TECHNICAL MEMORANDUM

422

Systematic Errors in the ECMWF Forecasting System

Thomas Jung and Adrian Tompkins

Research Department

Paper presented to the 32nd Session of the ECMWF Scientific Advisory Committee

October 2003

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

Abstract

In this study some of the key systematic errors of the ECMWF model are described using a wide range of operational forecasts and research experiments. Focal points of this study are (1) systematic errors of the latest model cycle (26r1), (2) the evolution of systematic model errors during the last two decades, (3) characteristics of the growth of systematic errors throughout the forecast including medium-range, monthly, and seasonal time scales, and (4) the sensitivity of systematic model errors to horizontal resolution. Results are presented for temperature and humidity fields, cloud related parameters, the atmospheric circulation, and the kinetic energy of transient eddies. Furthermore, systematic errors of some key phenomena are discussed, for example, synoptic activity, Euro-Atlantic blocking, and the Madden-and-Julian oscillation (MJO).

A concise summary is somewhat hampered by the fact that systematic error characteristics strongly depend on the parameter, region, vertical level, season, and the forecast range being considered. In general, though, it has been found that the magnitude of systematic model errors has been reduced considerably, particularly in the short-range and medium-range. The spatial structure of systematic model errors, however, remains more or less unchanged for most of the parameters. Model improvements are primarily manifest through a reduction of the rate at which systematic errors grow. Moreover, for many parameters the spatial structure of systematic errors undergoes considerable changes from the short-range to the medium-range and extended-range. The sensitivity of systematic model errors described in this study to the resolution is relatively small—the only exception being the MJO.

The most prominent systematic errors described in this study are the following: a warm upper-stratospheric bias; the loss of humidity in the tropical troposphere; lack of stratocumulus and too little cloud cover in midlatitudes; the development of a large anticyclonic bias in the central North Pacific; the underestimation of kinetic energy of the transient eddies; a pronounced underestimation of synoptic activity in high-latitudes; the underestimation of the frequency of occurrence of blocking in the Euro-Atlantic sector; and the failure of the model to simulate slowly-propagating MJO-related velocity potential anomalies.

Finally, possible directions for future research are outlined, aiming to understand the origin and reduce some of the systematic model errors described in this study.

1   Introduction

The aim of numerical weather forecasting is to predict the weather several hours to several days in advance using numerical models of the atmosphere. Uncertainties in the initial conditions, or analyses, as well as model deficiencies lead to the development of forecast errors. The time-dependent forecast error e_t can be expressed as the difference between the forecast f_t and our best estimate of the truth o_t (e.g, the analysis), that is, $e_t = f_t - o_t$. Continuous monitoring of forecast errors at ECMWF reveals that e_t has been substantially reduced during recent years (e.g. Simmons and Hollingsworth, 2001). This reduction of forecast errors is partly due to improved initial conditions and partly due to improvements of the numerical model. In general it is not straightforward, though, to separate the influence of improved analyses from those due to improved model formulations to the total reduction of e_t. (This has become even more the case as models are used in data assimilation schemes to determine the analysis.)

A relatively simple way to identify model errors is to focus on systematic differences between forecasts f_t and estimates of the true state of the atmosphere o_t. The estimated systematic error for the diagnostic d is given by the following expression:

$$\hat{d}_{se} = \hat{d}(f_t) - \hat{d}(o_t),\qquad(1)$$

where the hat in $\hat{d}$ indicates that we are dealing with an estimate of the true value of the diagnostic d. For $\hat{d}_{se}$ to reflect systematic *model* errors, two requirements must be fulfilled. First, the validation dataset o_t must be unbiased. Obviously, the degree to which o_t can be considered to be unbiased depends on the parameter that is diagnosed. In this sense Eqn. (1) is likely to be a reliable diagnostic for the geopotential height, for example,

whereas this is less clear so for a quantity like the cloud liquid water content for which direct measurements are very limited. Second, Eqn. (1) is subject to sampling variability due to the finite length of the time series. To illustrate this point, consider one month of daily D+10 forecasts of Northern Hemisphere 500 hPa geopotential height fields (Z500, hereafter). If this month was anomalously blocked, for example, then this is reflected in $\hat{d}(o_t)$. Further suppose that the predictability of Z500 has been largely lost for this particular month at D+10. Then $\hat{d}(f_t)$ is likely to represent less blocked conditions. As a consequence, even for a perfect model, $\hat{d}_{se}$ is likely to (wrongly) suggest that the model has a zonal bias.

The perhaps simplest and most widely used diagnostic to describe systematic errors is the average μ for which Eqn. (1) becomes $\hat{\mu}_{se} = \hat{\mu}(f_t) - \hat{\mu}(o_t)$. $\hat{\mu}_{se}$ is an estimate of the average systematic model error and will be used extensively throughout this study. However, given the importance of *variations* for numerical weather forecasting like those associated with synoptic systems we will also consider systematic errors for other diagnostics like, for example, the variance or the power spectrum.

Throughout the years systematic model errors have been extensively monitored and documented at ECMWF. Zonal average systematic temperature and wind errors, for example, are discussed by Klinker and Sardeshmukh (1992) along with their possible origins. Further, systematic errors of the ECMWF model to simulate the observed blocking characteristics over the Northern Hemisphere are described by Tibaldi and Molteni (1990) and a discussion of systematic errors of extratropical variability is given in Ferranti et al. (1994). A detailed description of systematic errors of model cycle 18r6 is given by Brankovic et al. (2002), and the influence of atmosphere-ocean coupling on systematic errors is discussed in Anderson et al. (2002).

As in the previous studies described above, here systematic errors of key parameters and key phenomena are described for the latest cycles of the ECMWF model. However, rather than being merely an update of previous work we go a step further by also addressing the following questions:

- How did systematic model errors evolve over the years?

- How do systematic model errors grow throughout the forecast?

- How sensitive are systematic model errors to horizontal resolution?

To this end a variety of different operational ECMWF forecast products and research experiments are diagnosed.

Given the large number of different forecast products issued by ECMWF, however, a comprehensive study of systematic errors for all available parameters would be prohibitive to carry out. Therefore, while trying to answer the above questions, the focus is on some of the key parameters and key phenomena. Systematic errors of temperature, humidity, and the geopotential height structure are discussed is some detail throughout this study. Apart from being parameters of general interest, relatively long and reliable validation datasets are available for these quantities—thus making Eqn. (1) a robust estimator of systematic model errors. In recent years much effort at ECMWF has been put into the improvement of physical aspects of the model. However, in general it is not straightforward to relate improvements of the representation of physical processes to parameters like, for example, geopotential height. Therefore, a separate section of this paper gives an assessment of the ECMWF model in simulating cloud related features of the atmosphere.

Moreover, there are atmospheric key phenomena which have to be simulated accurately because of their paramount direct or indirect importance to the users of numerical weather forecast products. Whereas it is important to properly simulate the large-scale atmospheric flow, local weather conditions in the extratropics are more directly related to the passage of synoptic systems. Therefore, the ability of the ECMWF model to simulate synoptic activity is described. This is particularly important given that the development of reliable severe weather forecasts is one of the main long-term goals at ECMWF. Furthermore, Euro-Atlantic atmospheric

blocking is addressed. Such an undertaking is justified by its enormous influence on weather parameters in the European area, particularly during boreal winter. Finally, the ECMWF model's ability to simulate the Madden-and-Julian oscillation (MJO, hereafter) is addressed. The MJO is associated with continental-scale organization of convection propagating eastward across the Indian and western Pacific ocean region. At ECMWF, there are at least two reasons why the MJO should be simulated well. First, there is some indication that westerly wind bursts can trigger ENSO events (e.g. Slingo, 1998). Therefore, seasonal ENSO forecasts at ECMWF might crucially depend on the model's skill in simulating intraseasonal variations. Second, there is an indication that medium-range forecast skill in the Northern Hemisphere extratropics depends crucially on how well the Tropics in general, and the MJO in particular, are simulated (Ferranti et al., 1990).

The present study is primarily diagnostic and descriptive. However, possible origins of the systematic model errors will be briefly discussed in passing. Furthermore, a more detailed description of future research to unravel the origins of systematic model errors will be given in the Discussion section.

The manuscript is divided into three main sections. First, the datasets used throughout this study are described in section 2. Then, the results will be presented in section 3. This section contains a description of the systematic errors of temperature, specific humidity, cloud parameters, the atmospheric circulation, kinetic energy, synoptic activity, atmospheric blocking, and the MJO. Finally the main conclusions of this study will be summarized and discussed in section 4.

2   Data

2.1   Forecast Experiments

In order to address the questions posed in the Introduction a large range of different operational ECMWF forecast products were diagnosed along with a series of research experiments.

One dataset that has been used extensively within this study is based on a set of 6-month long integrations that were started on 1 October and 1 April, respectively, of the years 1962–2001. These integrations were carried out using the latest model cycle 26r1, which was operational since 29 April 2003 to the time of writing of this paper, at a resolution of T_L95L60 and cover the extended winter and summer season. Observed sea surface temperature fields were used as lower boundary condition. These integrations were motivated by three issues. By using the latest model cycle (26r1) for diagnosis it is ensured that the systematic error structure of the most advanced ECMWF model is described. Additionally, by performing the integrations for 40 different years (1962–2001) a large forecast sample is available, thereby reducing sampling uncertainties. Finally, by performing 6 month integrations it is possible to investigate the asymptotic systematic error structure by discarding the first two months of the integration during which most of the systematic error growth takes place. It should be noted that diagnosis of these long integrations also helps to summarize systematic errors that might affect seasonal-forecast performance at ECMWF.

To study systematic errors in the medium-range, primarily operational deterministic forecasts are used. The deterministic forecasts are augmented by control forecasts from the Ensemble Prediction System (EPS) and extended-range forecasts based on the Monthly Forecasting System in order to study the influence of resolution and to quantify how systematic forecast errors grow beyond the medium-range. The EPS control ($T_L255L40$) and monthly forecasts ($T_L159L40$) last 20 days and 30 days, respectively. The Monthly Forecasting System is described in more detail in Vitart (2003). In this study, for each of the winter and summer seasons, 108 monthly ensemble forecasts from the period 1990–2002 were used. Most of these forecasts are based on two model cycles, that is, cycle 25r1 and 26r1. Each ensemble comprises of one control and five perturbed forecasts.

The experiments used to assess simulated cloud related parameters are conducted for the summer and winter seasons of 1987 and 1988 partly because of the availability of verification data. The main model cycles being considered are 23r4 and 26r1. The winter period was unusual in that it was marked by a strong El-Nino event which considerably altered the circulation in the Pacific region. The experiments consist of a limited ensemble of six experiments, initialized on the 1st, 2nd and 3rd day of the months of May and November, 1987, for the summer and winter cases, respectively. Each forecast lasts for four months, and the statistics are analysed for the final three months of the ensemble.

The small size of the ensemble, limited to just two seasons, could imply that the conclusions may not be generally applicable to summarize the model's characteristic biases. However, we shall see that many of the problems identified in these experiments corroborate those identified in the literature reviewed in section 3.3. Moreover, the computational frugality of this experimental design has allowed a long term archive to be built up, documenting major model cycle releases over the past 7 years, which will also be briefly examined. In any case, the 6 month comparison period still exceeds that used in many previous studies. For example, Hogan et al. (2001) use a 3 month period for their comparison to radar data, Chevallier and Morcrette (2000) use six weeks of data to make their conclusions, while Tselioudis and Jakob (2002) examine just the month of April 1992.

2.2 Verification Datasets

Two main sources of verification datasets are used. First, operational medium-range forecasts are compared with operational analyses only. This ensures that the models are compared against their own analyses. Monthly and seasonal integrations are compared to ERA-40 data. Brief overviews of ERA-40 and its quality are given elsewhere (see Uppala, 2002, and other papers in this issue). The ERA-40 model is based on a slightly modified version of model cycle 23r4 which was in operational use at ECMWF during the period 12 June 2001 to 22 January 2002. The horizontal resolution of the model used for reanalysis is $T_L 159$, and 60 levels in the vertical were employed.

The model radiation diagnostics are compared to the top of the atmosphere (TOA, hereafter) shortwave albedo and longwave net fluxes, which are both derived from Earth Radiation Budget Experiment (ERBE) products (Ramanathan, 1987). The total cloud cover is compared to the C2 product of ISCCP (Rossow and Schiffer, 1991; Rossow and Garder, 1993; Rossow and Schiffer, 1999). Total column water vapour, surface winds and liquid water path (LWP) amounts are compared to SSM/I retrievals (Wentz, 1997)[1].

3 Results

3.1 Temperature

In this section systematic errors of the ECMWF model in simulating the thermal structure of the atmosphere are described.

3.1.1 Systematic Errors

To start with, systematic temperature errors are considered using 6 month integrations of the ECMWF model (cy26r1 at $T_L 95L60$) over the period 1962–2001. For every year forecasts were started on 1 October and

---

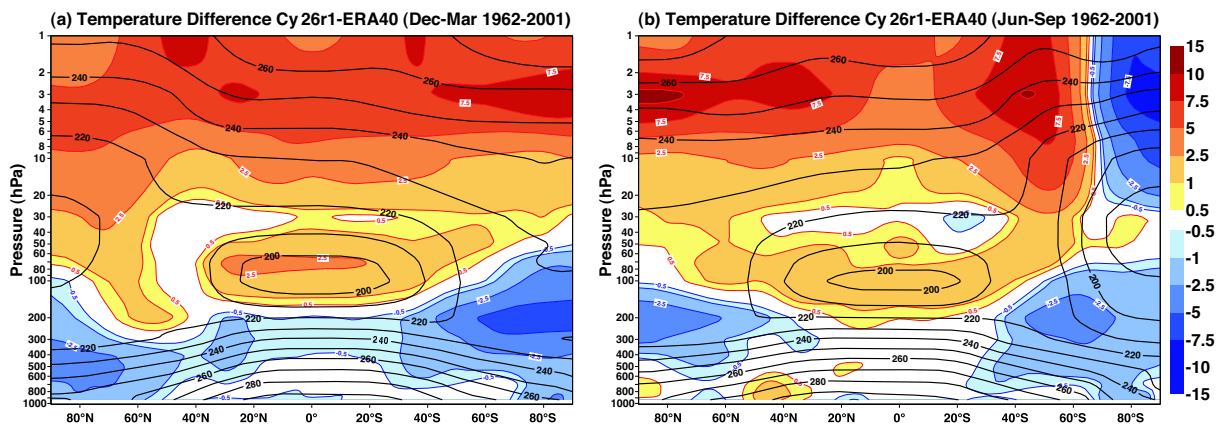[1]Remote Sensing Systems version 4 product at http://www.remss.com/

Figure 1: Zonal mean temperature structure from ERA-40 (contours in K) and difference in the zonal mean temperature structure (shading in K) between the ECMWF model (cy26r1 at T_L95L60) and ERA-40 data for the period 1962–2001: (a) winter (Dec–Mar) and (b) summer (Jun-Sep) season. Forecasts for the winter (summer) season were started on 1 October (1 April) 1962–2001.

1 April, respectively. Since the focus is on the boreal winter (Dec–Mar) and summer (Jun–Sep) season[2], systematic errors described in the following had time to *grow* for more than 2 months.

Differences of zonally averaged temperature fields between the ECMWF model and ERA-40 data are shown in Fig. 1 for winter and summer (shaded). Also shown are zonally-averaged mean temperature fields from ERA-40 (contours). In the upper stratosphere (10–1 hPa) the model produces higher temperatures than those obtained by ERA-40. This holds for both the winter and the summer season. It is not straightforward to identify from these diagnostics whether the model has a warm bias in the upper stratosphere. First, estimates of the thermal structure of the stratosphere that are based on ERA-40 partly depend on the the radiative transfer model used for the reanalysis. In April 2002 a revised shortwave scheme was introduced that is likely to produce higher temperatures in the upper stratosphere than the shortwave code used in ERA-40 (J.-J. Morcrette, personal communication). Therefore, stratospheric temperature differences between model cycle 26r1 and ERA-40 may partly be explained by differences between the two shortwave schemes. (A more detailed description of the difference between the two schemes will be given in section 3.3). Moreover, upper stratospheric temperature estimates in ERA-40 are largely constrained by satellite data. Thus, upper stratospheric temperature fields in ERA-40 are largely determined by model physics before the satellite era (i.e., before the early 1970s). The above diagnostics have been repeated focussing on the last two decades when satellite data were available (1980–2001). Temperature differences between the model cycle 26r1 and ERA-40 are very similar to those for the full period 1962–2001 (not shown).

In the troposphere systematic temperature differences generally are of smaller magnitude. In both seasons negative temperature differences can be found around the tropopause level in the polar regions. These differences tend to reduce the static stability close to the tropopause. Finally, positive temperature differences can be found in both seasons in the Tropics close to the tropopause. In general, though, the ECMWF model seems to perform realistically in the troposphere given that the magnitude of systematic temperature errors barely exceeds 2 K after more than 2 months into the integration.

---

[2]Hereafter, winter and summer refer to boreal winter and summer, respectively
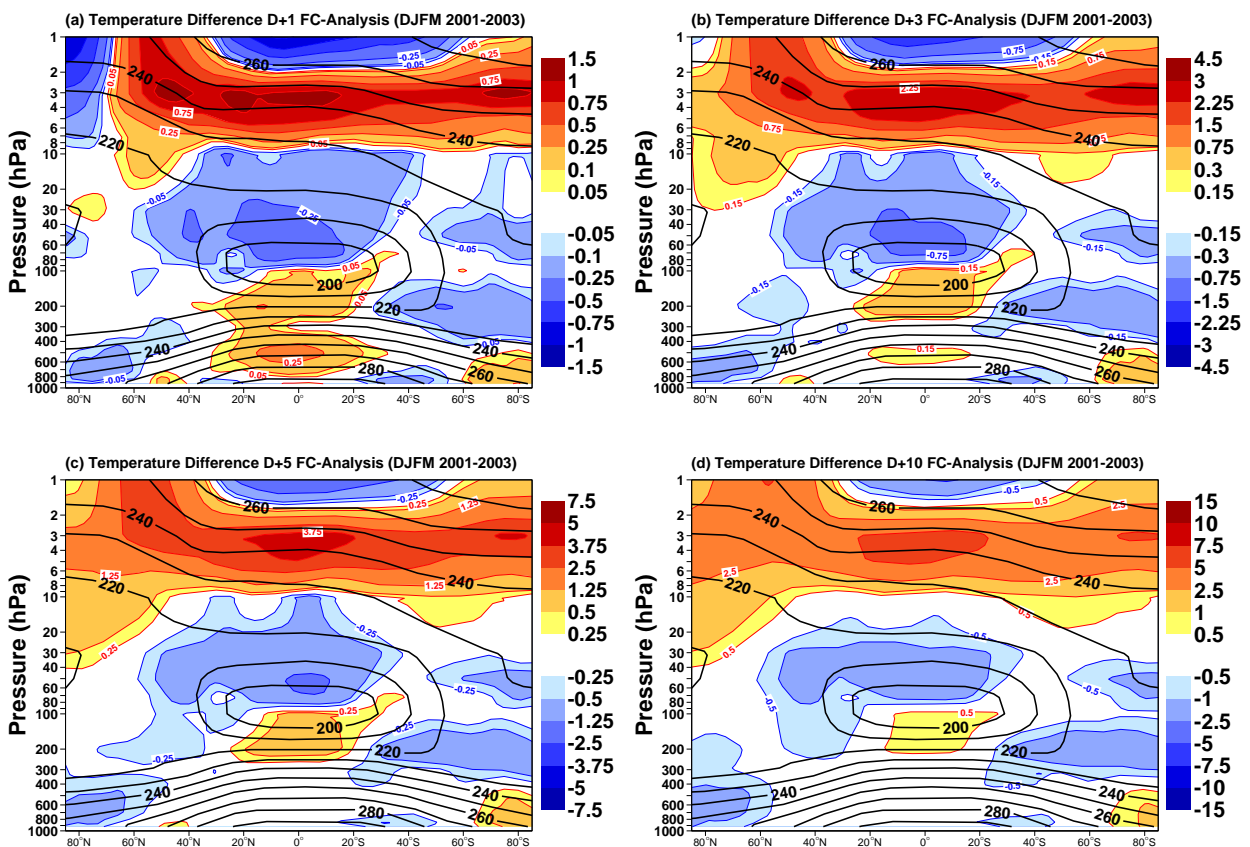
Figure 2: Zonal average temperature structure from operational analyses (contours in K) along with the difference between model forecasts and verifying analyses (shading in K): (a) D+1, (b) D+3, (c) D+5, and (d) D+10 forecasts. Statistics are based on the winters (Dec–Mar) 2001–03. Note, that the shading interval in (a)–(d) changes linearly with the forecast range.

3.1.2 Medium-Range Errors

The evolution of wintertime systematic errors of zonally-averaged temperature fields throughout the medium-range are shown in Fig. 2 for the period 2001–03. As for the asymptotic systematic errors (Fig. 1a), medium-range systematic temperature errors in the troposphere are relatively small; their magnitude at D+10 remains below 1 K. Relatively large errors can be found at D+1 in the tropical belt at around 600 hPa, where positive biases exceed 0.3 K. As will be shown in a later section, in this region the model tends to lose specific humidity quite quickly throughout the first couple of hours of the integration. This positive temperature bias, therefore, is a transient phenomenon that is likely to be associated with the anomalous strong release of latent heat by excessive convective precipitation during the first few hours of the integration.

Most of the other temperature errors exist throughout the whole medium-range and grow either at a linear or sub-linear rate. Systematic temperature errors in the upper (middle) troposphere over the high-latitude Southern (Northern) Hemisphere are examples of errors that grow at an approximately linear rate. The dipole-like temperature error at around the tropopause level in the tropical belt, on the other hand, growths at a sub-linear rate, at least in the medium-range. This dipole is associated with upward displacement of the temperature minimum around 100 hPa.

In the stratosphere the most obvious feature is that the model produces warmer upper stratospheric temperatures

than the operational analysis. The error growth is more or less linear throughout the medium-range. At D+10 the errors are still 2–3 K smaller than the asymptotic systematic errors after two months of integration shown in Fig. 1a. As a follow-up to the discussion of the influence of the shortwave scheme on stratospheric temperatures it is worth mentioning that the temperature comparison for the medium-range (Fig. 2) is clean in the sense that forecasts and analyses are based on the same shortwave schemes. The fact that a temperature error similar to those for the 6 months integrations described in the previous section can be seen in the medium-range suggests that this warm temperature difference is indeed a warm systematic temperature bias.

The negative temperature error in the tropical lower stratosphere is unique in the sense that it does not show up after 2 months of integration (compare Fig. 1a). One way to explain this difference is that the very recent model cycle (26r1) is superior to older model cycles (cy23r3 to cy25r3). It cannot be excluded, though, that the stratospheric warm bias seen in cy26r1 simply takes longer to develop and to dominate the cold bias still prominent in the medium-range. The slow evolution of the warm bias is also consistent with the slow spin-down of the Brewer-Dobson circulation which is know to be too strong in ERA-40 (A. Simmons, personal communication). Such a spin-down leads to reduced Tropical ascent and, thus, due to reduced adiabatic cooling to warmer temperatures.

The structure and growth rate of systematic temperature errors in the medium-range during summertime is very similar to those for the winter season (not shown). In summer, the only exception is for the high-latitude cold biases; the cold biases occur in the upper (lower) troposphere over the respective summer (winter) hemisphere.

3.1.3   Evolution of Temperature Errors

It it worthwhile discussing systematic temperature errors seen in the most recent model cycles in the context of older model versions. Such an intercomparison helps to clarify to what degree systematic model errors have changed in recent years. To this end the evolution of systematic D+3 temperature errors has been studied for three different periods (i.e., 1986–88, 1993–95, as well as 2001–03 and 2000–02). The results are shown in Fig. 3 for the winter (left column) and summer (right column) season. Note, that the verification has been carried out using operational analyses.

From Fig. 3 it becomes evident that zonal mean temperature errors have been reduced considerably during the last couple of years. This conclusion holds for the troposphere as well as for the lower stratosphere. (The upper stratosphere is not shown since these levels were not available in the early 1980s.) Furthermore, the error structure has partly changed since the mid-1980s. The lower-stratospheric warm bias at D+3, for example, that was prominent during mid-1980s is no longer evident. Rather, the most recent model cycles (Fig. 3, lower panel) show a small cold bias at D+3. Also noticeable is the improvement of systematic temperature errors in the tropical troposphere compared to the period 1986–88.

