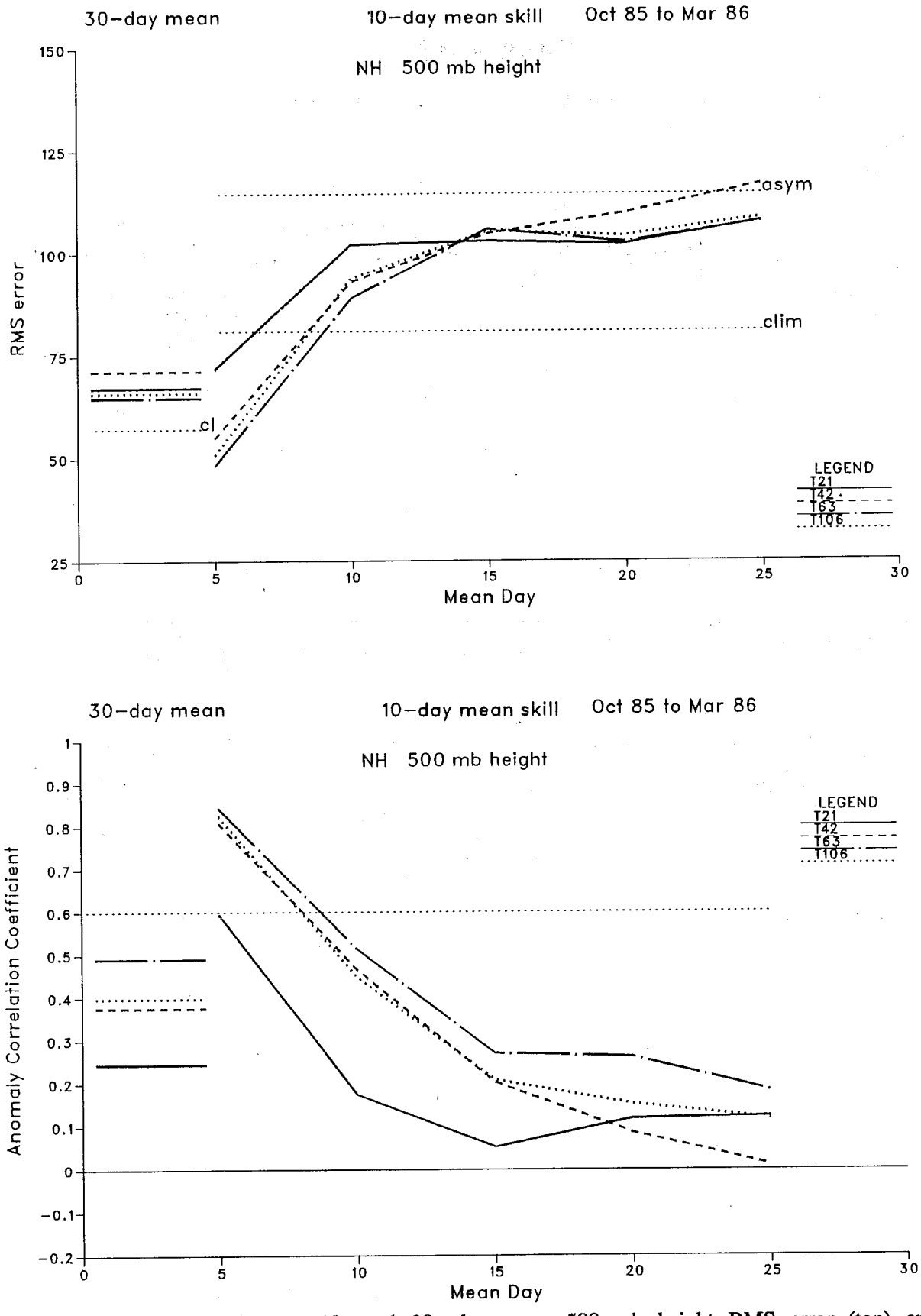

Fig.18 Northern hemisphere 10- and 30- day mean 500 mb height RMS error (top) and ACC (bottom) during OM period. Solid line T21; dashed line T42; dot-dashed line T63; dotted line T106.

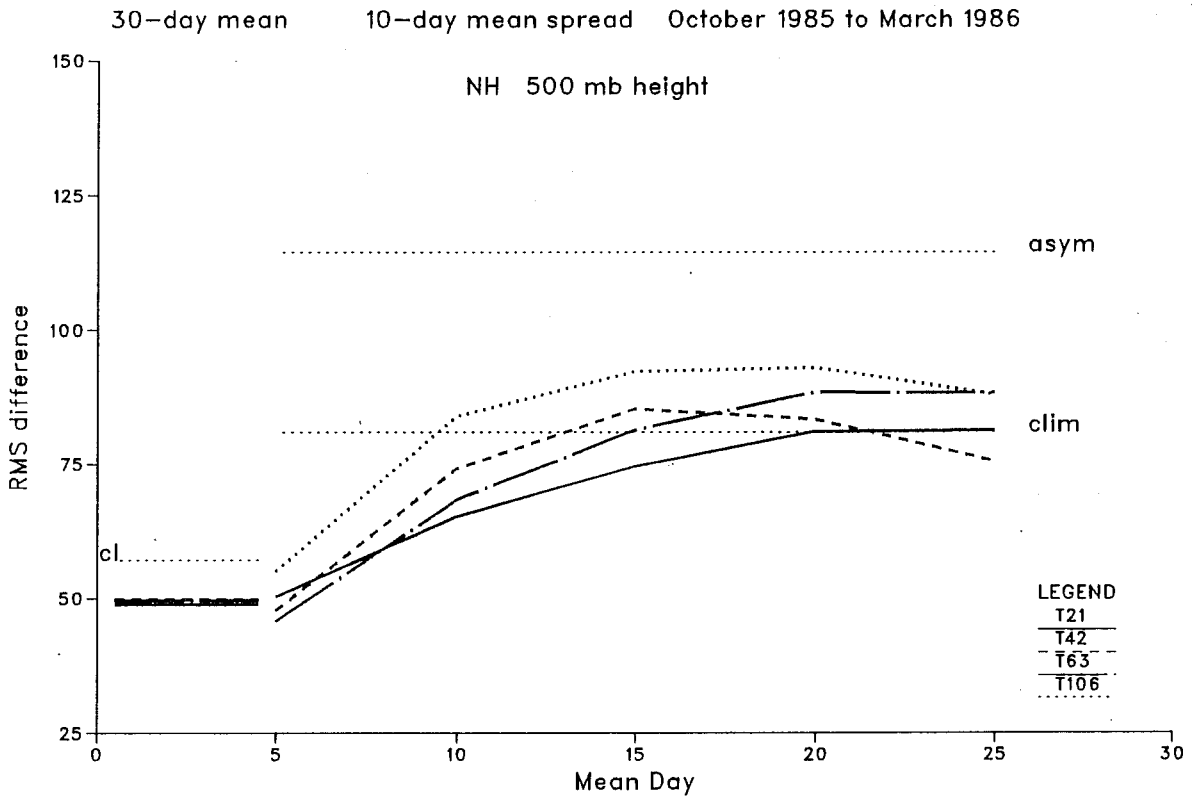

Fig 19. As Fig 18a but for RMS spread between forecasts initialised 24hrs apart.