So far, the discussion on secular changes of temperature errors has been rather subjective. In order to allow a more quantitative discussion of secular temperature changes, Euclidean norms of the temperature difference between model forecasts and verifying analyses have been determined separately for the troposphere (1000–300 hPa) and the lower stratosphere (200–10 hPa). Note that the norms were computed taking grid point values for each of the levels considered. In this way, both zonal mean and stationary wave components of systematic temperature errors are taken into account.

Results for the winter season are shown in Fig. 4 for each of the three periods 2001–03 (solid), 1993–95 (dash-dotted), and 1986–88 (dashed). Tropospheric temperature errors have been improved considerably since the mid-1990s. This is particularly true in the short-range, where improvements are partly due to lower values in the very short-range (D+1) and partly due to lower growth rates up to D+3. As a consequence, systematic temperature errors of D+10 forecasts during the period 2001–03, for example, have the same magnitude as systematic
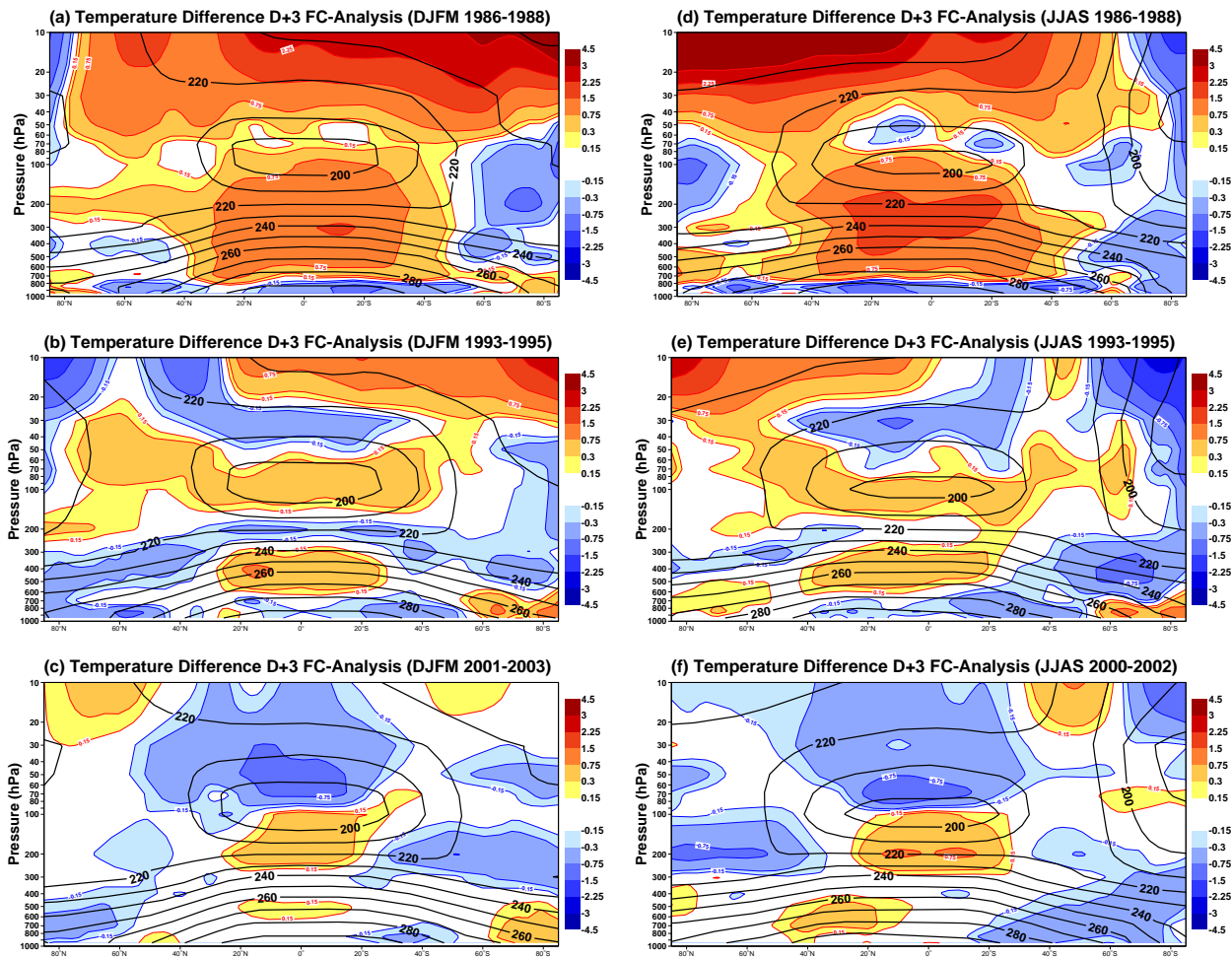
Figure 3: Same as in Fig. 2, except for D+3 forecasts only for the three periods 1986–88 (upper panels), 1993–95 (middle panels), and 2001-03 and 2000-02, respectively (lower panels). Results are shown separately for the winter (Dec–Mar) and summer (Jun–Sep) season in the left and right column, respectively. Note, that in contrast to Fig. 2 results are shown only up to pressure level 10 hPa.

temperature errors at D+4 (D+2) used to have during the period 1993–95 (1986–88). Note, however, that the estimated norms of systematic temperature errors are partly influenced by sampling variations due to the finite length (3 winter seasons each comprising four months) of the time series used. Further, one might expect the sampling fluctuations to be slightly smaller for more recent model cycles given recent improvements in forecast skill (Simmons and Hollingsworth, 2001). Similarly, improved analyses reduce forecast errors particularly in the short-range leading to slightly smaller sampling variations. Overall, however, it seems reasonable to assume that these temporal inhomogeneities are only of minor importance given relatively large differences of the size of systematic errors between the different periods.

A significant reduction of systematic temperature errors is also evident in the lower stratosphere (Fig. 4b). Here, though, the improvements seem to be more gradual over the whole period considered (1986–2002).

It is also interesting to quantify how the spatial structure of systematic temperature errors changes throughout the whole forecast range. To this end, spatial correlations coefficients between D+3 temperature errors
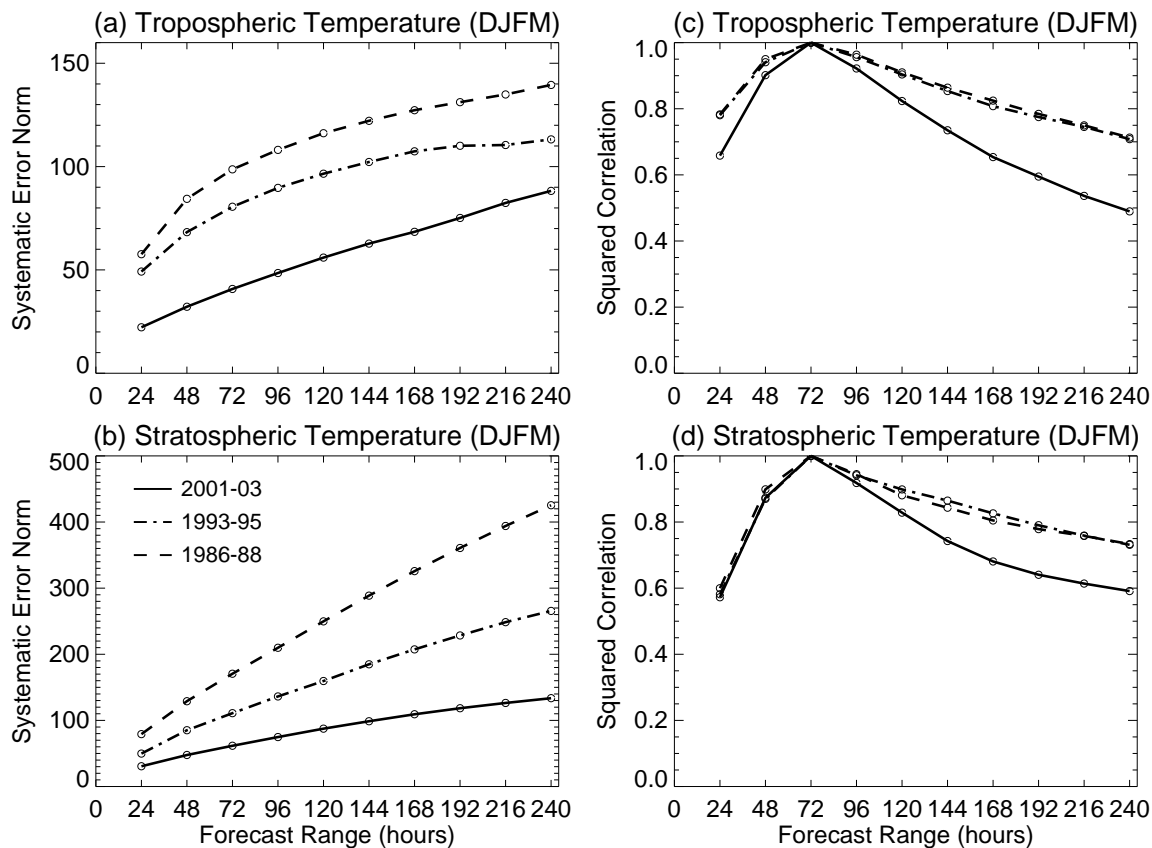
Figure 4: Euclidean norm of wintertime (DJFM) (a) tropospheric (1000–300 hPa) and (b) lower stratospheric (200–10 hPa) systematic temperature errors for different forecast steps and the three periods 2001–03 (solid), 1993–95 (dash-dotted) as well as 1986–88 (dashed). Also show are squared correlation coefficients for (c) tropospheric and (d) lower stratospheric systematic temperature errors between D+3 forecasts and all other forecast ranges (D+1 to D+10).

and those at all other forecast steps (D+1 to D+10) were computed[3]. As for the Euclidean norm, the estimated correlations coefficients are based on all tropospheric and stratospheric grid points, respectively. Squared correlation coefficients, which give the amount of variance explained by a linear model, are depicted in Fig. 4 (right column) for the troposphere and stratosphere. These diagnostics show that the gross features of systematic temperature errors are very similar throughout the first 10 days of the forecast. (Note, that the true systematic error patterns are likely to show smaller changes in their structure, given that the above estimates are influenced by noise.)

The same diagnostics applied to summertime data lead to very similar conclusions (Fig. 5). Notice, though, that the reduction of systematic tropospheric temperature errors during summertime primarily took place from the mid-1980s to the mid-1990s.

---

[3]D+3 temperature errors were taken as a reference to minimize the influence from analysis errors in the very short-range and relatively large sampling fluctuations in the far medium-range.
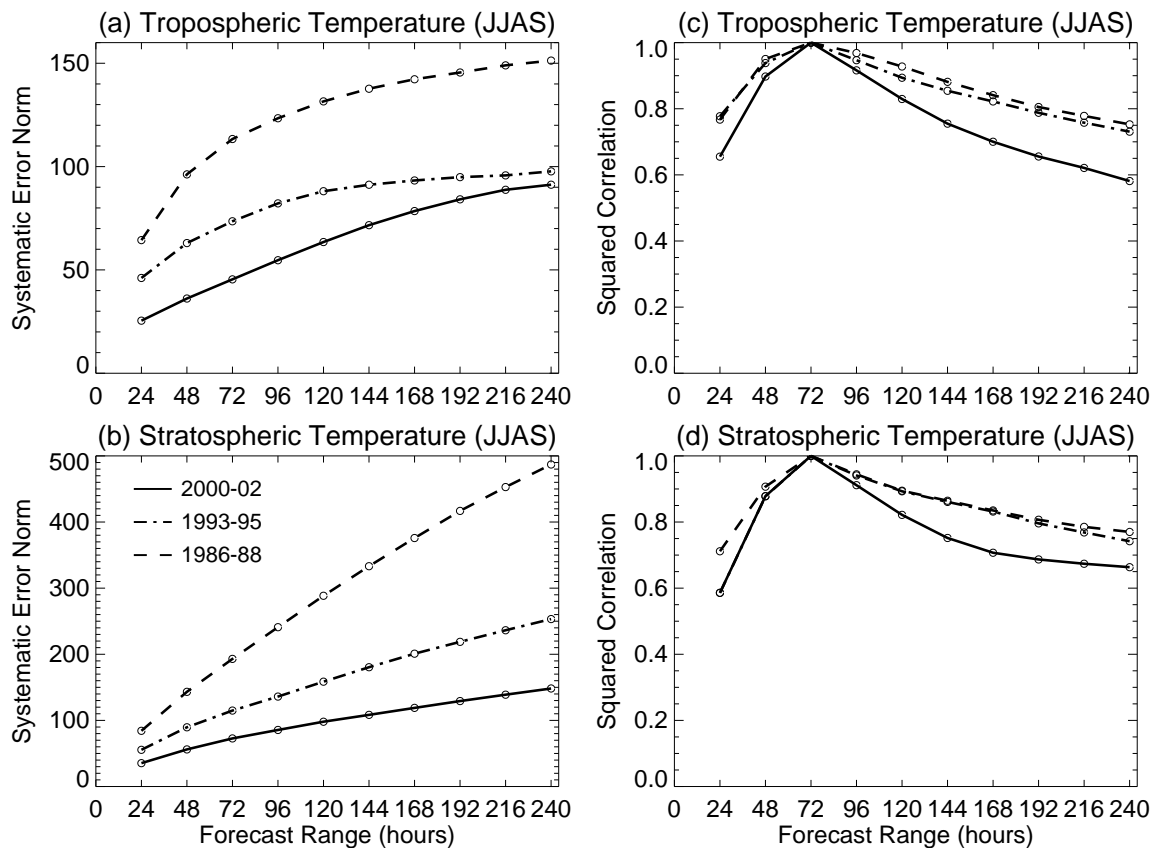
Figure 5: Same as in Fig. 4, except for the summer season JJAS.

3.2 Humidity

3.2.1 Systematic Errors

Zonal averages of systematic errors of specific humidity are shown in Fig. 6 for winter and summer. Results are based on 6 month integrations of model cycle 26r1 at a resolution of T_L95L60. Particularly during the winter season the tropical mid-troposphere is too dry compared to ERA-40 by up to 15%. Secondary maxima are found around 850 hPa at 20^oN and 37^oS. During summertime the spatial structure of the difference pattern is more complex. Still, the model is too dry compared to ERA-40; this difference, however, is less pronounced than it is during the winter season. Moreover, the boundary layer around 20^oN is too moist compared to ERA-40.

The above comparison shows that the model is too dry compared to ERA-40, particularly in the Tropics during boreal winter. This comparison has to be interpreted carefully, however, given that humidity fields from ERA-40 are considerably influenced by the model. In order to shed further light on the realisms of the humidity structure of the model, a comparison of total column water vapour fields between the seasonal runs of the model (cy26r1) and observational estimates based on SSM/I data were carried out for winter and summer of the period 1988-99. The results are shown in Fig. 7. Compared to the SSM/I estimates, the model significantly underestimates the total column water vapour by about 10–15% over the eastern Indian ocean, the Maritime Continent, and the western tropical Pacific during winter and summer. Other regions of significant underestimation include the stratocumulus regions in the eastern Pacific and the tropical North Atlantic. The fact that during wintertime the model is too dry compared to SSM/I data in large parts of the Tropics is consistent with zonal mean differences
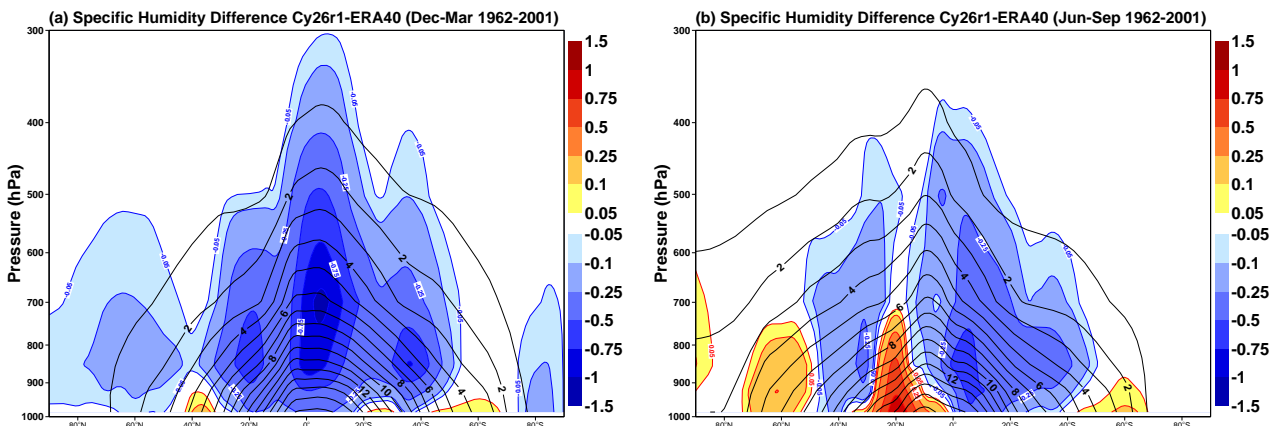
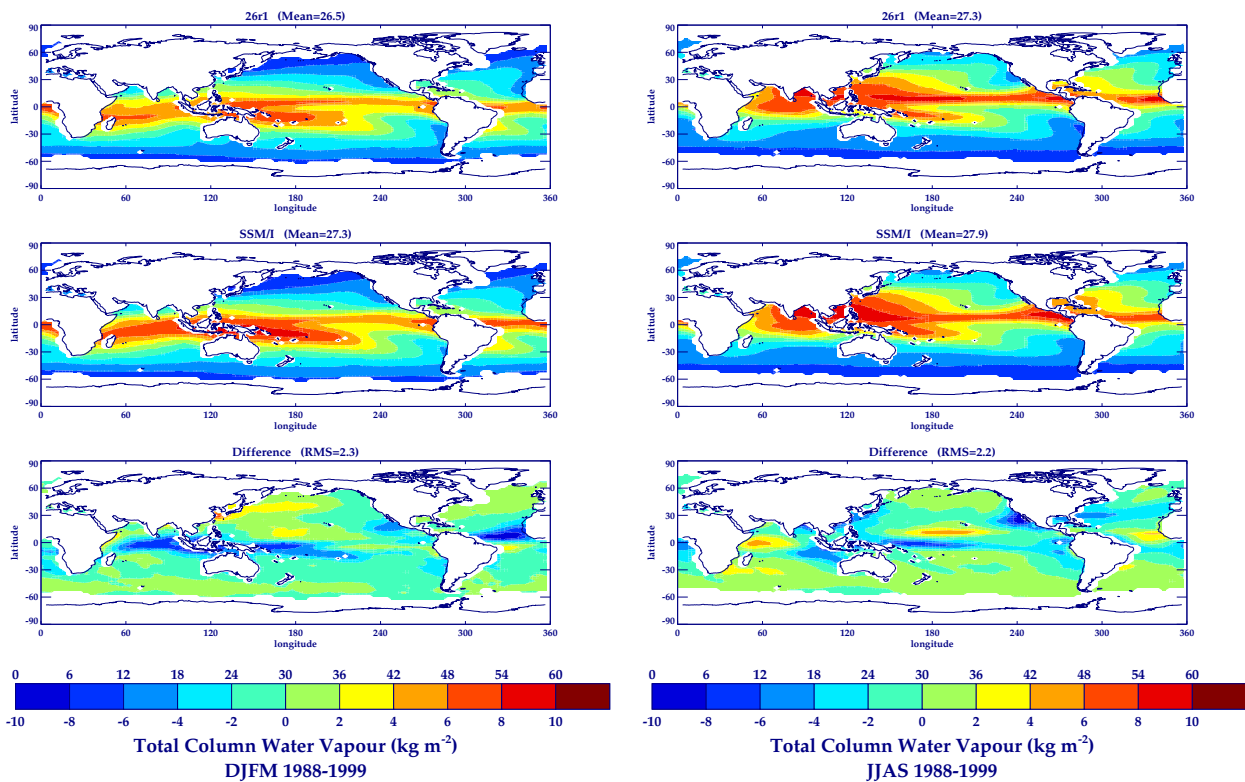*Figure 6: Same as in Fig. 1, except for specific humidity (g kg⁻¹).*

Figure 7: Total column water vapour fields (kg m⁻²) for seasonal runs of model cycle 26r1 (upper panels) and for SSM/I retrievals (middle panels). Also shown are asymptotic systematic total column water vapour errors (model minus SSM/I, lower panels). Results are based on winter (left column) and summer (right column) averages for the period 1988–99.

between the model and ERA-40 data shown in Fig. 6a. During summertime the dry bias of the ECMWF model is less pronounced, at least over the Indian ocean, the Maritime Continent, and the tropical North Atlantic. A pronounced underestimation of the total column water vapour is evident over subtropical oceanic regions of the Northern Hemisphere. Again, these results are consistent with those obtained from the comparison of the model with ERA-40 data (Fig. 6b)
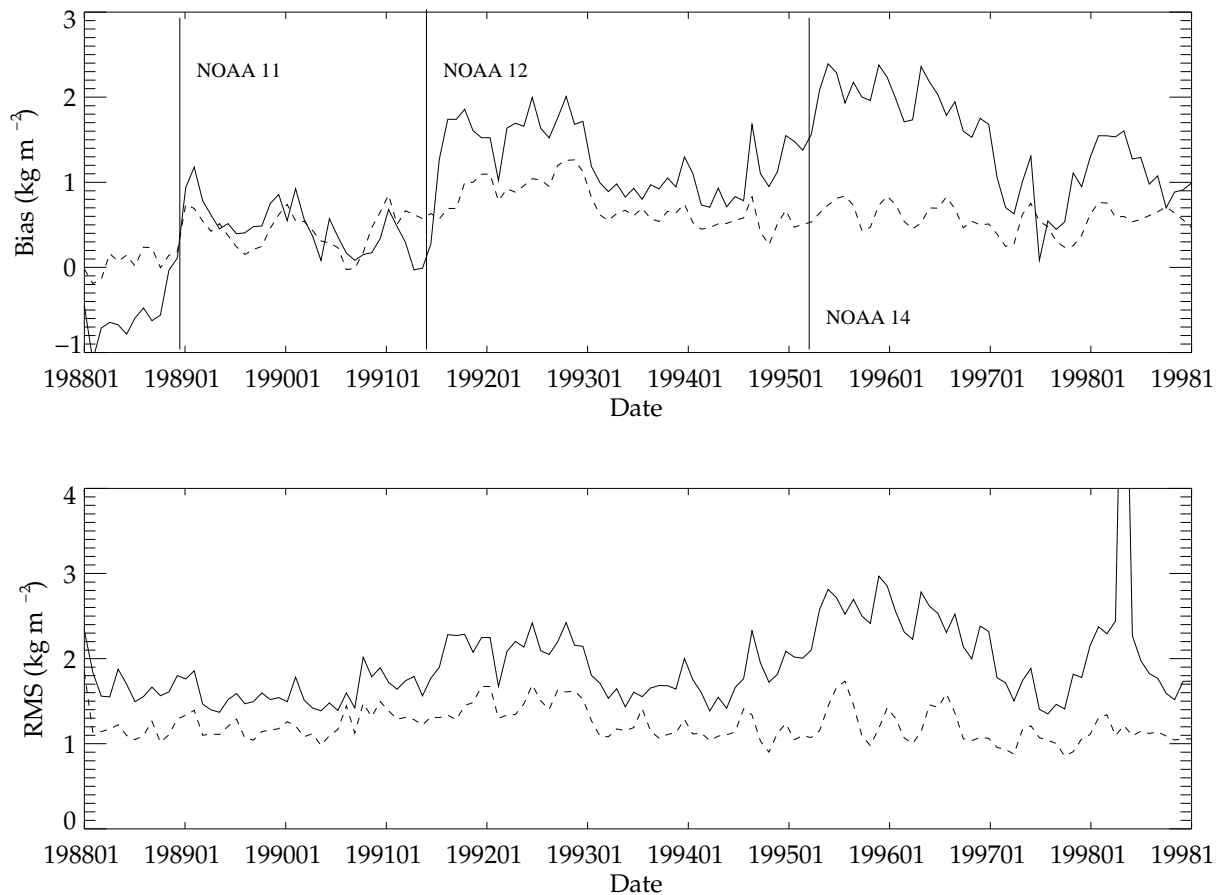
Figure 8: Comparison of 11 years of ERA40 reanalysis monthly mean total column water vapour over oceans to SSM/I retrievals. (Upper panel) time series of biases (positive meaning that ERA40 is too moist) and (lower panel) RMS error between reanalyzed and retrieved fields. The data is divided into the Tropics (latitude < 20º, solid lines) and extra-tropics (latitude > 20º, dashed lines).

Next we address the question how total column water vapour fields from ERA40 compare to those from SSM/I retrievals. This comparison is carried out on a monthly basis for the period 1988–98. It should be mentioned that these observational estimates are not independent, since the ERA40 systems assimilates SSM/I radiances from August 1987 onwards, and thus one would expect relatively small biases. That this is indeed the case can be seen from Fig. 8 (upper panel), showing time series of systematic differences of total column water vapour between ERA40 and SSM/I estimates. Biases in the Tropics amount to a maximum of 2 kg m^{-2}, which is about 5% of the observed mean values. It is notable that apart from 1988, the Tropics in ERA-40 are continually moister than the retrievals. This is surprising considering that the microwave based retrieval techniques tend to give more moisture than radiosonde soundings (Wang et al., 2002; Turner et al., 2003, e.g.). The bias appears to be related to the introduction of the NOAA series of satellites during the 1990s. Although these affect the mean humidity of the Tropics, they do not appear to affect the horizontal variability of the total column humidity as witnessed by the timeseries of RMS errors (Fig. 8, lower panel). In the context of the spin-up problems discussed earlier, the moist or dry biases have most influence in the atmospheric boundary layer, where they can be sufficiently sizable to have a large impact on convective available potential energy (CAPE) and therefore influence the location and intensity of convection (Guichard et al., 2000; Ciesielski et al., 2003).

In summary, the comparison based on SSM/I data suggests that in the tropical region the model indeed has a significant dry bias and that ERA-40 is slightly too moist.
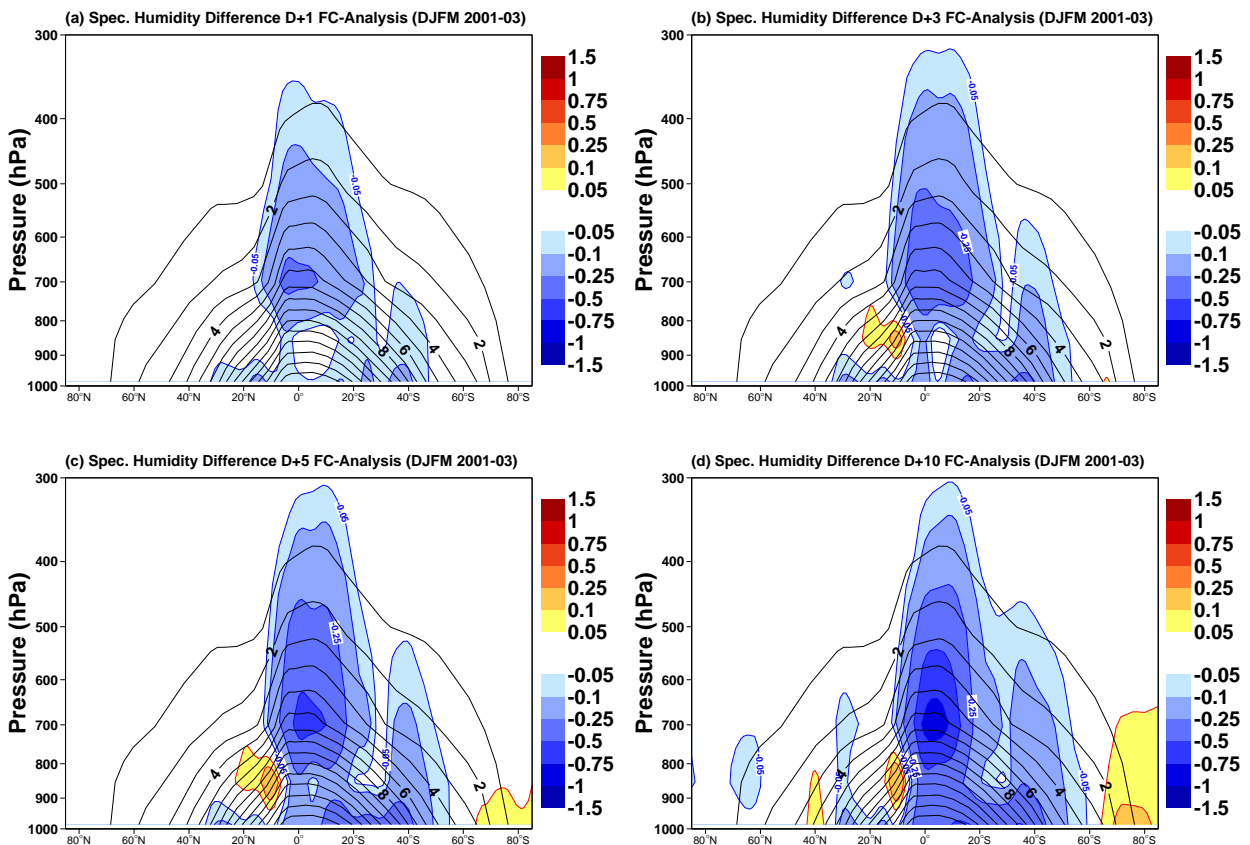
Figure 9: Same as in Fig. 2, except for specific humidity (g kg^{-1}).

3.2.2 Errors Growth in the Medium-Range

In order to shed further light on the model's characteristic to be too dry compared to the reanalysis, the growth of zonally averaged specific humidity differences in the medium-range has been considered. Results are shown in Fig. 9 for D+1, D+3, D+5, and D+10 forecasts for the winters 2001–03. The verification is based on operational analyses. It is evident from Fig. 9 that the model is already much drier (more than 0.2 g kg^{-1} around 700 hPa) than the analyses 24 hours into the forecast. This drying goes along with excessive amount of precipitation in the Tropics (not shown). Recall, that the tropical mid-troposphere shows a warm bias in the short-range (Fig. 2a). This warm bias is consistent with the release of relatively large amounts of latent heat by the exessive amount of convective precipitation during the first few hours of the integration. The drying continues, although at a slower rate, until D+10 by which time the tropical mid-troposphere is drier than the verifying analysis by about 0.75 g kg^{-1}. The fact that asymptotic systematic humidity differences (2–6 months into the integration), discussed in the previous section, are still larger than those seen at D+10 suggests that these differences tend to grow well into the extended-range. (A more detailed description of the growth rate in the medium-range will be given below.)

3.2.3 Evolution of Humidity Errors

Zonal averages of the difference between D+3 specific humidity forecasts and verifying operational analyses are shown in Fig. 10 for the three periods 2001–03 (2000–02), 1993–95, and 1985–88. Results for the winter
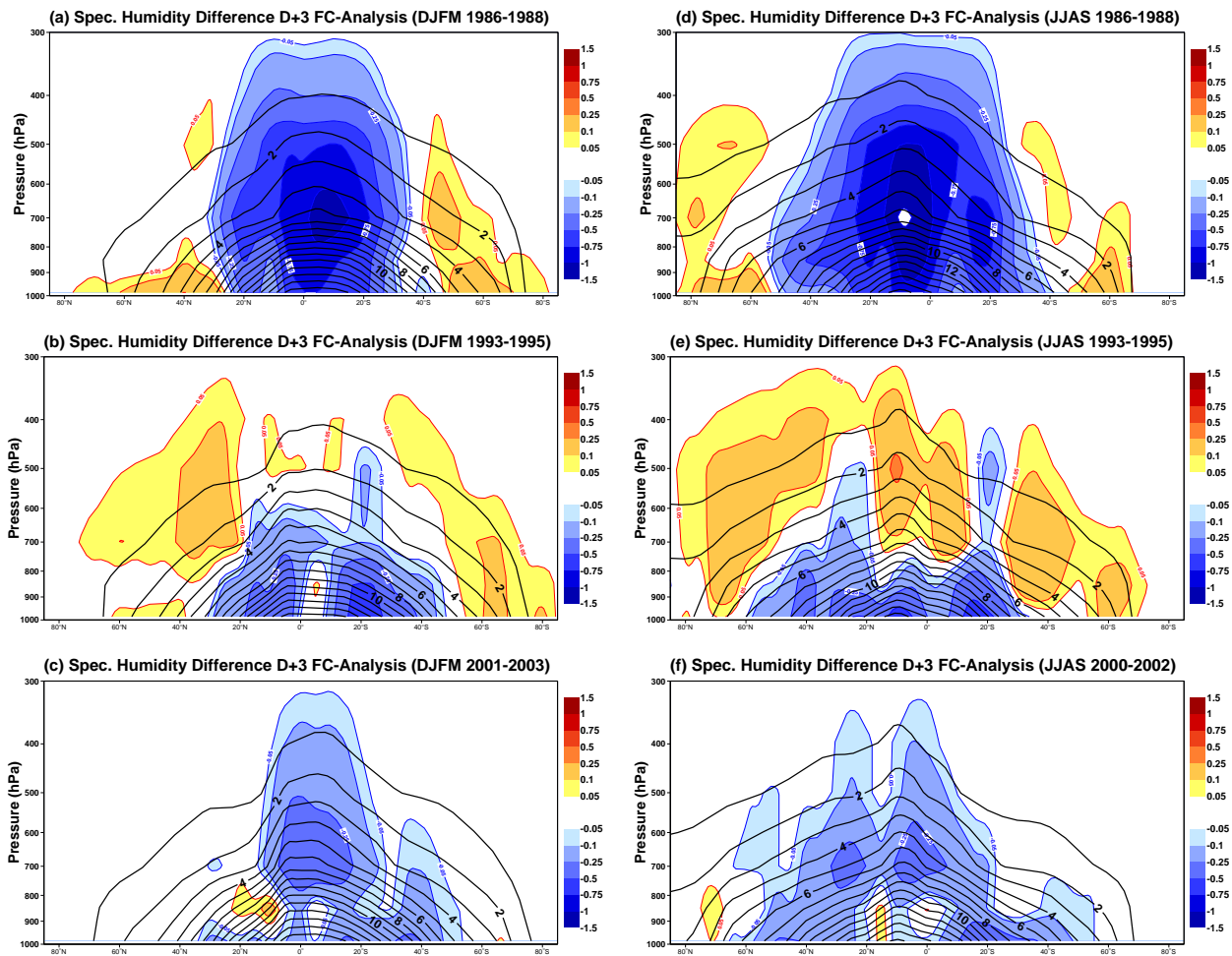
Figure 10: Same as in Fig. 3, except for specific humidity (g kg⁻¹).

(summer) season are shown in the left (right) column. By far the largest differences are found during the mid-1980s, when D+3 tropical specific humidity forecasts locally were drier than the analysis by more than 1 g kg^{-1}. The structure of the difference pattern is very similar for winter and summer. Systematic humidity differences between D+3 forecasts and verifying analyses were considerably reduced by the mid-1990s. During the later period it is primarily in the boundary-layer that the model was drier than the analysis; in the tropical mid-troposphere D+3 forecasts were more moist than than corresponding analyses, at least during summertime. The differences seen in the most recent model cycles have more resemblance with those seen during the mid-1980s. Their magnitude, however, is lower by more than a factor of two.

The growths of the magnitude of differences between tropospheric specific humidity forecasts and analyses throughout the forecast range, measured in terms of the Euclidean norm, is shown in Fig. 11 for the above three periods. First, it is evident that the largest growth takes place during the first 24 hours. Thereafter, the growth rate continues to decline. Based on the Euclidean norm it can be concluded that the magnitude, for example, of differences between D+10 forecasts and analyses of specific humidity of the most recent model cycles have been reduced so much that they are comparable to magnitudes seen during the period 1993–95 (1986–88) for D+2–D+3 (D+1–D+2) forecasts. Further, in absolute terms, the largest reduction of the norm took place from the mid-1980s to the mid-1990s.
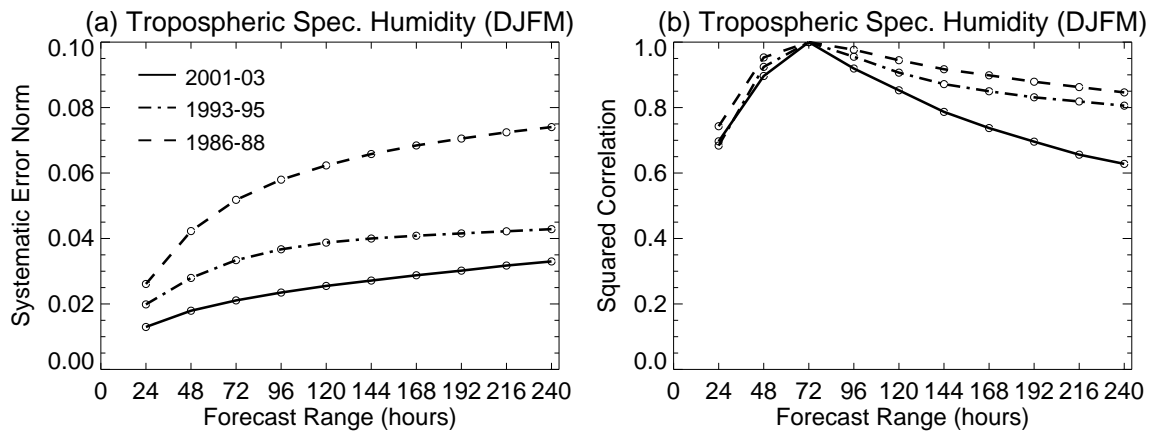
Figure 11: Same as in Fig. 4, except for tropospheric specific humidity (g kg^{-1}).

As for temperature, the spatial structure of systematic humidity differences undergoes rather little changes during the first 10 days of the integration (Fig. 11, right panel). For the summer season a very similar behaviour of the norm and squared correlation coefficient can be found (not shown).

3.3 Cloud Related Parameters

In this section the systematic model errors are examined for longer term atmospheric climate runs conducted at a lower resolution. Only a cursory examination of mean biases is conducted here. The main emphasis is on the principal diagnostics of top of atmosphere longwave and shortwave fluxes, total cloud cover (TCC), total column integrated liquid water (also referred to as liquid water path, LWP) and rainfall rates, which are compared to products derived from satellite observations. In the interests of brevity only mean fields are examined, and no critique is made of the quality of the retrieved observations, which are simplistically assumed to represent "truth".

After documenting the main features of the biases of the model climate, their evolution over the past 24 months is examined by comparing the current operational cycle to cycle 23r4, which became operational on the 12th June 2001. The latter cycle was chosen as a reference since it was one of the principal releases of the ECMWF model, forming both the atmospheric component of the currently operational seasonal forecasting system version 2, and was additionally used for the recently completed ERA40 reanalysis project.

Since part of the emphasis of this SAC report has been on the impact of horizontal resolution on biases, a range of climate experiments are conducted for a range of horizontal resolutions, and the changes in biases are compared to those obtained from improvements in physical parameterizations.

3.3.1 Past Studies

There have been a number of previous detailed studies that have examined the cloud climatology of the ECMWF model. It is useful to summarize these briefly. In the context of the forecast model, a recent summary by Lalaurette et al. (2003) shows a clear decline in European cloud cover biases after the implementation of the Tiedtke (1993) scheme in 1995. All of the following literature refers to the IFS model using this scheme, and most include the addition of ice sedimentation processes described by Gregory et al. (2000) and Jakob (2002)

The studies use a variety of observational sources. International Satellite Cloud Climatology Project (ISCCP) cloud cover observations were used to validate the cloud cover in the ERA15 reanalysis by Jakob (1999), who identified a lack of summer-time cloud cover over Eurasia, extratropical oceans and also in stratocumulus regions, underestimated in all three cases in the range of 10 to 15%, while cloud cover was too high in trade cumulus areas by about the same amount. The lack of stratocumulus is a common deficiency of many models (e.g. Teixeira and Hogan, 2002; Tompkins, 2002). Tselioudis and Jakob (2002) also used ISSCP data, but separated the statistics into different dynamical regimes for midlatitude clouds. They found a general overestimation of cloud optical depth, and an associated underestimation in cloud cover, particularly in subsidence regimes. High clouds were too thick while the cloud top pressure was too low. Low cloud in subsidence regimes had too small a cloud cover, and was optically too thick. Webb et al. (2001) also used ERBE and ISCCP data to highlight a lack of mid-level clouds in the Tropics (using model version 16r2).

In a series of studies, Chevallier et al. (2001), Chevallier and Kelly (2002), and Chevallier and Bauer (2003) used model fields to simulate radiances observed by various satellite platforms, and concluded that generally cloud systems were correctly located in the short term forecasts. However, they highlighted an overestimation of cloud liquid water, especially in the Tropics, and an underestimation of ice, prominently in midlatitudes. They again emphasized the lack of stratocumulus.

The ability of the short-term forecasts to correctly locate clouds was also identified by Mace et al. (1998); Klein and Jakob (1999); Miller et al. (1999); Hogan et al. (2001) and Morcrette (2002) using both space-borne lidar, ISCCP clouds and ground based radar observations. This is reassuring given that cloud in midlatitude winter is strongly dynamically forced, and the dynamical fields are well represented by the analysis system. Once again, Klein and Jakob (1999) pointed out the tendency of the ECMWF model to overestimate the optical depth of low clouds in midlatitudes, while ice clouds appeared to be underrepresented.

To summarize the conclusions of the above investigations, the clouds in the ECMWF model are assessed to be well captured in the short term forecast, but the following systematic biases were identified:

- The liquid water is too high, especially in subtropics;

- the ice amount is too low, especially in midlatitudes;

- cloud cover is too low in midlatitudes and subsidence regions, by 10 to 15%;

- there is a lack of stratocumulus clouds, with cloud cover too low by 10 to 20%;

- there is too much high cloud associated with tropical deep convection;

- and too little cloud over Eurasia in summer

It should be emphasized that the above studies almost exclusively examined short-range model forecasts (up to 2 days). One aim of this study is to examine how these biases are reflected in the atmospheric model's climatology (i.e., after 1 month into the integration).

3.3.2 Cycle 23r4

The catalog of statistics described above are examined in turn for the model cycle 23r4, with particular emphasis on model biases that are associated with either geographical location or regime type.

The first field to be examined is the map of total cloud cover (Fig. 12), which reveals a number of obvious problematic features. Perhaps the most glaring defect of the cloud simulation for this period is the almost complete lack of stratocumulus clouds off the western coasts of the Americas and Africa. Compared to the ISSCP C2 data, the cloud cover error exceeds 50% in places, and is over 30% for a significant region. It is
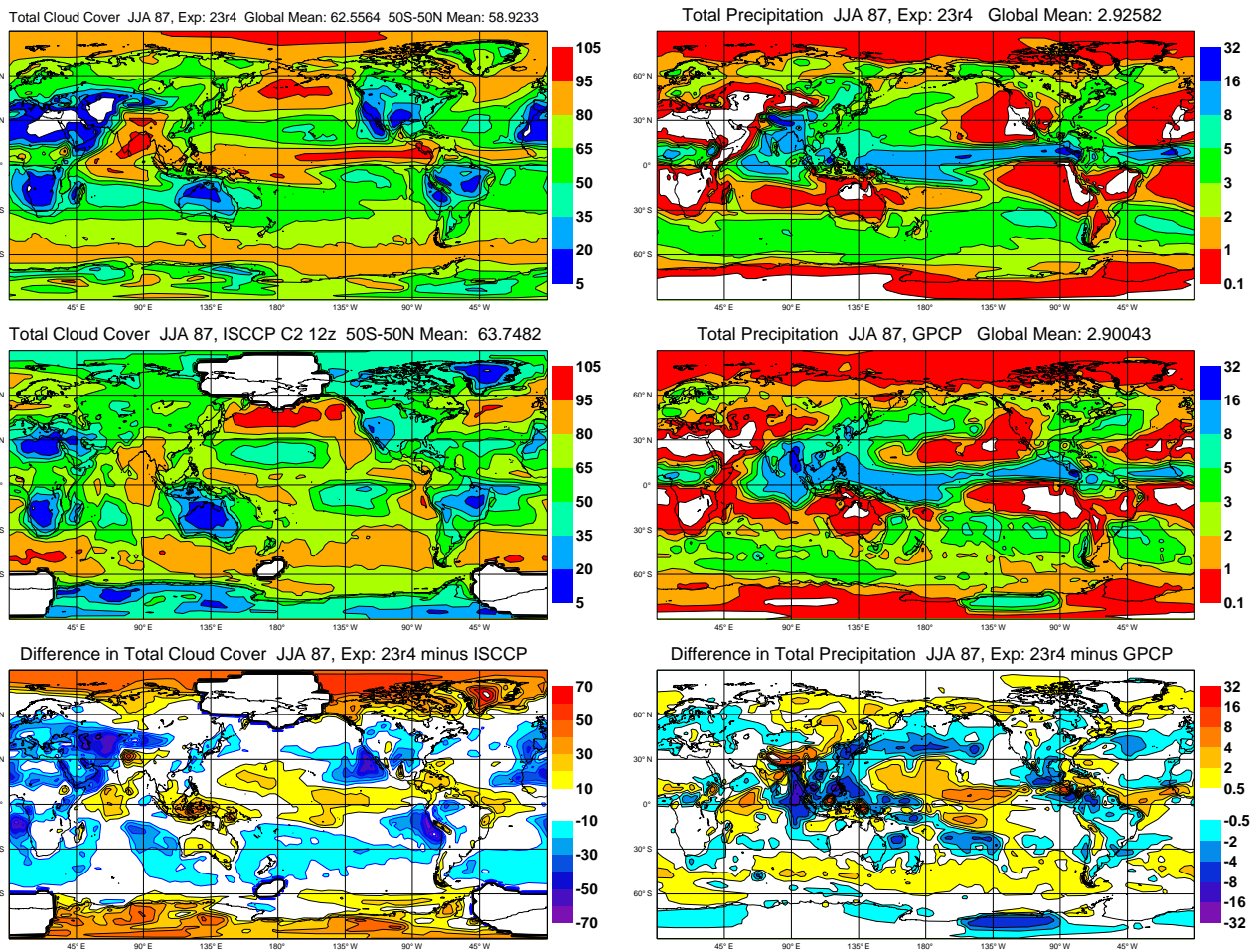
Figure 12: Left: Mean total cloud cover (%) for the period June-August 1987 based on model version 23r4 (upper panel) and ISCCP data (middle panel). Also shown is the difference between the model and observational estimates (lower panel). Right: As in the left column but for total precipitation (mm day^{-1}). Observational precipitation estimates are based on the GPCP climatology. Model results are based on 6 ensemble members.

clear therefore that the climate of the model performs significantly worse than the ERA40 reanalysis, which had errors of around 15% in these areas (Jakob, 1999). The second notable regime is the equatorial belt, associated with the Intertropical Convergence Zone (ITCZ). This cloud results from deep convective high level detrainment. The tendency for deep convection to remain fixed in location over the warmest SSTs has the consequence of producing too much high cloud in these regions. In contrast, most of the midlatitude winter hemisphere suffers from too little cloud, as does the Eurasian Continent. The errors in these regions, ranging from 10 to 20%, agree with the radar investigation of Hogan et al. (2001) which cited similar shortfalls in high cloud amounts at a UK site. All these systematic biases in cloud cover were identified in the previous studies discussed above, indicating that the asymptotic biases also affect the short-term forecast performance.

Fig. 12 also compares the model precipitation fields to the Global Precipitation Climatology Project (GPCP) climatology (Arkin and Xie, 1994). Despite the three month averaging period the fields are noisy as expected. However, it is clear that compared to the precipitation retrievals the model appears to over-estimate the ITCZ precipitation in the deep convective zones, situated over the mid Pacific due to the developing El-Nino conditions. This appears also to be the case for the midlatitude zone in the Southern Hemisphere, although the

accuracy of the retrieved product is questionable over the oceans.

Retrievals for August 1987[4] from SSM/I are compared to the ensemble mean model for the total column water vapour and liquid water path (Fig. 13). The model mean total water vapour agrees very well with the observations, but a number of regional differences are apparent. The Tropical Western Pacific appears too dry, as does the stratocumulus regions off the West American coasts. The Pacific sub-tropical regions on the other hand are too moist. There is also more distinction between the main and southern branch of the ITCZ in the Western Pacific; a historical characteristic of the ECMWF model. It is notable that many of these systematic errors are co-located with biases in the LWP. For example, the moist anomaly in the subtropical Pacific coincides with a strong positive bias in LWP. In fact, in agreement with previous findings, this LWP bias is widespread through most of the subtropical ocean regions usually subject to shallow cumulus convection. Likewise the LWP is overestimated in the deep convective regions, while it is severely underestimated in the stratocumulus regions. Overall, these biases effectively cancel, and the mean model LWP over the ocean points for which SSM/I data is available is only 7 g m^{-2} too high. The mean bias should be treated with caution due to inaccuracy of the observational product. However, it is possible to state with some confidence that the model fails to represent the contrast between the trade and stratocumulus regimes. Moreover, it is reassuring that the LWP biases with respect to SSM/I information are broadly consistent with the cloud cover comparison with ISCCP data, both in terms of geographical location and the sign of the biases.

The TOA shortwave and infra-red fluxes are now compared to observations for the summer months (Fig. 14). The infra-red fluxes indicate many of the biases apparent in the cloud cover comparison to ISCCP. The Tropics are too opaque, consistent with the excessive cloud cover in these regions. Chevallier and Morcrette (2000) also found similar biases in their comparison with CERES-TRMM data. Likewise the Eurasian Continent and the Southern hemisphere midlatitude regions emit more than indicated in the ERBE observations, consistent with the underestimated cloud cover, and also the conclusions of Chevallier and Kelly (2002) that the ice content in these clouds was too low.

Generally, the match of the shortwave radiation to observations appears worse than for the infra-red. In particular, the excess liquid water in the Trades significantly enhances the TOA albedo, while obviously the lack of cloud in stratocumulus areas has the opposite effect. The outgoing shortwave radiation is too low over Saharan Africa, which could be caused by a variety of reasons, including too low a surface albedo, or a lack of regional aerosols. It is unlikely that the annual-mean aerosol climatology can capture the high temporal variability of aerosol loading in the region (Pandithurai et al., 2001).

If the model would have a tendency to produce incorrect surface winds, this could lead to biases in cloud regimes via the link with surface fluxes. In the context of seasonal forecasting, the correct prediction of near surface winds along the equatorial belt is one of the key ocean forcing factors that is imperative to predict correctly. To examine this, Fig. 15 compares the ensemble mean 10m wind speed for July and August 1987 to the product retrieved from SSM/I. The correspondence is striking, both in terms of the mean (7.3 compared to 7.2 m s^{-1}) and regional variations, especially when the low resolution of the model is considered. The relative error is shown, since small absolute errors can still have a significant impact in low wind conditions. Indeed, this reveals that the area of the Tropical Western Pacific which suffered from dry biases and low cloud amount, also has under-predicted surface winds, despite the low wind conditions prevailing in the observations. It should be noted, however, that the effective wind speed used for surface fluxes is enhanced in these regions by the implementation of a subgrid gustiness parameterization, which dramatically improved the model climate in this part of the Tropics (Miller et al., 1992).

Apart from a slight over-prediction of surface winds in limited areas of the southern oceans and the subtropical

---

[4]August 1987 is the first month in ERA40 for which SSM/I data are assimilated (Chevallier et al., 2001). This version of the retrieval product is available from July 1987, one month after launch.
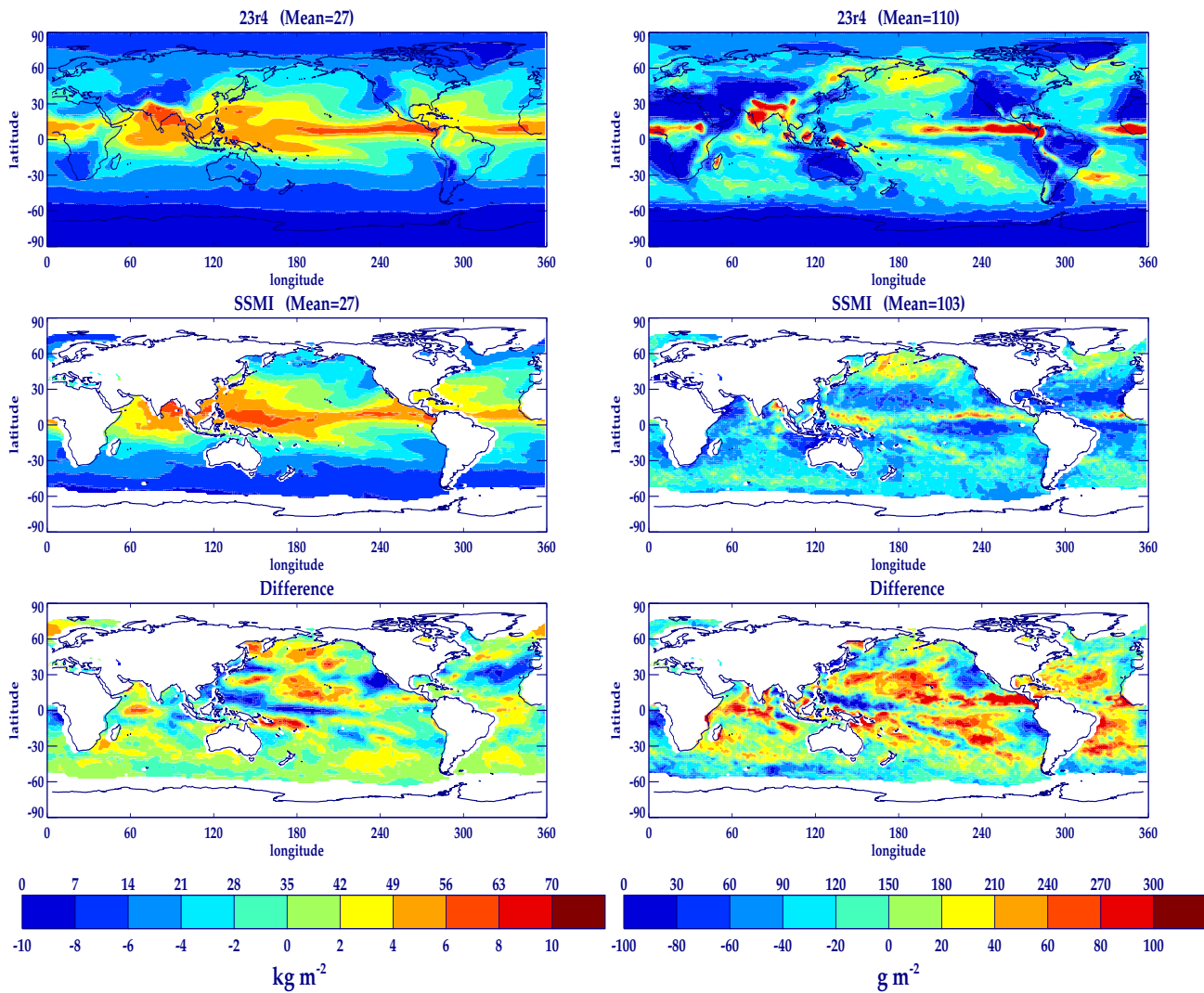
Figure 13: Left: Mean total column water vapour (kg m^{-2}) for August 1987 based on model version 23r4 (upper panel) and SSM/I estimates (middle panel). Also shown is the difference between the model and observational estimates (lower panel). Right: Same as in the left column, except for cloud liquid water path (g m^{-2}).

Atlantic, the only other region with significant biases are the stratocumulus regions, although it is unclear how much of this stems from the difficulty of retrieving information in these virtually overcast areas.

Finally, the corresponding experiment conducted for the winter season of 1987/1988 is briefly discussed. The majority of the main biases highlighted above were also seen in the winter season, and thus for reasons of brevity, only the TOA fluxes are shown for illustrative purposes in Fig 16. These again show the high opacity of deep convective clouds, the lack of stratocumulus, and the under-prediction of midlatitude ice clouds.

To summarize, it is apparent from this cursory comparison with satellite retrieved products that virtually all cloud related biases that have been identified in the short-term forecasts by previous investigations are reflected in the model climatology at T95 resolution. These biases, in particular the lack of stratocumulus, the deep convective clouds that are too thick, and the over-prediction of liquid water in the shallow cumulus regimes, are present in both the summer and winter integrations.
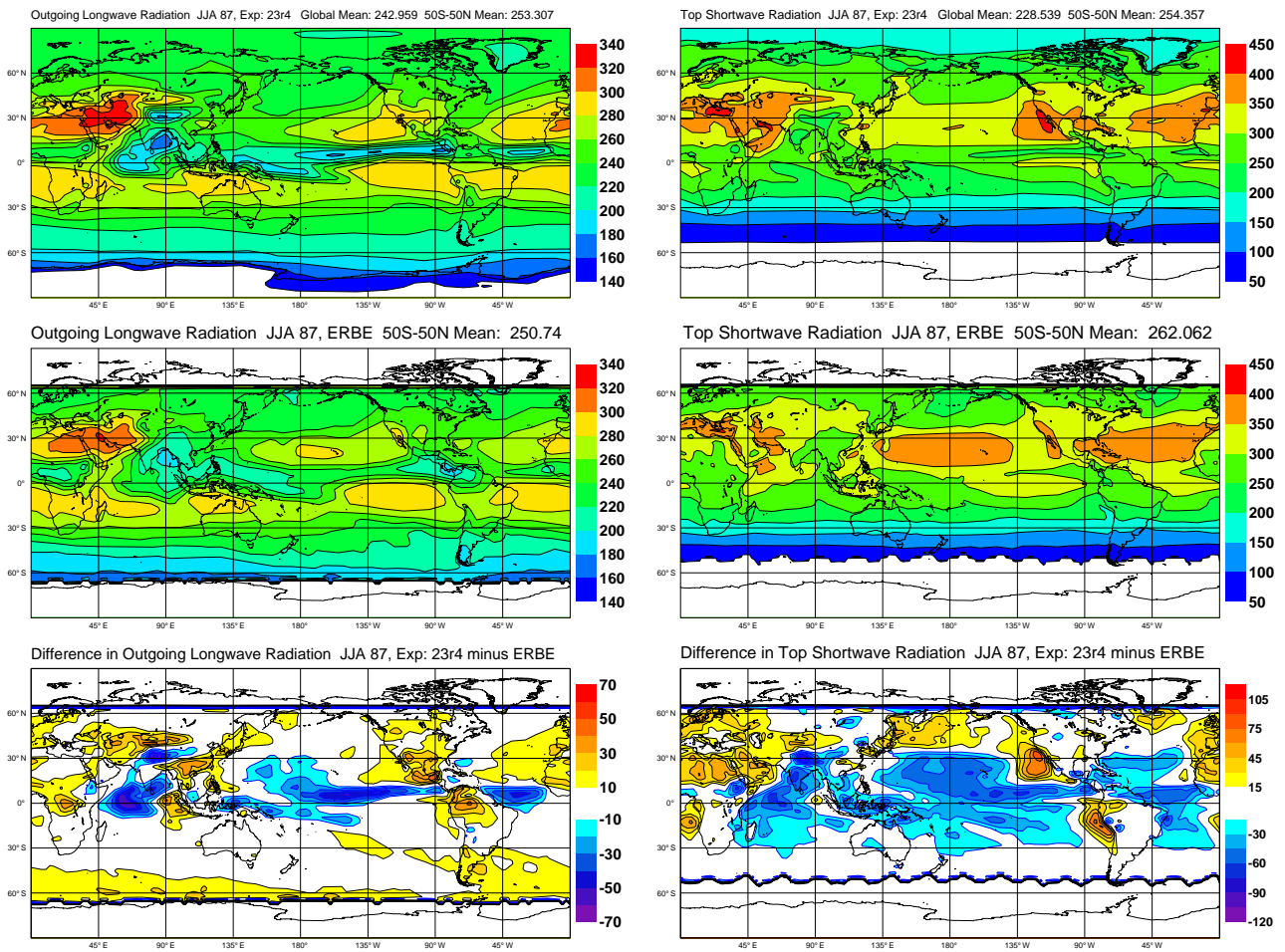
Figure 14: Mean top of the atmosphere (TOA) net fluxes (W m^{-2}) for the period June–August 1987 based on model version 23r4 (upper panels) and ERBE estimates (middle panels) along with differences between the model and observational data (lower panels). Left: Outgoing longwave radiation. Right: Shortwave radiation (positive values indicating a net influx).

3.3.3 Improvement with the latest cycle

During the 24 months that have passed since the ERA40 cycle 23r4 was operational, a considerable number of modifications have been made to ECMWF IFS system, not least to the model physics. This section examines how these physics changes have modified the model climate and its cloud associated biases. The inextricable link between the processes of convection, cloud and radiation render it impractical to assign changes in systematic biases to any particular modification without conducting further sensitivity experiments. Since this is beyond the scope of this report, the modifications are considered together as a complete package.

The modifications made between cycles 23r4 and 26r1 can be summarized as follows:

Convection

The convection scheme has undergone a number of modifications. The first pertains to the microphysical treatment of liquid water in the updraughts. In the ERA40 cycle model, precipitation formation was specifically inhibited in the lowest 150hPa above cloud-base over oceans and 300hPa over land. This inhibition was deemed no longer to be necessary due to the inclusion of a more complex microphysical scheme based on
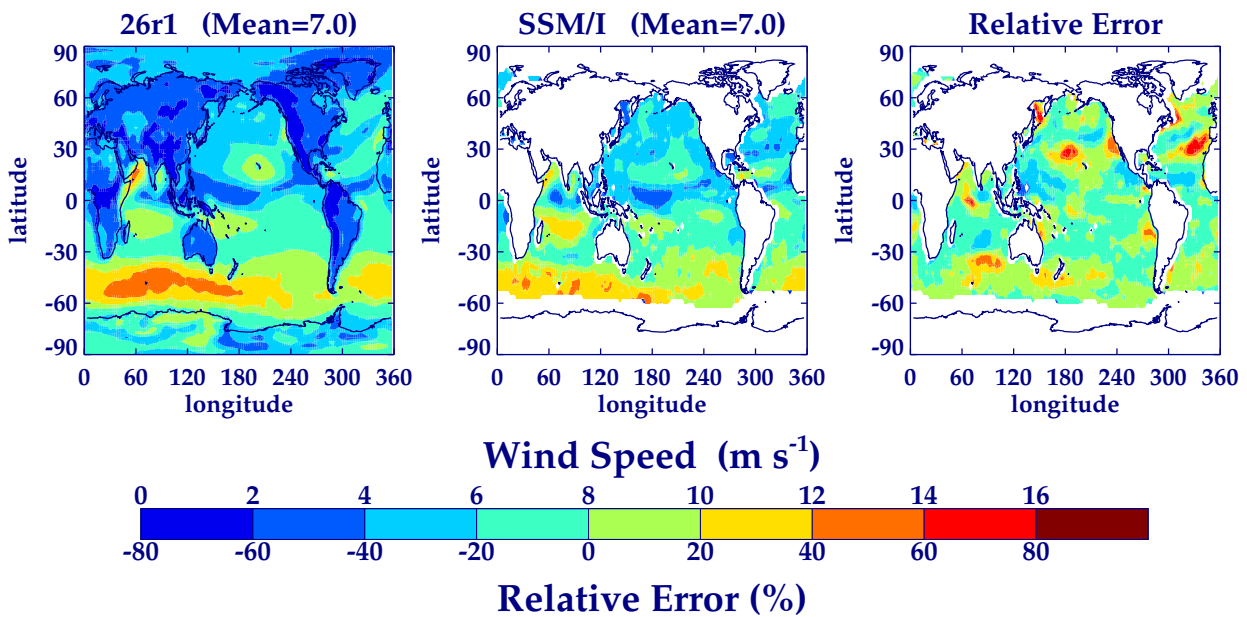
Figure 15: Model version 23r4 mean 10m wind speed for July/August 1987 compared to SSM/I product

Sundqvist et al. (1989) and was therefore removed.

The initiation of the convection scheme (switching) was also modified. In the original scheme, a negative buoyancy of 0.5K at cloud base for an undiluted parcel ascent was used as a threshold. The undiluted parcel ascent was also used for selecting whether shallow or deep convection would take place. Cloud base vertical velocity was fixed to 1m s^{-1} for the parcel.

The modified scheme introduces an entraining plume assumption for the parcel ascent and subsequent convective decision. The parcel is initialized with thermodynamic excess properties and a vertical velocity according to surface layer similarity. Surface driven convection (shallow and deep) now switches on if heat fluxes and cloud base parcel vertical velocity are upward. Cloud base vertical velocity in the final ascent is set consistently with the sub-cloud entraining plume. The discrete cloud base is now set to the closest model level to the analytic cloud base instead of the level below cloud base. The entraining plume formulation used in the cloud type decision process is specified to closely resemble the full cumulus ascent calculation.

In cases of deep convection, instead of simply using parcel properties originating from the lowest model level, all model levels below 700 hPa are now tested for instability. Additionally the parcels used in the ascent are now mixed layer mean values, rather than single layer values. Finally, further adjustments were made to increase the entrainment into the updraught parcels, and also the precipitation efficiency, which had beneficial effects on upper-tropospheric winds.

Clouds

The cloud scheme based on Tiedtke (1993) was also extensively modified. The principle aim was to upgrade the numerical implementation of the scheme, such that all the cloud related processes (sources and sinks) are assumed to occur simultaneously. In the original scheme used in version 23r4, the processes were instead treated sequentially, and as such, precipitation formation assumed the lowest priority. In other words, if a process such as evaporation would remove all liquid water over a timestep, no precipitation could form. The opportunity was taken to ensure that the new solution conserved total water and moist static energy.

In addition a number of physics changes were made. This includes the treatment of ice sedimentation, identified
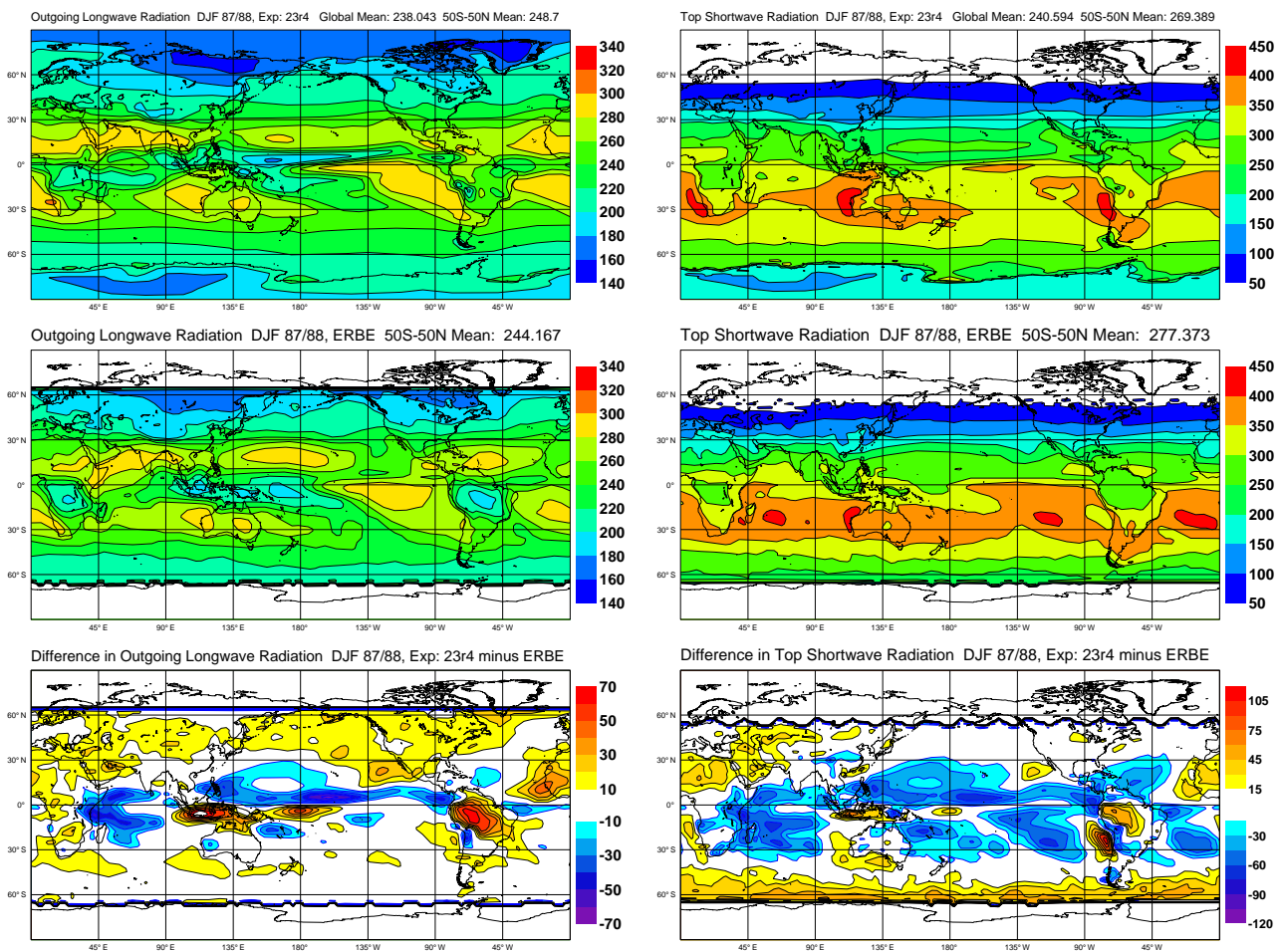
Figure 16: As Fig. 14, except for months of December to February of winter 87/88.

by Jakob (2002) as a key parameterization. The previous implementation restricted the large ice particles to fall at a speed related to the vertical resolution and timestep for stability purposes. However, the stable implicit numerical solution did not require such a restriction, which was also not imposed on the small ice particles. The net result was that large ice particles effectively fell slower than their smaller counterparts, and the net mean fall speed for all ice was consequently almost independent of ice mass mixing ratio q_i, in stark contrast to observations (Heymsfield and Donner, 1990). This treatment was modified to assume that large ice particles fall according to Heymsfield and Donner (1990) as $V = Kq_i^{0.16}$, where the constant K assumes a reduced value of 2.3. Small ice particles are given a residual non-zero fall-speed of 0.15 m s^{-1}.

Other cloud physics changes included the cloud top entrainment process which was reformulated in terms of total water and dry static energy, and solved implicitly. The cumulus subsidence-induced cloud water advection process was also implicitly solved, along with the cloud erosion by turbulent mixing term. Finally, changes were made in the way supersaturation was handled, which is now calculated at semi-Lagrangian arrival point profiles.

Radiation

The short-wave radiation scheme was modified to increase the spectral resolution. The model used in 23r4 used four spectral intervals with transmission functions based on statistical models. This was replaced by a

six band code using transmission functions derived from a line-by-line code (Dubuisson et al., 1996). Part of the requirement for a new scheme stemmed from an earlier change in model vertical resolution, which raised the model top from 10 to 0.2 hPa, and improvements in the surface scheme. The Rayleigh scattering optical thickness is now solar-zenith angle dependent and all optical properties have been recomputed for the six new bands.

Results

The two model fields that clearly illustrated the salient features and problems with systematic biases at 23r4 were the total cloud cover and LWP. Fig. 17 reproduces the comparison to ISCCP and SSM/I for model cycle 26r1. The clear improvement is in the Tropics where there is a distinct reduction in cloud cover. There is also a small improvement in the stratocumulus regime off the West coast of north America, although this cloud type is still poorly represented. The general under-prediction of cloud cover in the midlatitude regions is still present, and the cloud cover in the subsiding branches of the Hadley circulation appears to have deteriorated.

The LWP shows a notable improvement over these model cycles. Firstly, it is obvious that the reduction in cloud cover in the deep convective regimes is accompanied by a dramatic drop in liquid water, to bring the model much closer to observations. Additionally it is clear that the widespread overestimation in the subtropical trade and shallow cumulus regimes has been almost completely eradicated, both over the Pacific and Indian Oceans. This, in addition to the continued lack of stratocumulus clouds implies that there is no longer a strong cancellation of errors, and the mean bias increases from $+7 \text{ g m}^{-2}$ to -15 g m^{-2}. However, later it will be shown that the improvement is reflected in a reduction in spatial root mean square (RMS) error of LWP.

Do these improvements in cloud properties reflect in the radiative biases? Fig. 18 shows that the improved cloud properties, in conjunction with the new radiation code, do indeed reduce radiative biases considerably. The reduction of cloud liquid water and cover in deep convective regions improves both the SW and IR biases there. The continuing lack of ice water in midlatitudes means that the IR bias is retained. The SW field is dramatically improved throughout the trade regions due to the reduction in LWP and the direct modifications to the radiation scheme itself.

To summarize the changes between cycle 23r4 and 26r1, Tab. 1 gives the mean bias and spatial RMS error for these main quantities of cloud cover, LWP, and TOA radiative fluxes for the three member ensembles. Two calculations are shown for the RMS error: the first is the mean of the spatial RMS error for each ensemble member, with the standard deviation providing an estimate of the inter-ensemble variability, while the second calculation is for the RMS error between the ensemble mean field and the respective observation. For most smooth fields these figures are very similar, but for spatially variable fields such as the LWP, there is a considerable difference. The standard deviation of the mean biases is also given. It is apparent that the new cycle improves on all quantities for both seasons, with the exception of the TOA IR fluxes, which is neutral.

To set this improvement in a longer-term context, Tab. 2 shows the evolution of cloud cover, LWP, and TOA net flux RMS errors for main model cycles over the past 7 years. This comparison is conducted at a lower horizontal resolution of T63, and for one summer integration only. However, the differences between model versions generally exceed the inter-ensemble variability stated in Tab. 1. The table indicates that at least in terms of these variables, there has not been a monotonic improvement during this period. For example, in terms of the cloud liquid water climatology, the RMS error indicates that model cycle 23r4 was further from observations than any previous cycle for August 1987, while the latest model cycle gives by far the best performance. For the other variables the impact appears neutral, although cycle 26r1 does also improve the TOA SW fluxes.

The analysis of RMS errors is repeated for the winter 1987/1988 to confirm these trends (with the exception of cycle 23r4 for which no T63 experiment exists). For this season, cycle 26r1 not only improves the LWP representation, but also gives the best performance for the TOA net fluxes in both infra-red and short-wave.
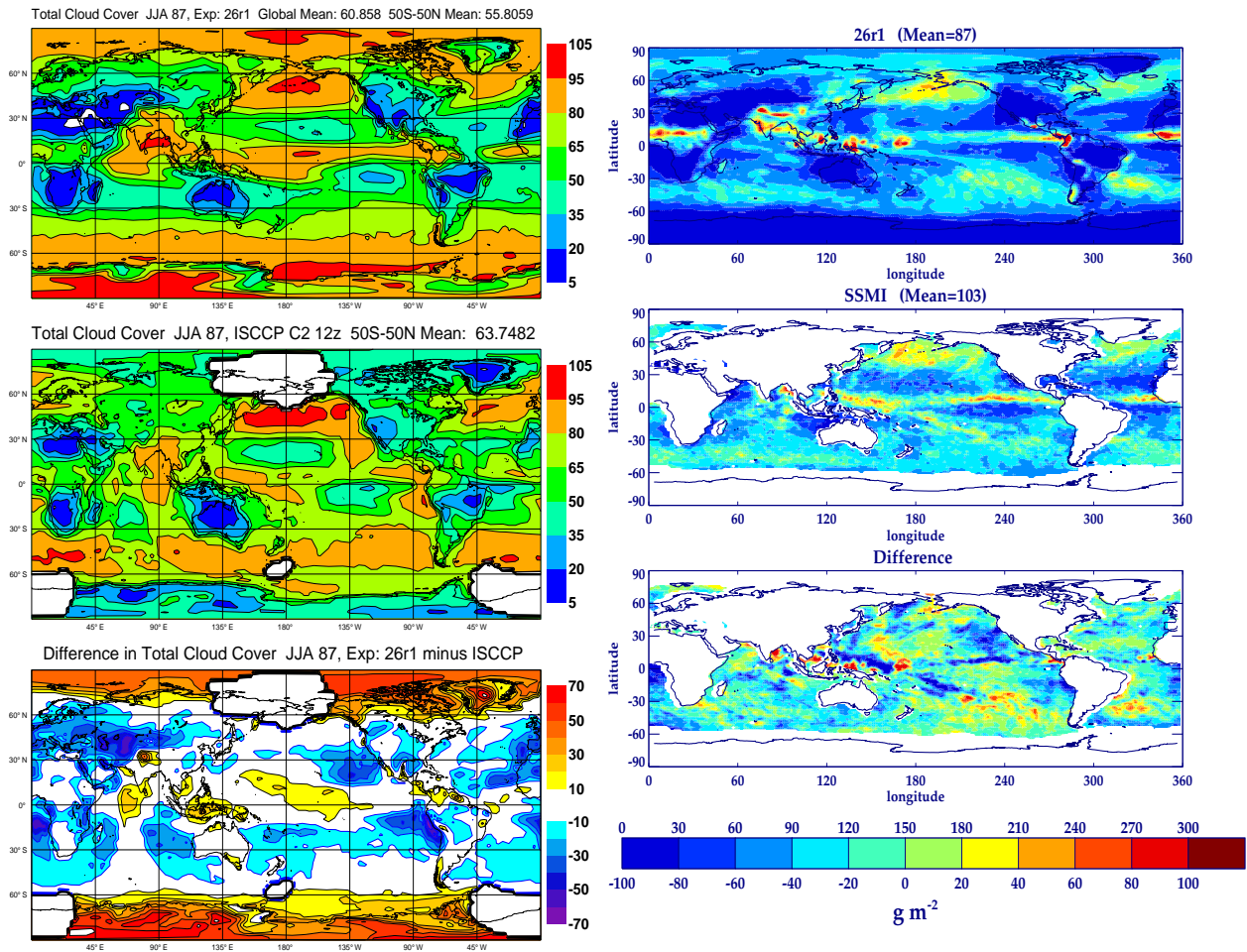
Figure 17: Left: Mean cloud cover (%) for June–August 1987 for model cycle 26r1 (upper panel), ISCCP observations (middle panel), and the difference between model and observations (lower panel). Right: As in the left column, except for cloud liquid water path (kg m^{-2}). Observational LWP estimated are based on SSM/I retrievals.

It should be pointed out that a steady improvement in the long-term climatology of these moist-process related quantities should not be taken for granted. Unlike climate models, forecast models are traditionally judged on their ability to reproduce observed dynamical and thermodynamical (specifically temperature related) fields in short-term high resolution forecasts. The fact that the climatology of these quasi-independent metrics is superior in the latest model version is therefore gratifying.

Finally, we consider possible problems of the latest ECMWF model (cy26r1) in simulating precipitation. Results for wintertime (left) and summertime (right) precipitation are shown in Fig. 19 for the period 1988–99. The verification dataset is the global precipitation climatology of Xie and Arkin (1997). The most notable systematic differences between model precipitation and the observations is found in the Tropics and in the Northern Hemisphere extratropics. Compared to the Xie and Arkin (1997) climatology, the latest ECMWF model underestimates (overestimates) wintertime precipitation south (north) of the Equator in the western tropical Pacific. Moreover, the model tends to underestimate precipitation over the Maritime Continent by about 30%. During boreal winter the model shows higher values of precipitation than the Xie and Arkin (1997) climatology over midlatitude North Pacific and North Atlantic ocean. This must not necessarily be due to model deficits, how-
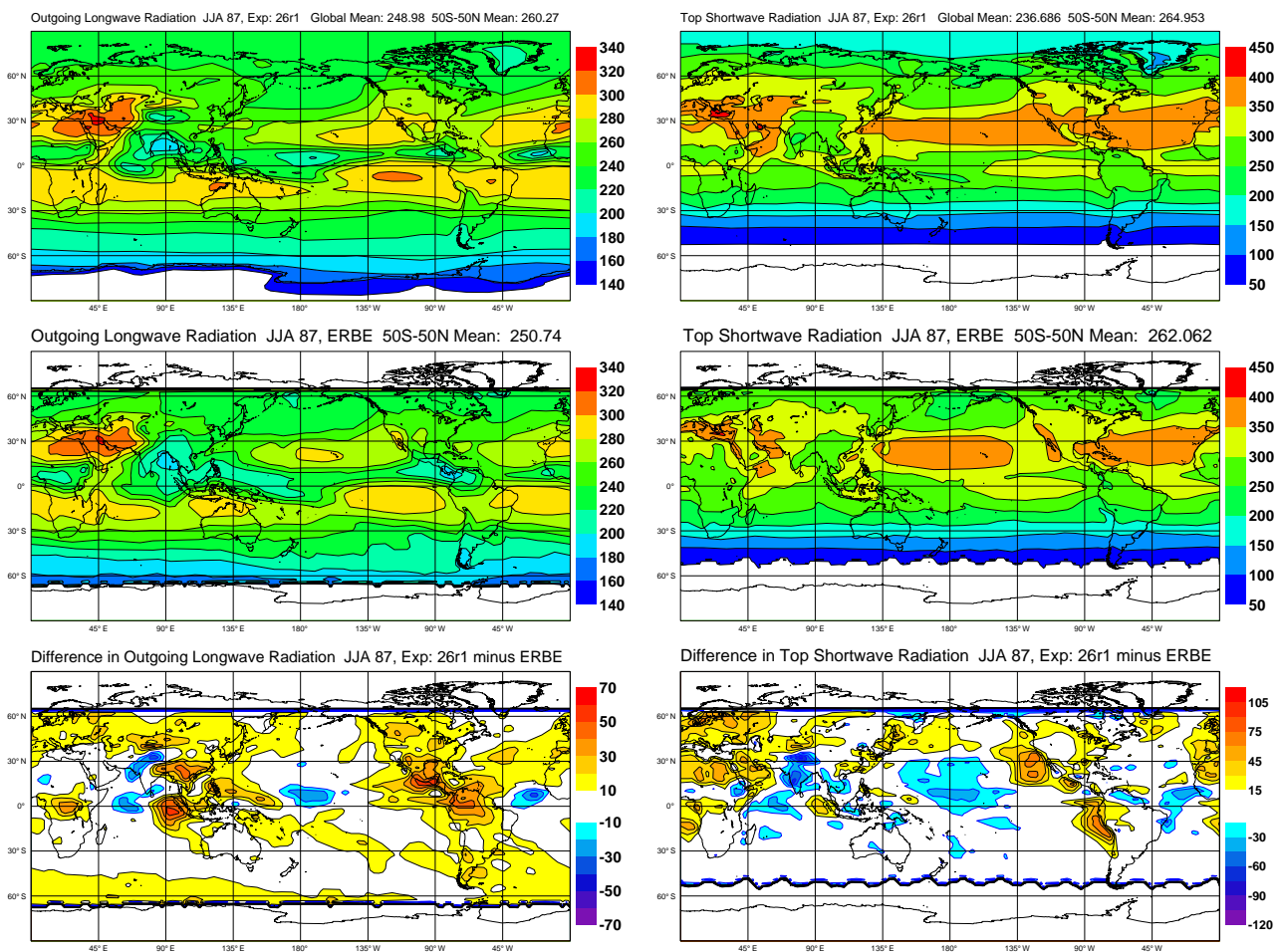
Figure 18: As in Fig. 14, except for model cycle 26r1.

ever, given that the Xie and Arkin (1997) climatology is likely to have a low precipitation bias in these region (Adler et al., 2001).

The most pronounced summertime precipitation differences between the latest model cycle and the Xie and Arkin (1997) climatology is associated with the tendency of the model to overestimate the splitting of the ITCZ over the central tropical Pacific, and with too low values of precipitation over Indonesia and the eastern tropical Pacific. A similar assessment of simulated global precipitation fields has been carried out by Brankovic et al. (2002) for model cycle 18R6 (operational from June 1998 to March 1999) with 60 levels in the vertical. The difference they found during wintertime and summertime are virtually the same as those found for the latest model cycle 26r1. This suggests that model improvements during the last three or four years had a rather small influence on the characteristics of simulated precipitation.

3.3.4 Influence of horizontal resolution

The last section shows how modifications to physics have systematically improved various aspects of the model climatology over the past 7 years, and particularly the last two. Over this period of time, the expanding computational power has been used to introduce a finer grid in the operational forecast model; from 31 to 60 vertical levels, and from T213 to $T_L 511$ horizontal resolution. The improvement in climate by the modified physics

Table 1: RMS errors of Cloud cover, LWP, and TOA fluxes for model cycle 23r4 and 26r1. Results are for an ensemble of six experiments at T_L95L60 resolution. For each field ϕ, the spatial RMS error is calculated for the ensemble mean $RMS(\overline{\phi})$ is shown. Additionally the mean of the separately calculated bias and RMS error is given, with the standard deviation providing an estimate of the inter-ensemble variability. With the exception of LWP, all comparisons are for June–Aug 1987 for the summer case, and Dec-Feb 1987/1988 for the Winter case, in the region $-50^o <lat<50^o$. Instead, LWP is for all ocean points, and for the summer case is only for the two months of July/August 1987 and the Winter only for January/February 1988.

Model Experiment Date	23r4 Summer 87	26r1 Summer 87	23r4 Winter 87	26r1 Winter 87
Cloud Cover (%)				
Mean Bias	-5.4 ± 0.2	-8.3 ± 0.4	-4.0 ± 0.2	-6.3 ± 0.4
$\overline{RMS}(\phi)$	17.4 ± 0.4	16.4 ± 0.3	16.0 ± 0.4	15.2 ± 0.2
$RMS(\overline{\phi})$	16.9	15.9	15.4	14.5
LWP (g m^{-2})	(July/Aug only)	(July/Aug only)	(Jan/Feb only)	(Jan/Feb only)
Mean Bias	9.6 ± 2.4	-14.0 ± 1.4	12.8 ± 0.8	-8.1 ± 1.2
$\overline{RMS}(\phi)$	48.2 ± 1.0	41.6 ± 1.0	51.4 ± 2.8	43.9 ± 3.4
$RMS(\overline{\phi})$	44.0	38.4	45.5	34.8
TOA IR Flux (W m^{-2})				
Mean Bias	2.9 ± 0.3	9.5 ± 0.4	4.9 ± 0.3	11.1 ± 0.4
$\overline{RMS}(\phi)$	14.7 ± 0.5	14.8 ± 0.4	16.8 ± 0.9	16.1 ± 0.3
$RMS(\overline{\phi})$	13.7	14.3	16.0	15.4
TOA SW Flux (W m^{-2})				
Mean Bias	-7.6 ± 0.5	2.6 ± 0.6	-7.3 ± 0.3	3.6 ± 0.4
$\overline{RMS}(\phi)$	26.8 ± 0.4	18.9 ± 0.4	26.4 ± 0.6	17.4 ± 0.3
$RMS(\overline{\phi})$	26.1	18.1	25.6	16.4

Table 2: RMS errors of Cloud cover, LWP, and TOA fluxes for various model cycles over the past 7 years. All experiments are conducted at T63 resolution for one experiment only. All comparisons are for June-Aug 1987, and $-50^o <lat<50^o$, except LWP, which is for Aug 1987 only. [1]Note: Model versions 15r7 and 18r4 use the 31 level vertical grid, rather than the 60 level current operational mesh.

Model Version Date	15r7[1] 12/96	18r4[1] 02/98	21r4 10/99	22r3 06/00	23r4 06/01	26r1 04/03
Cloud Cover (%)	17.1	16.4	16.1	15.9	16.0	17.5
LWP (g m^{-2})	54.5	51.5	53.6	52.5	63.3	46.7
TOA IR Flux (W m^{-2})	18.9	12.8	14.6	14.4	14.4	16.5
TOA SW Flux (W m^{-2})	27.5	28.3	29.6	29.6	29.5	25.1

Table 3: As Tab. 2 but for winter 1987 experiments. Note, that no T63 integration is available for model version 23r4. LWP RMS errors are for February 1988

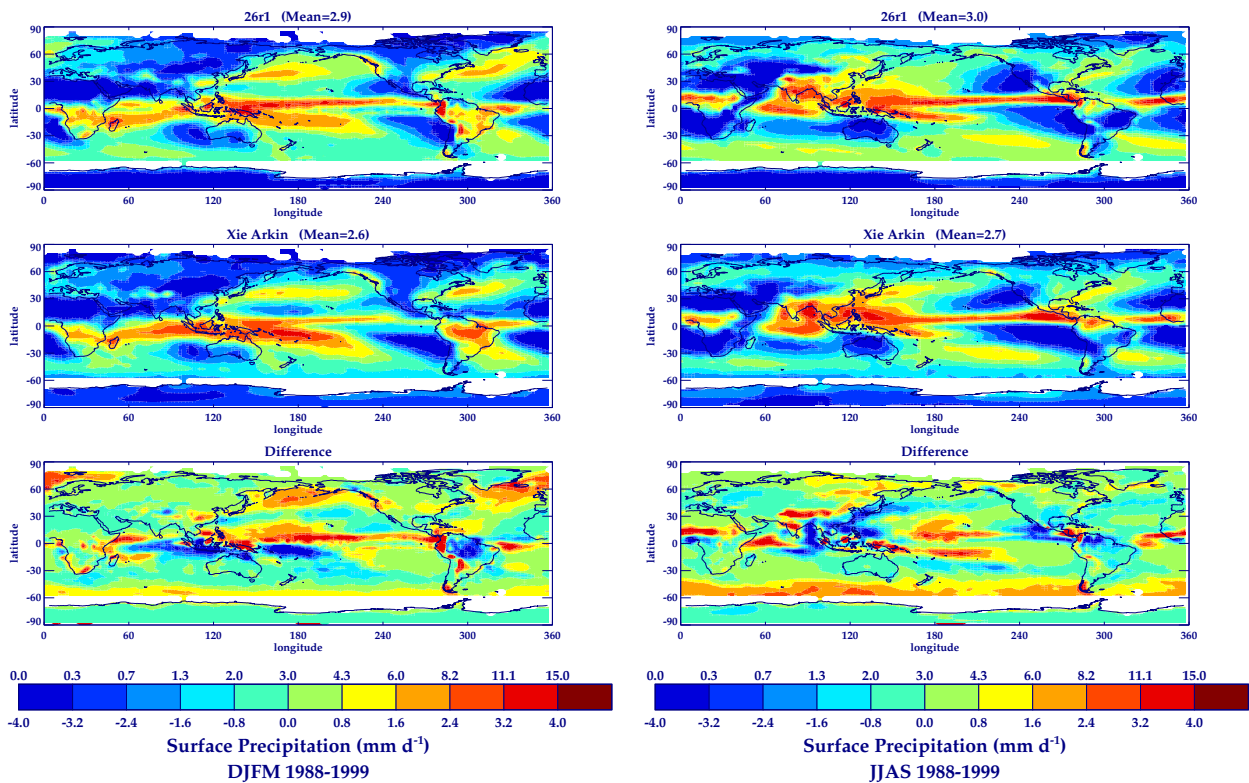| Model Version Date	15r7[1] 12/96	18r4[1] 02/98	21r4 10/99	22r3 06/00	26r1 04/03
Cloud Cover (%)	17.2	15.3	16.1	15.5	16.5
LWP (g m^{-2})	54.6	48.8	55.5	53.0	42.4
TOA IR Flux (W m^{-2})	21.0	16.8	18.3	16.6	15.2
TOA SW Flux (W m^{-2})	24.5	25.5	25.9	26.8	19.1

Figure 19: Average precipitation fields (mm day^{-1}) for seasonal runs of model cycle 26r1 (upper panels) and for observational estimates based on the climatology of Xie and Arkin (1997) (middle panels). Also shown are asymptotic systematic precipitation errors (model minus observations, lower panels). Results are based on winter (left column) and summer (right column) averages for the period 1988–99.

parameterizations is compared to changes obtained from horizontal resolution increases. Using model version 26r1, four ensembles of three members each are conducted at T63, T_L95, T_L159 and T_L255 resolutions, using 60 levels for the summer case only.

The total cloud cover is depicted in Fig. 20, along with the error with respect to ISCCP observations. It is apparent that the lowest resolution simulation at T63 suffers from increased biases in the Tropics. The increasing resolution also aids slightly the representation of the stratocumulus regions. Otherwise, there is no clear trend, which also seems to be the case for other model quantities which are examined in Tab. 4. In fact, it appears that there is a slight deterioration at the highest resolution of T_L255. It should be remembered that the model is used with its "default" settings at each resolution, so that the timestep is shorter at T_L255, and the lowest resolution of T63 uses a different grid. (Note that the convective CAPE closure adjustment timescale is constant throughout this resolution range at one hour.) Recent model cycles have mainly been developed using T_L511 forecasts and T_L95 climate runs, implying that the choice of model tunable parameters may be optimised to suit these two resolutions. Nevertheless, the insensitivity of the model climate to horizontal resolution confirms the findings elsewhere in this document, which show little impact of model resolution on many model diagnostics. The stability with regard to horizontal resolution may be regarded as a positive model attribute, since it implies that results of numerical investigations can often be safely generalized, without resorting to expensive replications for a range of resolutions.
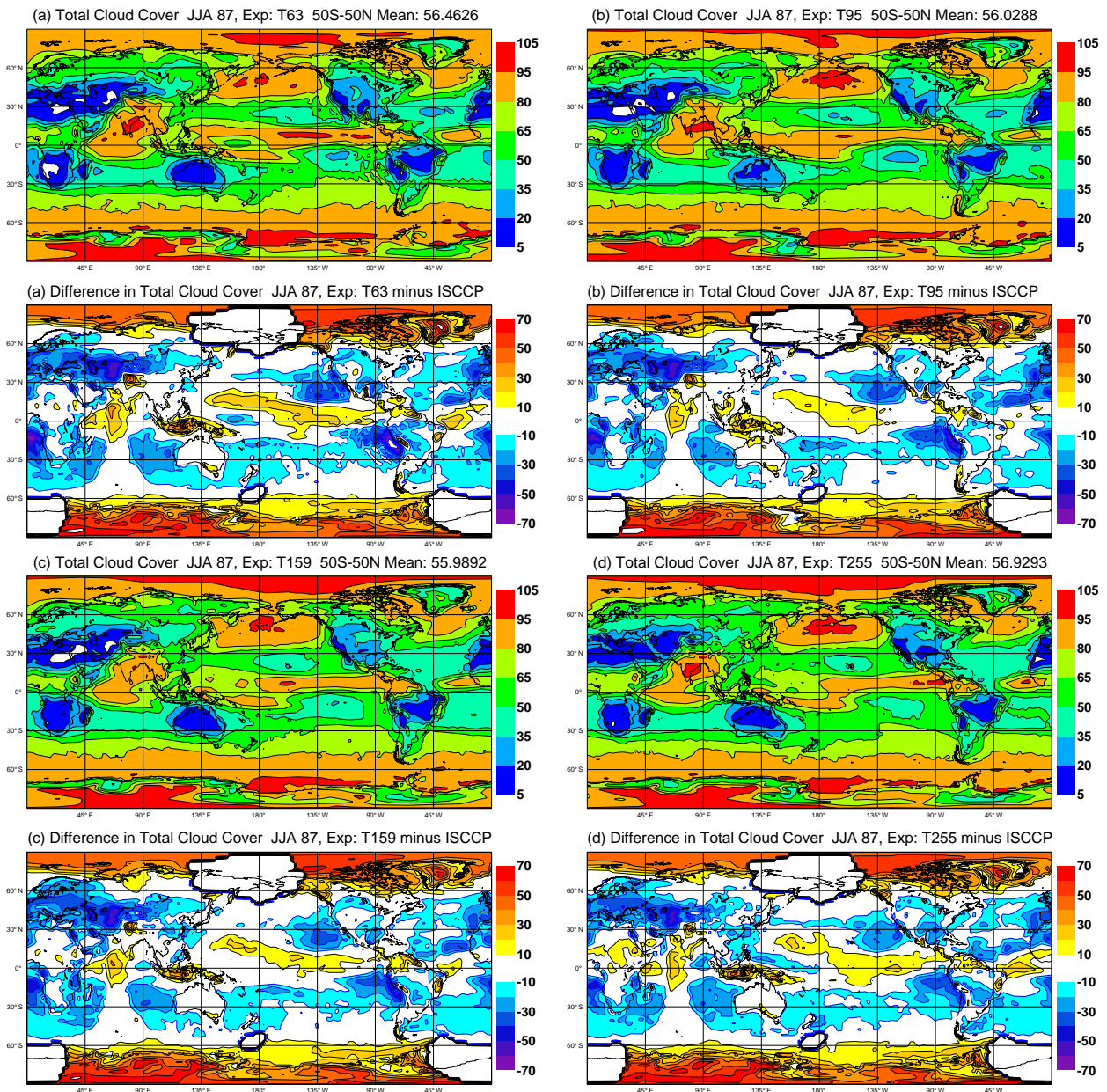
Figure 20: Mean total cloud cover (%) and cloud cover error (%) compared to ISCCP observations for June–August 1987 and model cycle 26r1 at (a) T63 (b) T_L95 (c) T_L159 and (d) T_L255 horizontal resolutions. The model results are based on 3 ensemble members.

3.4 Atmospheric Circulation

3.4.1 Systematic Errors

Asymptotic systematic wind errors from the seasonal run at 1000 hPa for winter (left column) and summer (right column) season are shown in Fig. 21 together with long-term mean winds from ERA-40. For the winter season the most notable feature is associated with very large systematic anticyclonic wind errors in the North Pacific. In some regions the wind errors are of the same size as the mean wind itself (e.g., close to Japan). During

Table 4: As Tab. 2 but for various model resolutions of cycle 26r1. Note that the RMS errors are for the ensemble mean fields, hence the generally lower values for the LWP results.

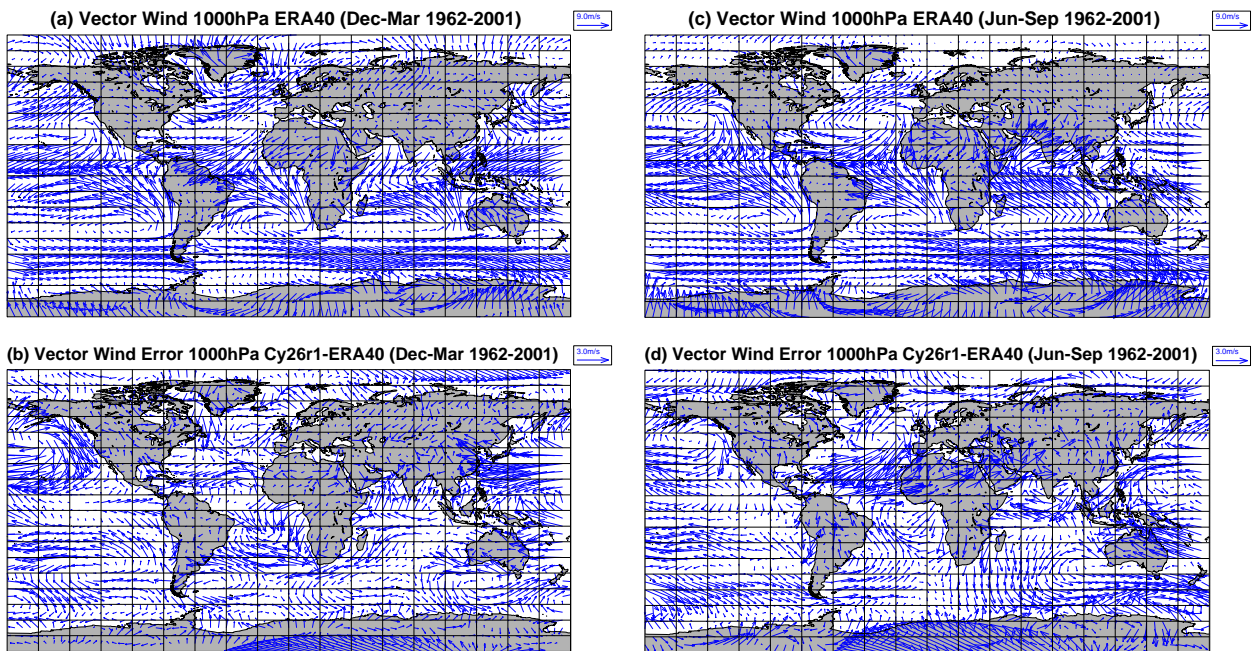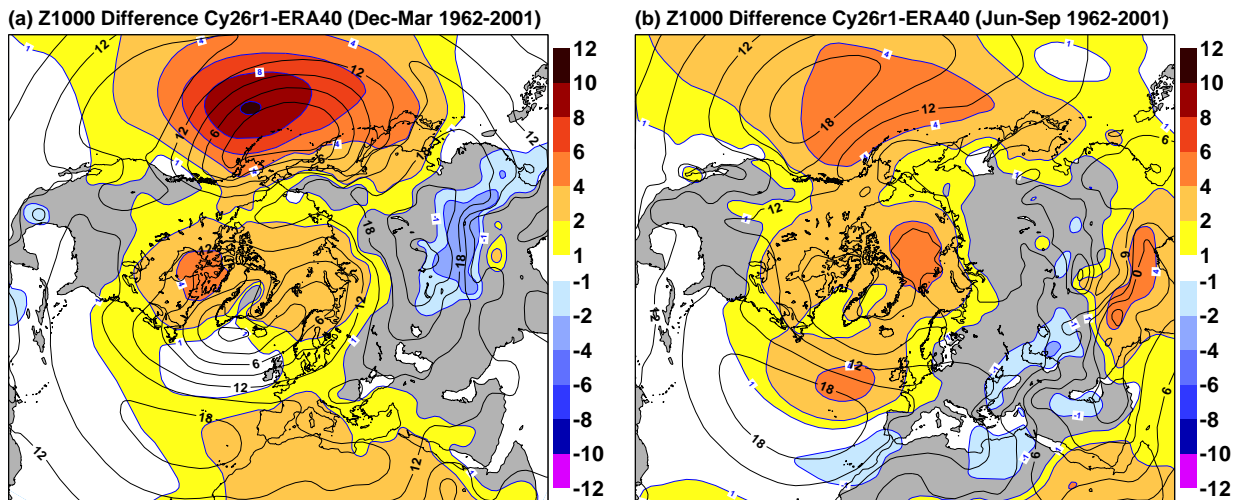| Resolution	T63	T_L95	T_L159	T_L255
Timestep (s)	3600	3600	3600	2700
Cloud Cover (%)	17.1	15.9	15.4	15.5
LWP (g m^{-2})	41.2	40.4	41.7	43.6
TOA IR Flux (W m^{-2})	15.7	14.3	14.9	16.5
TOA SW Flux (W m^{-2})	24.3	22.1	22.3	23.1

Figure 21: Vector winds at 1000 hPa from ERA-40 for (a) winter (Dec–Mar) and (c) summer (Jun-Sep) along with the difference of mean vector winds between the ECMWF model (cy26r1 at T_L95L60) and ERA-40 data for the period 1962–2001: (b) winter and (d) summer season. Forecasts for the winter (summer) season were started on 1 October (1 April) 1962–2001. The reference vector length of the wind difference is 1/3 of that for ERA-40 winds.

the summer season the most pronounced errors are found in the North Atlantic where the model considerably overestimates the northeasterly trade winds. Moreover, the model tends to produce a summertime circulation over the North Pacific which is too anticyclonic. Relatively large errors can also be found over the Middle East. Throughout the remainder of this section the focus will be on the Northern Hemisphere where near-surface wind errors appear to be more prominent than over the Southern Hemisphere.

Systematic errors of 1000 hPa geopotential height fields (Z1000, hereafter) are shown in Fig. 22 for the winter and summer season. The strong wintertime Z1000 bias, which has developed over the North Pacific during the first months of the integration (2–6 months forecasts), is clearly evident. In the central North Pacific the model overestimates Z1000 by as much as 10 dam[5]. In the North Atlantic region wintertime systematic Z1000 errors are primarily associated with the north-western and north-eastern flanks of the Icelandic low pressure system.

---

[5]Seasonal integrations with cy26r1 were carried out without using the mass-fixer (the mass increases slightly during the course of the integration). This may explain parts of positive Z1000 errors. The pronounced systematic Z1000 error in the central North Pacific remains almost unchanged, however, if the mass-fixer is applied (L. Ferranti, personal communication).
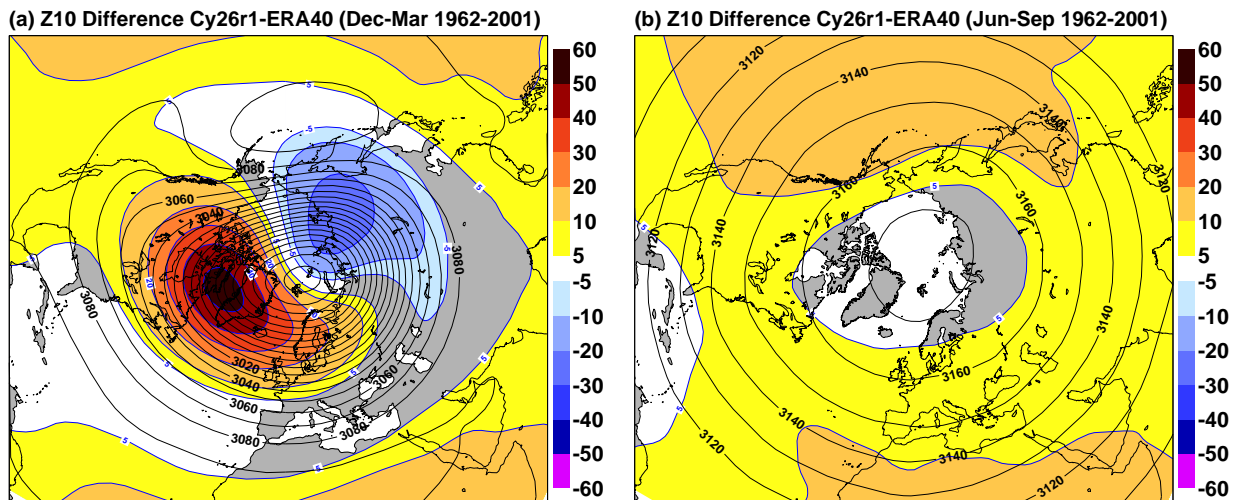
Figure 22: Average 1000 hPa geopotential height fields from ERA-40 (contours in dam) and difference of mean 1000 hPa geopotential height fields (shading in dam) between the ECMWF model (cy26r1 at T_L95L60) and ERA-40 data for the period 1962–2001: (a) winter (Dec–Mar) and (b) summer (Jun-Sep) season. Forecasts for the winter (summer) season were started on 1 October (1 April) 1962–2001.

Systematic Z1000 errors in the North Pacific, over north-eastern parts of Canada, and parts of the Arctic are evident in both seasons (Fig. 22). In the North Atlantic, on the other hand, systematic Z1000 errors show a relatively pronounced seasonality. In summer, like over the North Pacific, the Azores high pressure system appears to be too strong on its northern side. In winter, the Azores high is slightly biased towards the east.

Systematic Z500 errors are shown in Fig. 23. During wintertime the largest systematic Z500 errors are found over the North Pacific and western parts of North America. A comparison between Fig. 22a Fig. 23a further reveals that the strong wintertime bias in the North Pacific low-level flow extends well into the mid-troposphere; systematic errors in other areas are less equivalent barotropic.

Figure 23: Same as in Fig. 22, except for 500 hPa geopotential height fields.

Figure 24: Same as in Fig. 22, except for 10 hPa geopotential height fields.

It is worth stressing that the systematic Z500 error pattern for the latest model cycle only partly resembles the typical systematic Z500 error pattern described in previous studies, that is, the westerly wind bias in the North Pacific and North Atlantic region (Palmer et al., 2000; Brankovic et al., 2002). Whereas systematic Z500 errors in the North Pacific region seem to be slightly more pronounced in model cycle 26r1 compared to the cycles used in the 1990s, the latest model cycle shows much reduced systematic Z500 errors in the North Atlantic region.

During summertime the systematic error structure of geopotential heights at 1000 hPa and 500 hPa is very similar, the only exception being systematic flow errors over North America and the Middle East.

So far, the focus has been on the atmospheric circulation in the lower and middle troposphere. To assess the model's ability to simulate the average stratospheric horizontal circulation, systematic Z10 errors were considered (Fig. 24). During wintertime the trough over Greenland is slightly underestimated and the centre of the simulated polar vortex is slightly displaced towards Siberia. During summertime the model tends to overestimate the north-south Z10 gradient. Overall, however, the model results are in relatively good agreement with estimates from ERA-40.

3.4.2 Medium-Range Errors

In the previous section it has been shown that the flow in the lower and middle troposphere shows substantial biases after more than two months of forecast time. This is particularly true in winter for the North Pacific region. An important question to answer is how fast these errors grow and whether the medium-range forecasts are already affected by such large systematic errors.

Systematic Z500 forecast errors are shown in Fig. 25 for the winters 2001–03 and different forecast time steps (D+1, D+3, D+5, and D+10). The rate at which systematic Z500 errors grow throughout the first 10 forecast days is different for different locations. The positive systematic Z500 error over north-west Canada, for example, grows more or less linearly during the first five days of the forecast. At D+10, however, this error is of minor importance compared to errors in other regions. Furthermore, the model tends to underestimate Z500 south-east of Iceland. This error remains quite small until D+3 and tends to grow at an approximately linear rate until D+10. This error is not so prominent in the 6 months winter integration with the T_L95L60 model
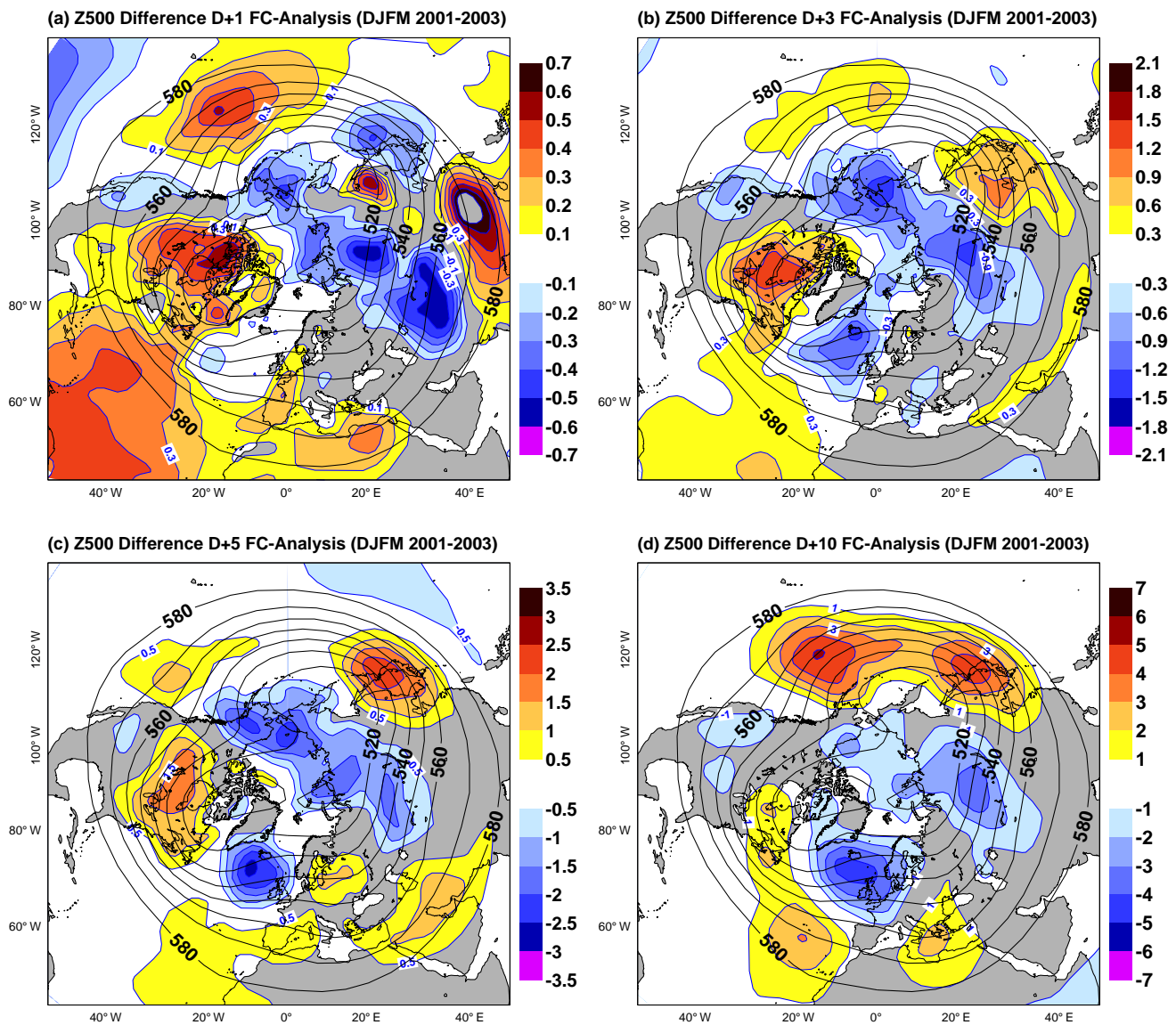
Figure 25: Difference of 500 hPa geopotential height fields (shading in dam) between operational D+3 forecasts and verifying analyses for the three periods 1986–88 (upper panels), 1993–95 (middle panels), and 2001–03 and 2000–02, respectively (lower panels). Results are shown separately for winter (left column) and summer (right column). Also shown are mean 500 hPa geopotential height fields (contours in dam) from the verifying analyses.

(Fig. 23a). This can be partly explained by the rather unusual wintertime flow conditions during the period 2001–03. For these years, Z500 was above (below) the long-term average by about 4 dam close to Iceland (the Azores). Notice, that these anomalies are as large as the "systematic forecast errors" shown in Fig. 25. Therefore, loss of predictability (of the anomalous conditions) might be one explanation for the underestimation (overestimation) of Z500 close to Iceland (the Azores) by the operational ECMWF model at D+10. As will be shown in a later section on atmospheric blocking, this error is also consistent with an underestimation of the number of blocking events events in the ECMWF model in the eastern North Atlantic. Finally, it should be mentioned that the positive Z500 bias over the North Pacific has developed to almost half of its asymptotic size. A comparison of Fig. 25 and Fig. 23a suggests that the growth of systematic Z500 errors continues well
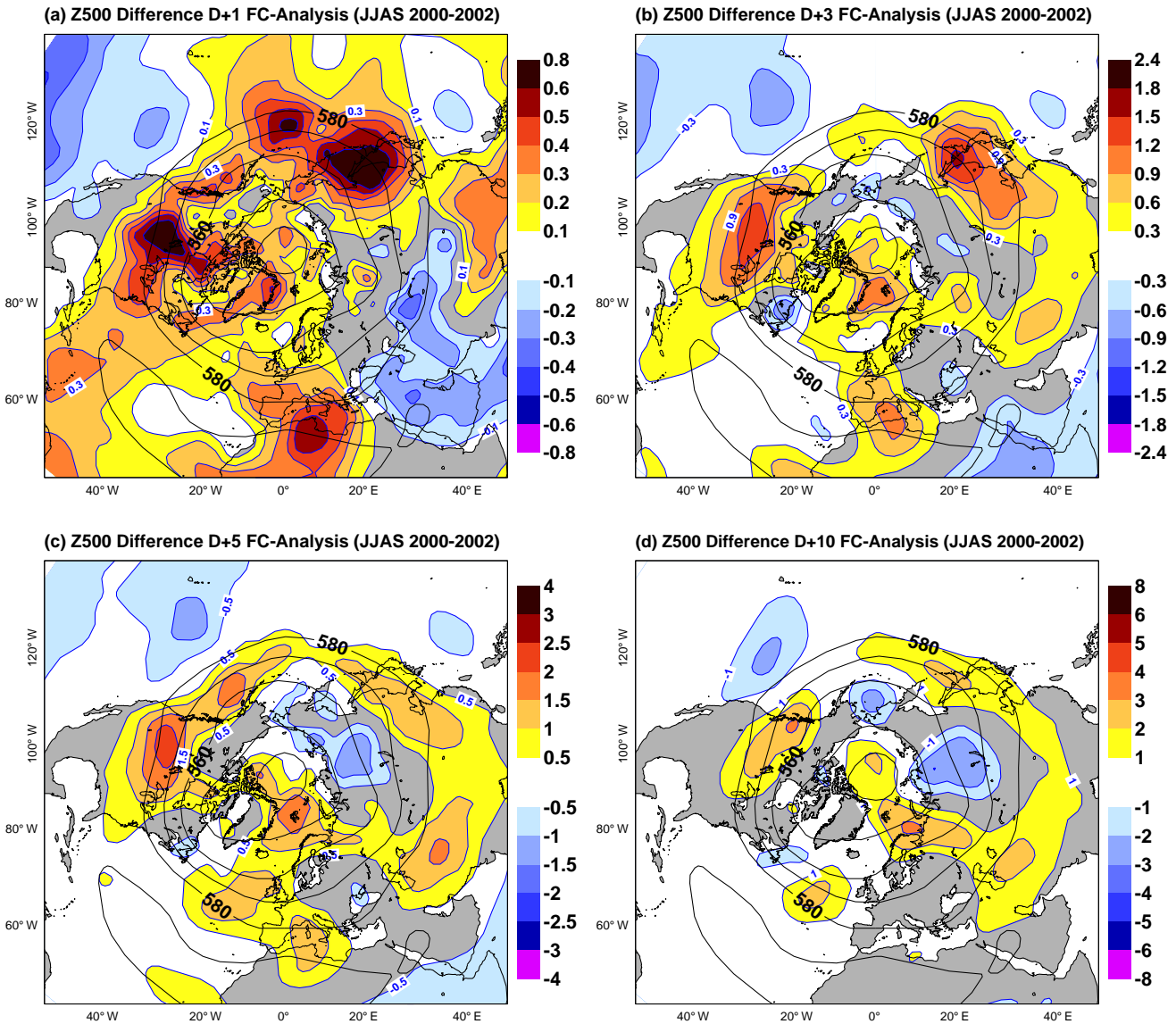
Figure 26: Same as in Fig. 25, except for summertime (Jun–Sep).

into the extended-range over the North Pacific and western parts of North America; the growth seems to have been more or less completed at D+10, however, in all other regions of the Northern Hemisphere.

The Z500 error growth during summertime (Jun–Sep, 2000–02) is shown in Fig. 26. For all different forecast steps the spatial structure of systematic Z500 error closely resembles those of the asymptotic errors (see Fig. 23b). However, systematic Z500 errors at D+10 are still substantially lower than their asymptotic counterparts (more than two months of integration). This suggests that systematic Z500 errors continue to grow beyond the medium-range and/or that there is some sensitivity of the size of the systematic errors to the horizontal resolutions used (T_L95 vs T_L511).
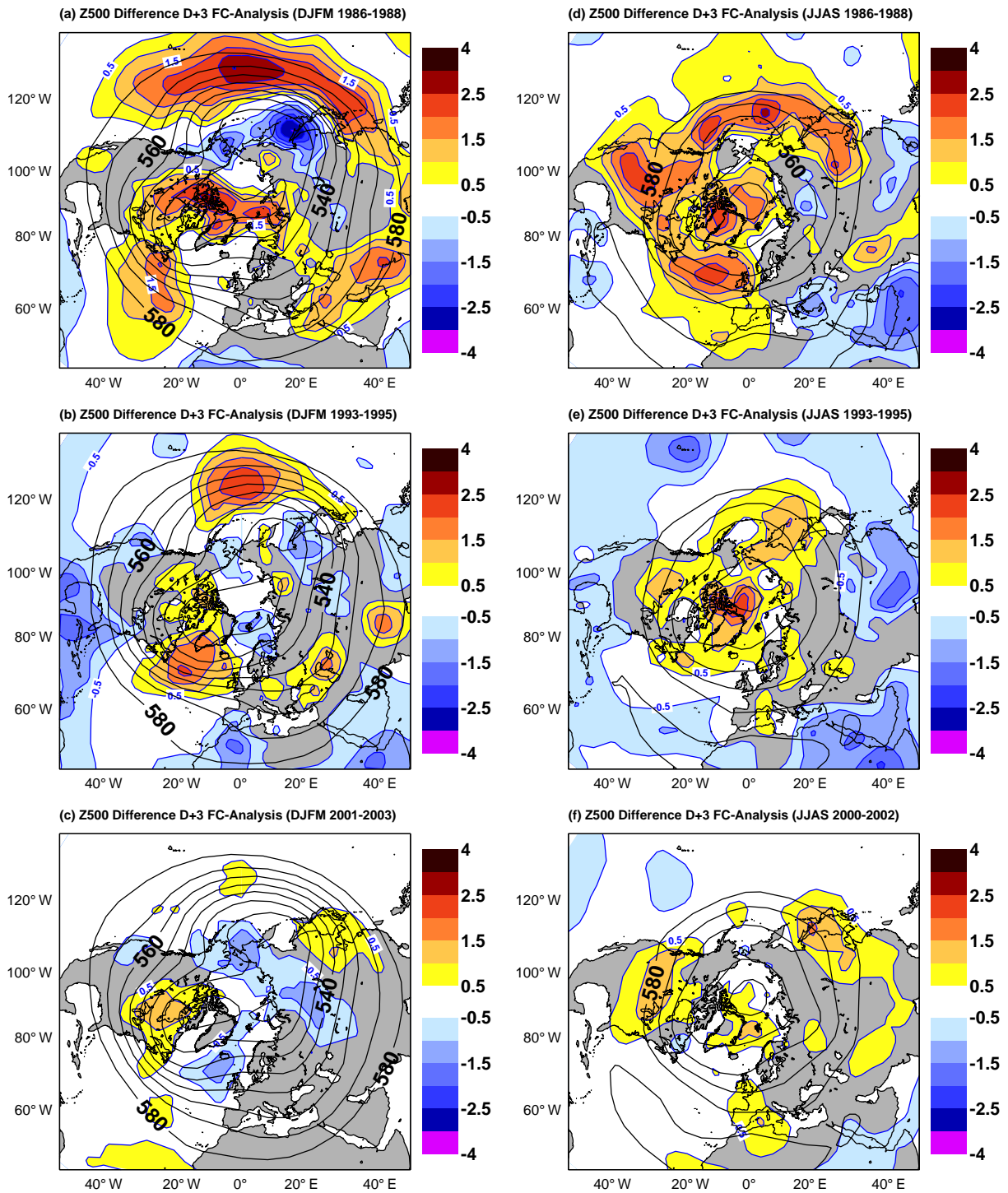
Figure 27: Systematic errors of D+3 forecasts of 500 hPa geopotential height fields (shading, in dam) along with average fields from operational analyses (contours, in dam) for the three periods (upper panels) 1986–88, (middle panels) 1993–95, and (lower panels) 2001–03 and 2000–02, respectively. Results are shown for the (left column) winter and (right column) summer seasons.
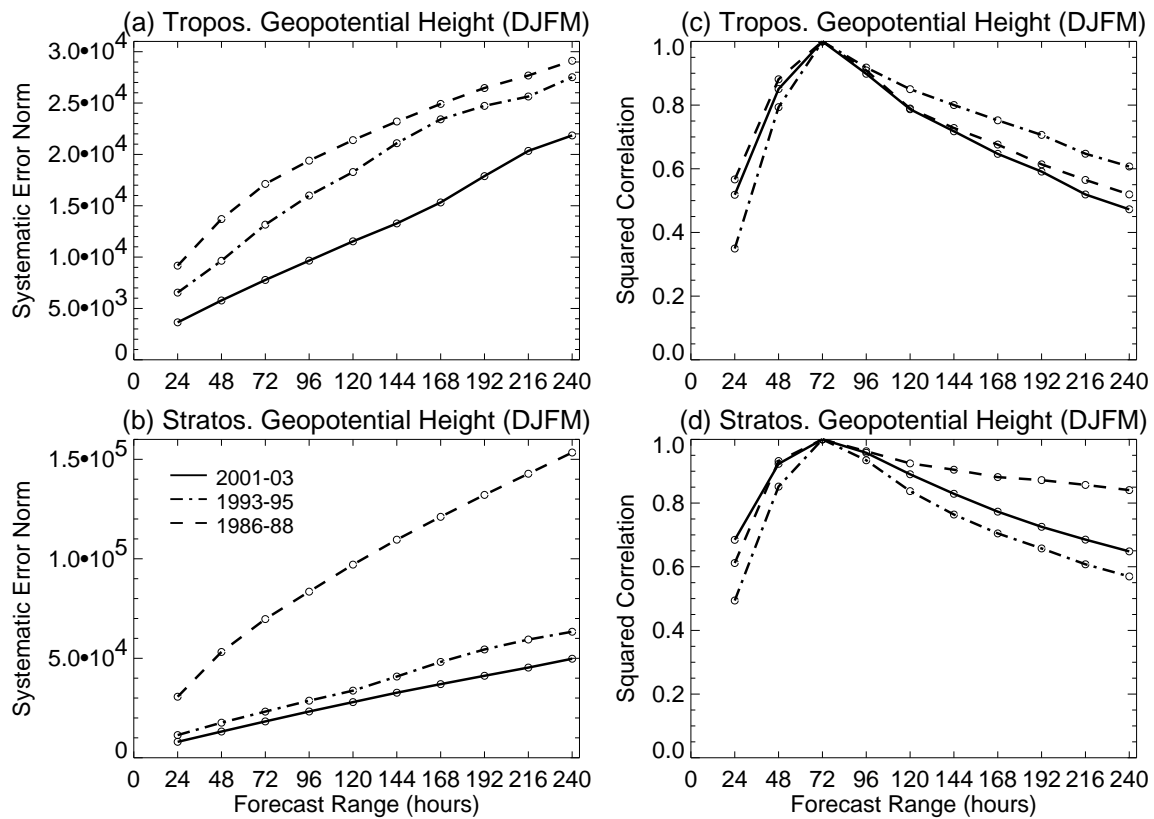
Figure 28: Same as in Fig. 4, except for geopotential height fields.

3.4.3 Evolution of Medium-Range Errors

Systematic Z500 errors for the three period 1985–88, 1993-95, and 2001–03 (2000–02) are shown in Fig. 27 for the winter and the summer season. Systematic Z500 errors have been considerably reduced in the medium-range since the mid-1980s. This is particularly true in the North Pacific. Moreover, the general systematic error structure is quite similar for all periods, both in winter and in summer. From the above results and the structure and magnitude of systematic errors (see Fig. 23), it can be concluded that it is primarily the rate at which Northern Hemisphere systematic Z500 errors grow that has been reduced since the mid-1980s. This is an important improvement, though, given that systematic Z500 errors in the *medium-range* are much smaller than they used to be in the mid-1980s and mid-1990s.

The growth characteristics of wintertime systematic geopotential height errors during the first 10 days of the forecast are shown in Fig. 28 for the troposphere and lower-stratosphere for the three different winter periods (1986–88, 1993–95, and 2001–03). Results for the summer season are very similar (not shown). Evidently, the largest reduction of tropospheric systematic Z500 errors took place during the last couple of years. This is particularly true for the medium-range. In the stratosphere, on the other hand, recent improvements are comparably small compared to those from the mid-1980s to the mid-1990s. Another interesting aspect revealed by Fig. 28 is that the growth of systematic tropospheric and stratospheric geopotential height errors shows little indication for saturation in the far medium-range. This is particularly true for the most recent model cycles (solid curve). These results are consistent with the conjecture made above, that substantial systematic error growth takes place beyond the medium-range.
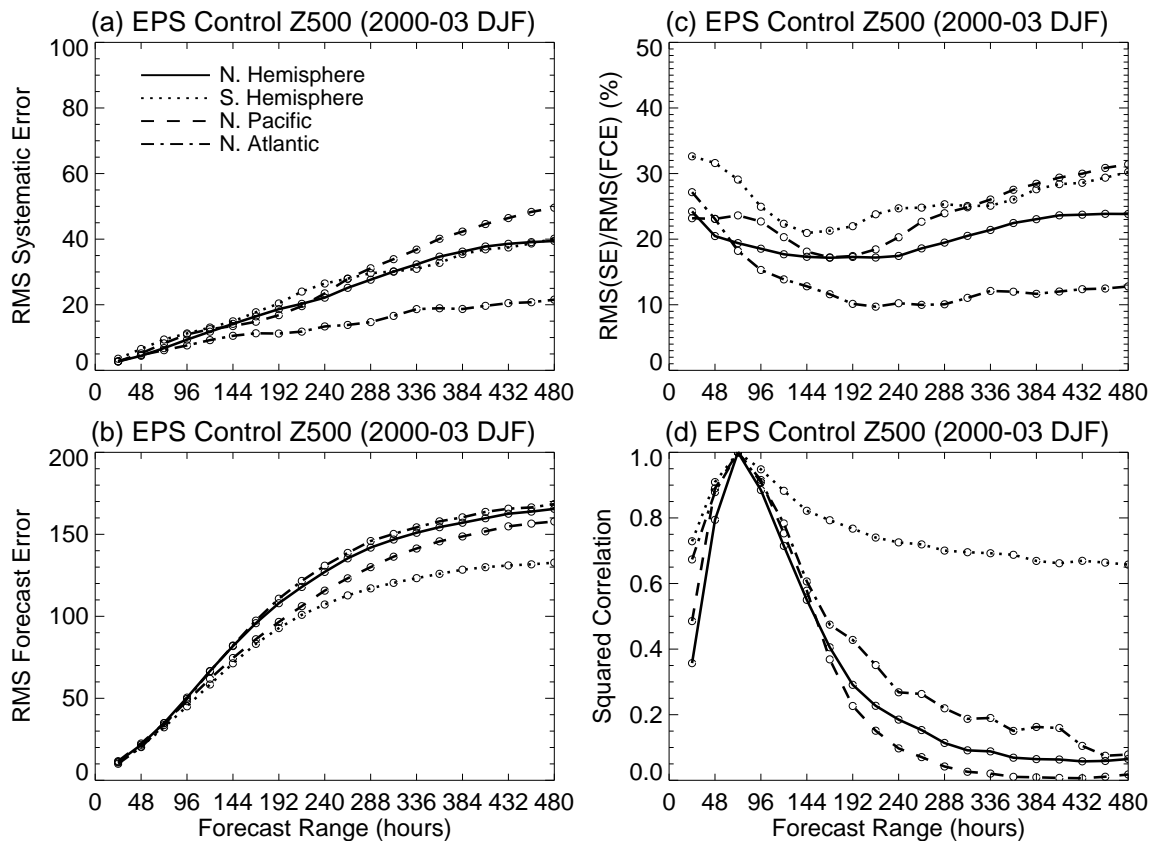
Figure 29: Standard deviation (m) of (a) systematic Z500 errors and (b) Z500 forecast errors for the winters (DJFM) 2000–03 and the Northern Hemisphere (solid), Southern Hemisphere (dotted), North Pacific (dashed), and North Atlantic (dash-dotted). (c) Ratio of (a) to (b) (in %). (d) as in Fig. 28, except for Z500 fields. Results are based on 20 day EPS control forecasts.

As for the other variables considered so far, systematic geopotential height errors show little evidence for significant changes in their spatial structure throughout the first ten days of the forecast (Fig. 28, right column).

The above conclusions are valid only on a *global* scale. From Fig. 25, however, it is evident that the spatial structure of North Pacific systematic Z500 errors changes significantly during the first 10 days of the forecast. Therefore, the above diagnosis has been repeated for the Northern Hemisphere only (not shown). In the short-range and medium-range the conclusions for the Northern Hemisphere are virtually the same as those for the whole globe. In the far medium-range, however, systematic Northern Hemisphere *tropospheric* geopotential height errors during the period 2001–03 were worse than those during the period 1993–95. Furthermore, systematic errors tend to change their spatial structure more strongly during the first ten forecast days over the Northern Hemisphere.

Changes in the magnitude of systematic geopotential height errors at 500 hPa during the first 20 days of the EPS control forecast are shown in Fig. 29a for the Northern Hemisphere, Southern Hemisphere, North Pacific, and the North Atlantic. For the first three regions the largest error growth can be found, particularly in the North Pacific. By far the smallest errors occur in the North Atlantic. There is some indication that the systematic error growth levels off around D+20. Forecast errors for different lead times and the same regions are shown in Fig. 29b. The ratio between the systematic error magnitude and the forecast error, which is shown in Fig. 29c, quantifies to what extent systematic errors contribute to total forecast errors. Depending on the region, system-

atic Z500 errors amount to 25–35% of the forecast error in the very short-range. Lowest values are found in the medium-range. Others than in the North Atlantic region, in the Northern Hemisphere (primarily due to the North Pacific) and the Southern Hemisphere the importance of systematic errors increases again between D+10 and D+20. This is due to the fact that forecast errors level off whereas systematic Z500 continue to grow at a linear rate.

3.4.4   Extended-Range Errors

In the previous section it has been argued that further growth of systematic Z500 errors is likely to take place in the extended-range. This conjecture was mainly based on comparisons between asymptotic systematic Z500 errors and those evident in the medium-range. To further highlight the systematic error growth during the first 30 days of the integration (medium-range and extended-range) data from the Monthly Forecasting System have been diagnosed.

Systematic Z500 differences between monthly forecasts and verifying analyses for the winter season are shown in Fig. 30 for different forecast ranges. Results are based on a total of 108 monthly forecasts, each comprising one control forecast and 4 ensemble members, started between November and February of the years 1990 to 2002. It is evident that the spatial structure of systematic Z500 errors hardly changes after the late medium-range. Second, the positive Z500 error in the central North Pacific continues to grow considerably in the extended-range. This is in contrast to systematic Z500 errors in other regions which appear to be more or less fully developed after D+15. Finally, it should be mentioned that the negative Z500 errors over Great Britain are relatively pronounced. This kind of error does not seem to be a result of the relatively low resolution used in the Monthly Forecasting System since similar errors are also apparent in high-resolution forecasts with recent model cycles (compare Fig. 25).

As for the medium-range, the model tends to underestimate Z500 in the Euro-Atlantic region beyond the medium-range, an error which does not occur in the 6 months integration over the 40-year period 1962–2001 (see Fig. 23). The perhaps simplest explanation is that the atmospheric wintertime circulation during the period 1990–2002, for which the Monthly Forecasting System has been diagnosed, has been relatively unusual, and, therefore, that loss of predictability in the extended-range leads to apparent systematic model errors that resemble the opposite of the observed Z500 anomalies. We have looked at Z500 epoch differences between 1990–2002 and 1975–1989 using ERA-40 data. During the former period, Z500 indeed was anomalously high over the United Kingdom. Therefore, at least partly, the negative Z500 forecast errors apparent in the Monthly Forecasting System in the Euro-Atlantic region are due to the anomalous circulation.

Results for 108 monthly forecasts started during the months May to August are shown in Fig. 31. Similar to the winter season, the structure of systematic Z500 errors hardly changes in the extended-range. Moreover, the error growth seems to be completed after two weeks into the integration. Further, the similarity of systematic Z500 errors of the Monthly Forecasting System and those from the high-resolution model, which is used to perform medium-range forecasts (cf. Fig. 26), suggest rather low sensitivity of these errors to horizontal and vertical resolution, and, indeed, ocean-atmosphere coupling.

3.5   Kinetic Energy

An important diagnostic to characterize the atmospheric circulation is the kinetic energy. In this section the ability of the ECMWF model to simulate observed kinetic energy characteristics is considered.

The zonally averaged mean total kinetic energy per unit mass K can be decomposed into the following expres-
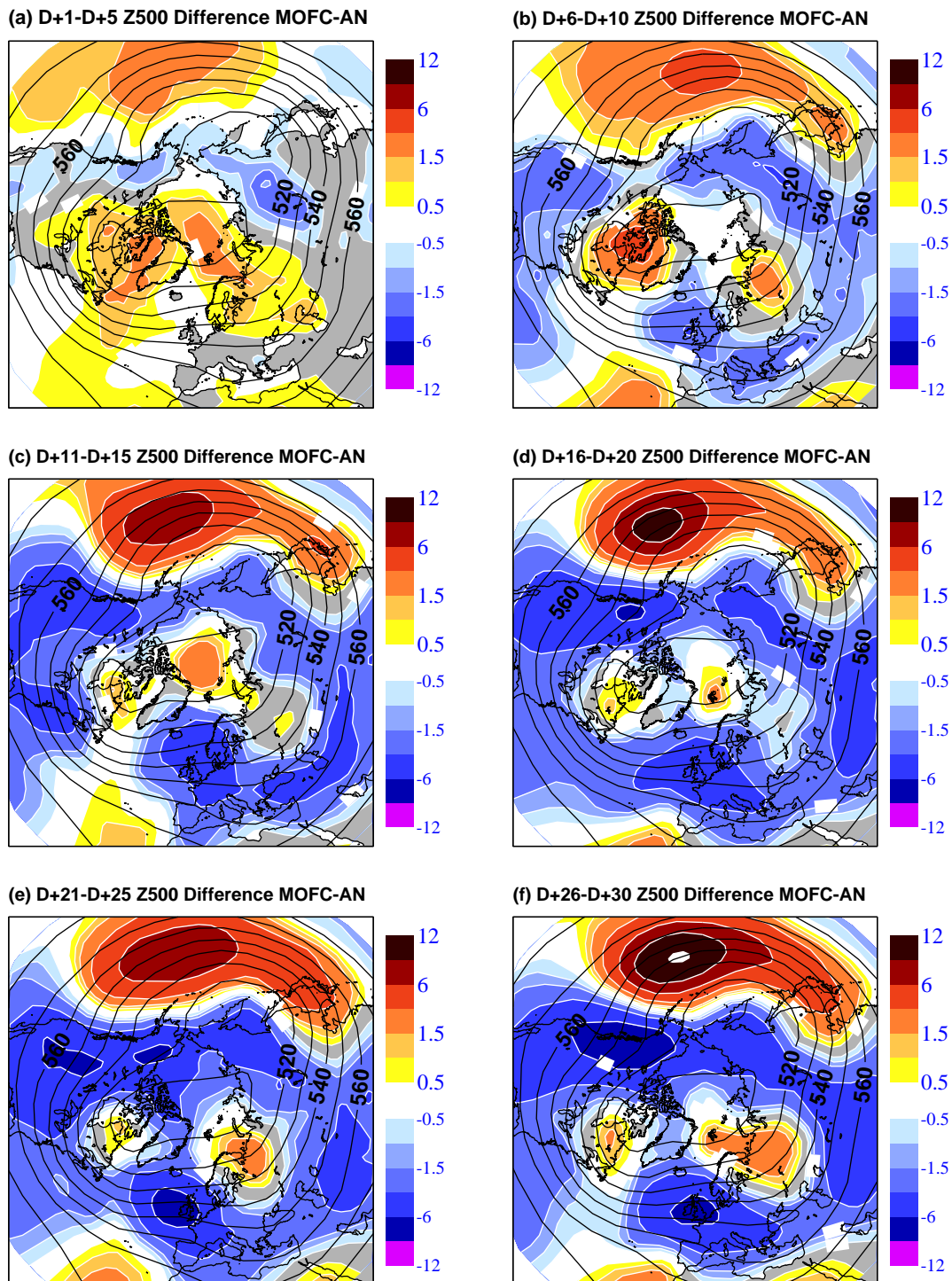
Figure 30: *Difference of 500 hPa geopotential height fields (shading in dam) between ($T_L159L40$) forecasts and verifying analyses: (a) D+1 to D+5, (b) D+6 to D+10, (c) D+11 to D+15, (d) D+16 to D+20, (e) D+21 to D+25, and (f) D+26 to D+30. Also shown are mean values of the verifying analysis (contours in dam). The results are based on 108 wintertime integrations of the Monthly Forecasting System over a total of 30 days each. Notice, that the values shown in (a)–(f) are based on averages over 5 consecutive analysis/forecast days.*
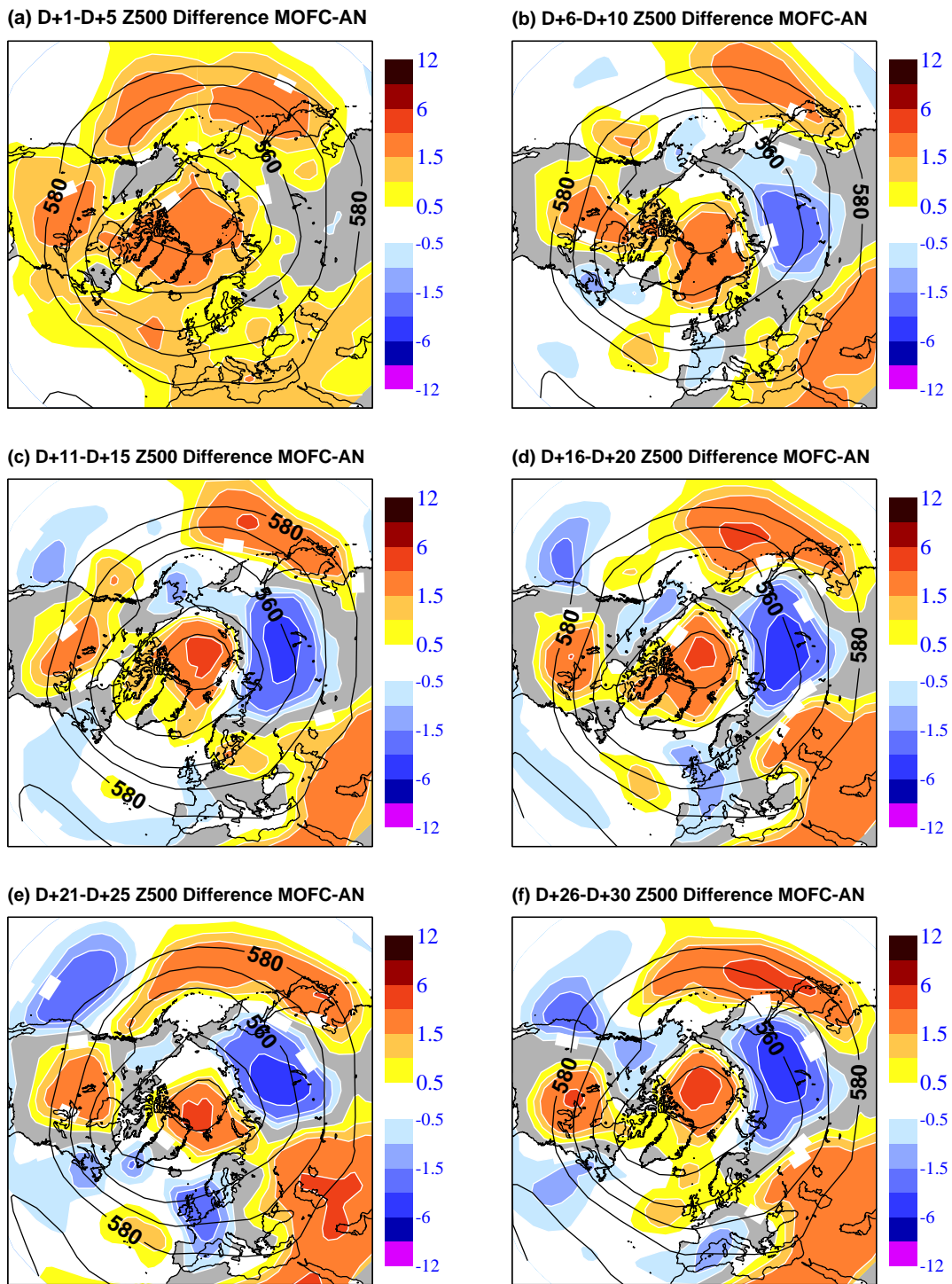
(a) D+1-D+5 Z500 Difference MOFC-AN

(b) D+6-D+10 Z500 Difference MOFC-AN

(c) D+11-D+15 Z500 Difference MOFC-AN

(d) D+16-D+20 Z500 Difference MOFC-AN

(e) D+21-D+25 Z500 Difference MOFC-AN

(f) D+26-D+30 Z500 Difference MOFC-AN

Figure 31: Same as in Fig. 30, except for summertime. Results are based on a total of 108 monthly forecasts.

sion (e.g. Peixoto and Oort, 1992, for details):
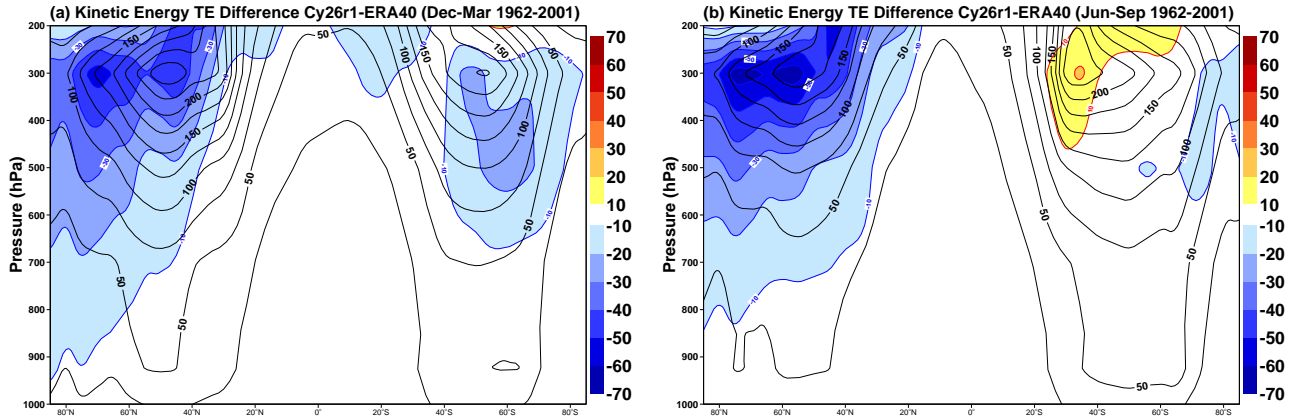
$$K = K_M + K_{TM} + K_{SE} + K_{TE}, \qquad (2)$$

Figure 32: Systematic errors of the kinetic energy of transient eddies (K_{TE}, shading in $m^2 s^{-2}$) for (a) winters (Dec-Mar) and (b) summers of the period 1962–2001. Also shown is the kinetic energy of transient eddies obtained from ERA-40 (contours). 6-month integrations based on model cycle 26r1 were started on 1 October of each calendar year.

with

$$K = \frac{1}{2}\left([\overline{u \cdot u}] + [\overline{v \cdot v}]\right), \tag{3}$$

$$K_M = \frac{1}{2}\left([\overline{u}] \cdot [\overline{u}] + [\overline{v}] \cdot [\overline{v}]\right), \tag{4}$$

$$K_{TM} = \frac{1}{2}\left(\overline{[u]' \cdot [u]'} + \overline{[v]' \cdot [v]'}\right), \tag{5}$$

$$K_{SE} = \frac{1}{2}\left([\overline{u}^* \cdot \overline{u}^*] + [\overline{v}^* \cdot \overline{v}^*]\right), \qquad \text{and} \tag{6}$$

$$K_{TE} = \frac{1}{2}\left([\overline{u'^* \cdot u'^*}] + [\overline{v'^* \cdot v'^*}]\right), \tag{7}$$

where brackets denote zonal mean averages, overlines temporal means, asterisks departures from the zonal mean, and primes departures from the temporal mean. K_M denotes the zonally averaged mean circulation and contains, for example, contributions from the subtropical jet stream; K_{TM} denotes temporal variation of the zonally averaged flow including, for example, temporal variations in the strength of the Hadley circulation; contributions from stationary waves are described by K_{SE}; whereas K_{TE} reflects contributions from transient eddies. Note, that K_{TE} takes into account both fluctuations due to synoptic-scale systems and low-frequency intraseasonal modes of variability (e.g., the North Atlantic Oscillation). In the following the focus will be on systematic errors of K_{TE} during wintertime.

3.5.1 Systematic Errors

Systematic K_{TE} errors (shading) are shown in Fig. 32 for winter and summer of the period 1962–2001 together with K_{TE} estimates obtained from ERA-40 (contours). It is evident that over the Northern Hemisphere the model systematically underestimates K_{TE} by up to 15% and 30% during the winter and summer season, respectively. The largest errors are found in the upper troposphere north of about 40^oN. Over the Southern Hemisphere systematic K_{TE} errors are much smaller. Moreover, and in contrast to the Northern Hemisphere, errors over the Southern Hemisphere show a relatively pronounced annual cycle.

As has been noted above, K_{TE} contains contributions from synoptic systems as well as from low-frequency
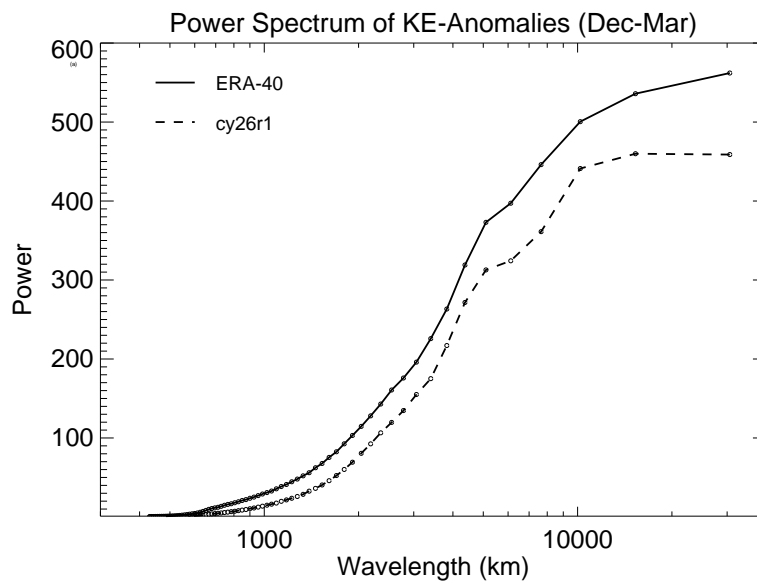
Figure 33: Average wavenumber spectra of temporal anomalies of wintertime (Dec–Mar) kinetic energy at 50°N and pressure level 300 hPa for ERA-40 data (solid) and model cycle 26r1 (dashed). Estimates are based on forty years (1962–2001) of daily data.

intraseasonal variations (weather regimes). Therefore, from the above diagnostics it is difficult to infer whether only one or both of these phenomena contribute to the systematic K_{TE} errors shown in Fig. 32. In order to further hightlight the nature of the underestimation of Northern Hemisphere K_{TE}, average wavenumber spectra have been estimate for temporal anomalies of the reanalysed (solid) and simulated (dashed) kinetic energy at 50°N at pressure level 300 hPa (Fig. 33). This longitude section coincides with the region of maximum underestimation of K_{TE}. Fig. 33 clearly shows that kinetic energy anomalies are underestimated on all spatial scales. The general structure of the power spectrum, however, is well-captured by the model. This includes the small spectral peak which is evident on spatial scales of 4000–5000 km.

3.5.2 Errors in the Medium-Range

Systematic K_{TE} differences between medium-range forecasts and verifying operational analyses are shown in Fig. 34 for the winters 2001–03. As for the asymptotic errors of the lower-resolution model, the high-resolution model clearly underestimates K_{TE} in the upper-troposphere over the Northern Hemisphere; the magnitude of these errors of the high-resolution model at D+10, however, is considerably smaller than the systematic errors of the T_L95 model suggesting that there is still some further error growth beyond the medium-range[6]. As for the low-resolution model, relatively small errors are found over the Southern Hemisphere.

Systematic K_{TE} differences between medium-range forecasts and verifying operational analyses for the winters 1993–95 are shown in Fig. 35. Over the Northern Hemisphere the model cycles used in the early 1990s show smaller systematic K_{TE} errors than the most recent model cycles. Over the Southern Hemisphere, however, the most recent model cycles perform better.

---

[6]Alternatively, the growth might be finished at D+10 for the high-resolution model. This would suggest that there is some sensitivity of error growth in the extended-range to horizontal resolution. Since currently no integrations of the T_L511 model are available going beyond the medium-range this question cannot be answered at present. Experiences with resolutions of T_L255 and T_L159 suggests, however, that the sensitivity to horizontal resolution is rather small.
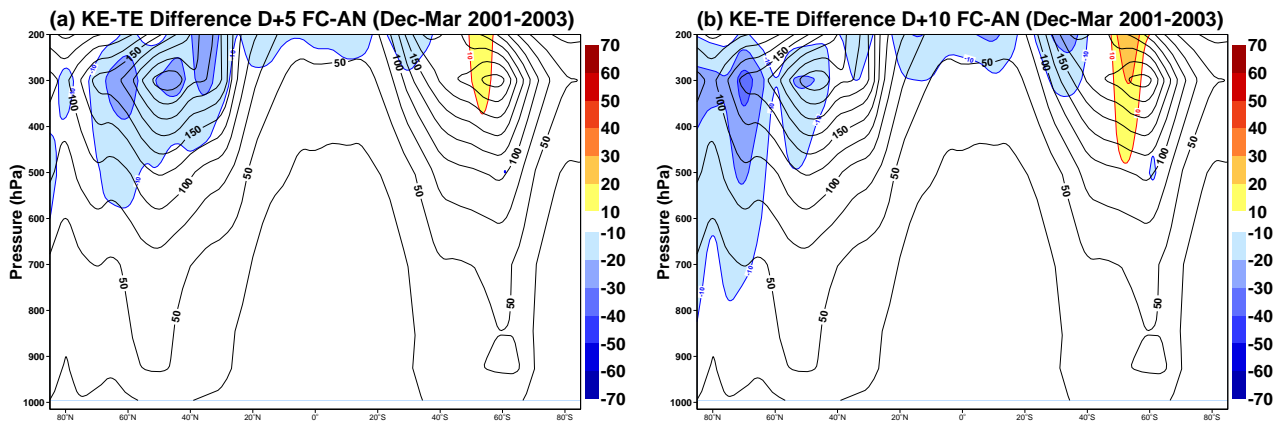
Figure 34: Medium-range systematic errors of the kinetic energy of transient eddies (K_{TE}, shading in $m^2 s^{-2}$) for (a) D+5 and (b) D+10 forecasts. Results are for the winter season Dec–Mar. Also shown are estimates of K_{TE} from operational analyses (contours).
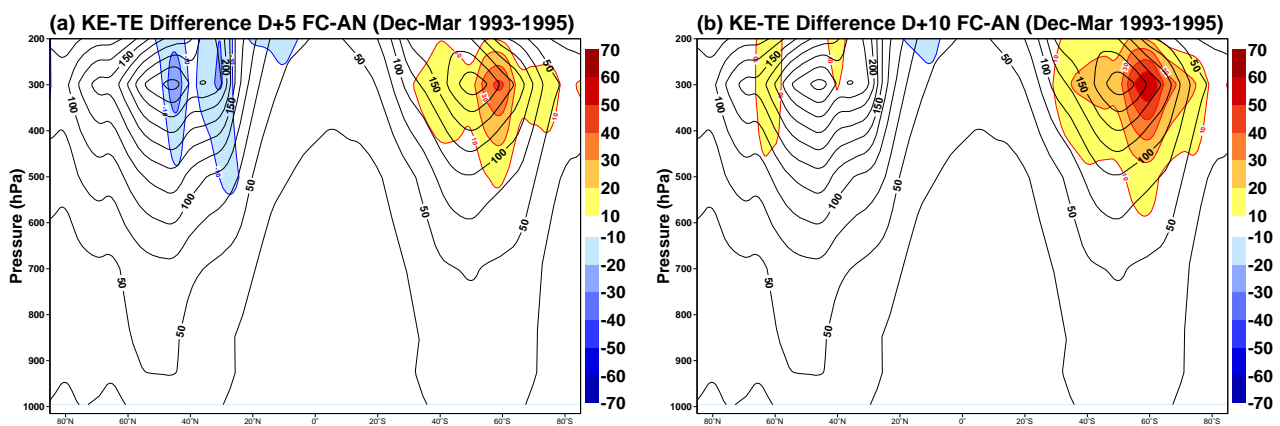

Figure 35: Same as in Fig. 34, except for the period 1993–95.

Average wavenumber spectra of kinetic energy anomalies provide evidence that the underestimation of K_{TE} in the medium-range is due to an underestimation of transient kinetic energy on all spatial scales (not shown)

3.6 Synoptic Activity

In this section the ECMWF model's ability to simulate synoptic variability is described. The focus is on systematic errors over the Northern Hemisphere during winter and summer.

There are different ways to characterize synoptic scale variability. Here, we chose to consider high-frequency variations of 500 hPa geopotential height fields (Z500). The filtering is accomplished by means of a simple differencing technique ($y_t = x_t - x_{t-\Delta}$). In this study a value of $\Delta = 1$ day is used for the analysis and for the forecast. Notice that for the forecast the difference is taken between consecutive days of a forecast started at the same date (e.g., D+3 minus D+2). The response function of the difference filter is shown in Fig. 36. The half-power width of the filter lies at 10 days, that is, the bulk of the variance of the highpass-filtered Z500 data arises from variations on time scales shorter than 10 days (the synoptic frequency band). The difference filter
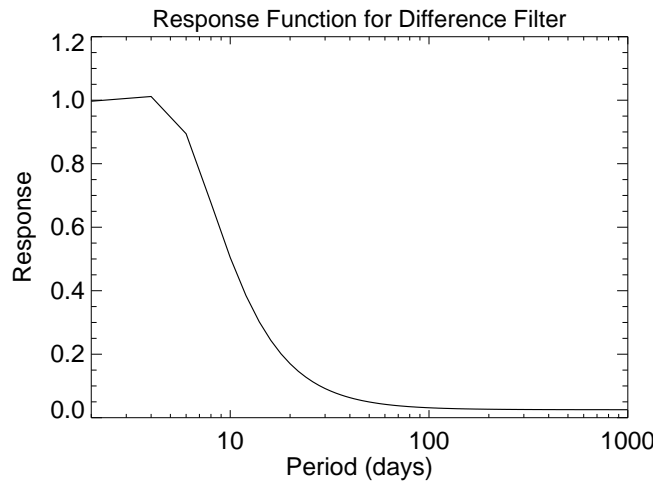
Figure 36: Frequency response function of the simple difference filter $y_t = x_t - x_{t-\Delta}$ which is used throughout this study with $\Delta = 1$ day to obtain synoptic scale variability. The response gives the amplification of a cosine wave of period T when it passes the difference filter.

is one of the crudest filters in the sense that the decline of the response function towards longer timescales is rather gradual. However, the advantage of this filter compared to more sophisticated digital filters lies in the fact that it can be used for very short timeseries. This is crucial when high-frequency components are to be determined from medium-range forecasts (i.e., the time series is no longer than 10 days). For the six month integrations it would be possible to use other filters. The use of different filters for different datasets, however, would inhibit a fair comparison of the results.
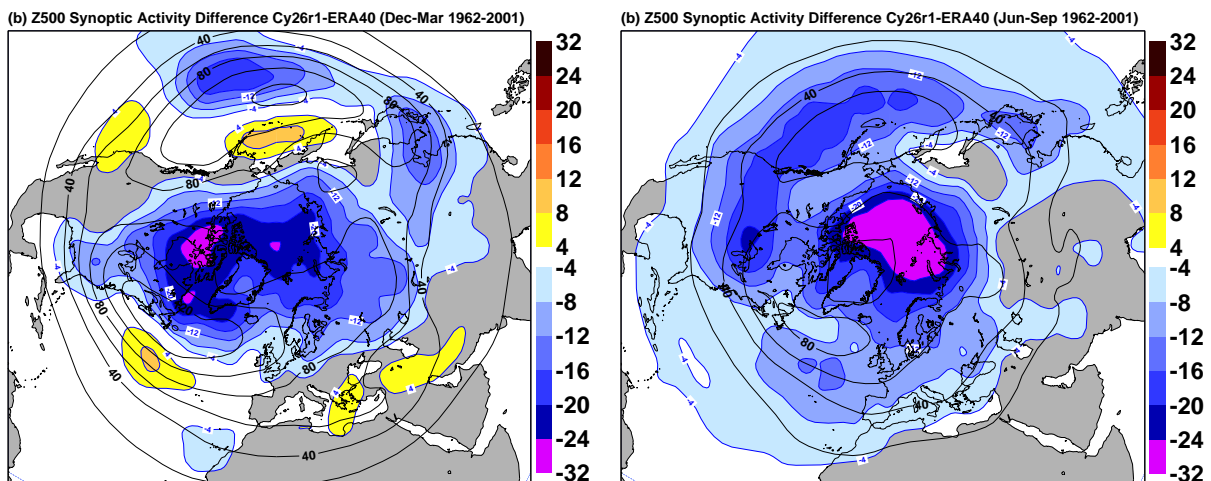
3.6.1 Systematic Errors

Figure 37: Synoptic Z500 activity (solid contours, in m day^{-1}) and difference in Z500 synoptic activity (colour shading, in m/day) between the model (T_L95L60, cycle 26r1) and ERA-40 data: (a) winter season (December–March) and (b) summer season (June–September). Synoptic activity has been determined by taking the standard deviation of day-to-day Z500 changes.
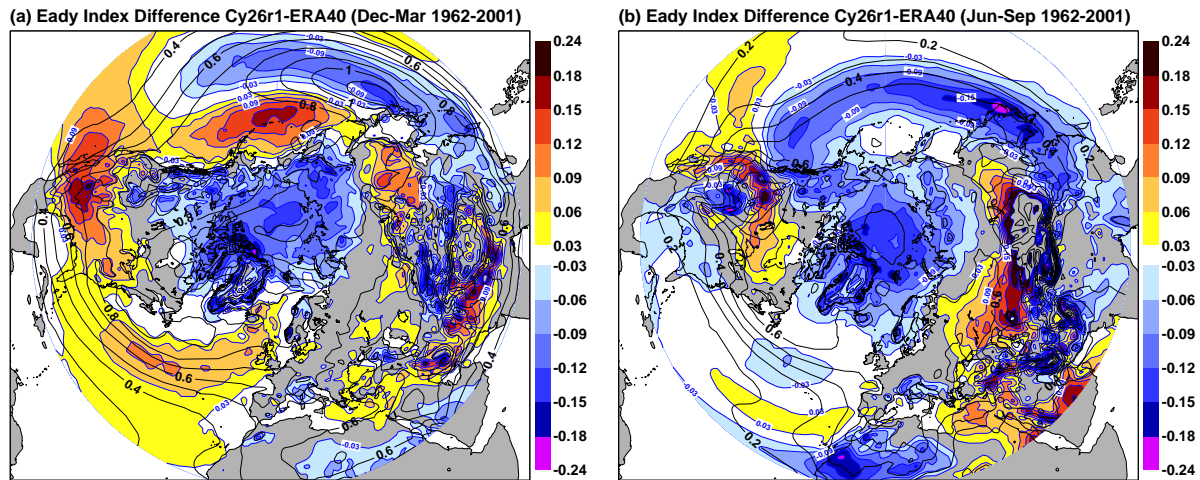
Figure 38: Local Eady indices for ERA-40 data (contours in 0.1 day^{-1}) along with systematic Eady index differences (colour shading) between the model and ERA-40 data: (a) winter season (Dec–Mar) and (b) summer season (Jun-Sep). Computations are based on pressure levels 500hPa and 850hPa (low-level baroclinicity). Notice, that the 850hPa level lies under the surface in the Rocky Mountains, Greenland and the Himalaya.

The standard deviation of high-frequency Z500 variability (referred to as synoptic variability) is shown in Fig. 37 for winter (Dec–Mar) and summer (Jun-Sep) for the ECMWF model (cycle 26r1 at T$_L$95L60) and ERA-40 data together with the systematic model error (difference between model and ERA-40). Winter (summer) results are based on 6 month integrations started on 1 October (1 April).

The difference filter clearly identifies the major Northern Hemisphere storm tracks over the North Pacific and North Atlantic (contours). Moreover, differences in synoptic activity between summer and winter are well captured.

During winter the ECMWF model clearly underestimates synoptic variability in high latitudes and over the Arctic by as much as 25% (Fig. 37a). Over the North Atlantic synoptic variability is slightly overestimated south of the mean storm track. This implies that the mean North Atlantic storm track is slightly shifted towards the south compared to the ERA-40 data. In the North Pacific, on the other hand, the model tends to place the storm track too far to the north. Furthermore, the North Pacific storm track is underestimated at its southern flank by up to 25%.

During summer the simulated North Pacific storm track is underestimated by up to 30% (Fig. 37b). Systematic errors of synoptic variability extend downstream and show a secondary maximum over North America. The core of the North Atlantic storm track, in contrary, appears to be well simulated. As for the winter season, synoptic variability is considerably underestimated over the Arctic (locally by more than 30%). In contrast to the winter season, no areas of positive systematic errors are found during summer over the Northern Hemisphere.

Some of the systematic errors of synoptic activity are consistent with those of Z500, particularly during the winter season (compare Fig. 23a with Fig. 37a). In the central (high-latitude) North Pacific, for example, systematic Z500 errors imply reduced (enhanced) westerly winds, which is consistent with reduced (enhanced) day-to-day variations of Z500. From the above diagnostics, however, it is difficult to separate between cause and effect. This is because synoptic eddies influence the mean circulation, and the mean circulation feedbacks onto the eddies.

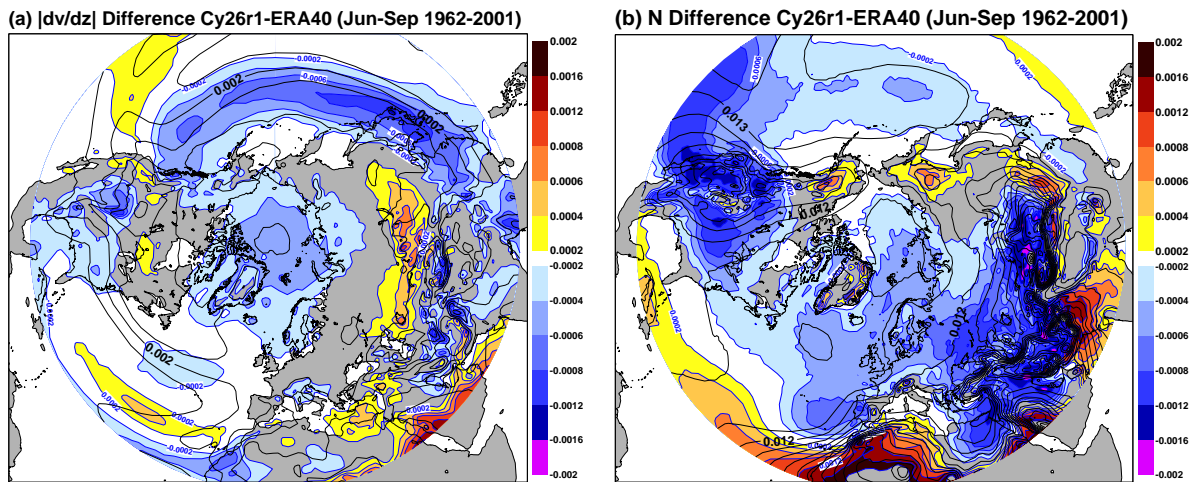One possible explanation for the above systematic errors is that the model has problems in a proper representa-

(a) |dv/dz| Difference Cy26r1-ERA40 (Jun-Sep 1962-2001)   **(b) N Difference Cy26r1-ERA40 (Jun-Sep 1962-2001)**

Figure 39: Same as in Fig. 38, except for (a) the magnitude of low-level vertical wind shear ($|\partial\mathbf{v}/\partial z|$) ($s^{-1}$) and (b) the Brunt-Vaisala frequency (s^{-1}) for summertime.

tion of low-level baroclinicity. In order to test this hypothesis daily values of the Eady index (Hoskins and Valdes, 1990) have been computed. The Eady index is a useful diagnostic to quantify low-level baroclinicity and takes into account contributions from the static stability and vertical wind shear (taken at 500 and 850 hPa). In Fig. 38 the difference between simulated and reanalysed values are shown for wintertime and summertime averages over the 40-year period 1962–2001. The ECMWF model underestimates low-level baroclinicity over the Arctic in both seasons. Furthermore, daily values of the Eady index are underestimated in the south-western part of the North Pacific storm track in both seasons. This underestimation, which amounts to about 10–15% of the mean for ERA-40, is consistent with the underactivity of the model over the Arctic and the central North Pacific. Too high values of synoptic activity in the model over the northern North Pacific, over California, and the central North Atlantic (Fig. 37a) during wintertime are also matched by too high low-level baroclinicity in the model runs. In summary, the gross structures of systematic synoptic activity errors are consistent with model failures to captured "observed" (ERA-40) values of low-level baroclinicity in both seasons.

The Eady index is directly proportional to the magnitude of the vertical wind shear and inversely proportional to the static stability. Thus, it seem worthwhile to seperately consider the influence of these two properties. The systematic errors in the magnitude of the vertical wind shear and static stability, expressed in terms of the Brunt-Vaisala frequency, are shown in Fig. 39 for the summer season. Over the North Pacific systematic errors of the Eady index (Fig. 38b) are largely due to an underestimation of the vertical wind shear. Similar arguments hold for the Artic. Over North America, North Africa and large parts of Asia (i.e., over the continents), on the other hand, systematic errors of the Eady index can largely be explained by systematic errors of the Brunt-Vaisala frequency. Note, that the underestimation of low-level static stability over North America amounts to more than 10%.

It should be emphasized, however, that it is difficult to infer causality from the above diagnostics. The large-scale flow, for example, influences the vertical wind shear which affects synoptic systems which in turn affect the large-scale circulation.

The development of synoptic systems is partly fueled by the release of latent heat (e.g. Hoskins and Valdes, 1990; Wernli and Davis, 1997). To test whether the supply of latent energy through the air-sea interface is underestimated in the model, surface latent heat fluxes have been diagnosed. We found little indication, however, for surface latent heat fluxes errors to be the major source of the above systematic errors of synoptic Z500

variability (not shown).

3.6.2   Evolution of Medium-Range Errors

Synoptic variability has been computed in the same way as above from operational medium-range forecast for three different periods. Earlier versions of the ECWMF model are covered by the period 1986–88. The second period considered encompasses the mid-1990s (1993–95) and the last period is based on the most recent model cycles (2001–03). Notice that the difference filter has been applied to daily forecasts from D+5 to D+10, that is, synoptic variability in the medium-range is covered[7]. The verification is based on operational analyses. Results are shown in Fig. 40 for the winter and summer season.

It is evident from Fig. 37 and Fig. 40 that systematic errors of synoptic activity in the medium-range during the mid-1980s have the same structure and magnitude as *asymptotic* (errors that have developed after two months of integration) systematic errors in one of the most recent ECMWF model cycles (cy26r1). In contrast, systematic model errors in the medium-range during the periods 1993–95 as well as 2001–03 (2000–02 for summer) are much smaller than the asymptotic errors. This comparison, thus, shows that the latest ECMWF model still has the same systematic errors in terms of synoptic variability as the model cycles used during the mid-1980s; the rate at which these systematic errors grow, however, has been considerably reduced between the 1980s and mid-1990s.

During the summer season systematic errors in simulating synoptic variability are lowest in 2000–02 compared to the two previous periods, suggesting that the growth rate has been further reduced after the mid-1990s. During the winter season, however, the underestimation of simulated synoptic activity in the medium-range over the Arctic as well as in the North Atlantic and North Pacific storm tracks appears to be more pronounced in the latest model cycles compared to those used in the mid-1990s.

It is interesting to note at this point that the model tends to overestimate synoptic activity west of Great Britain and in the northern North Pacific (Fig. 40). These regions are known for the frequent occurrence of persistent atmospheric blocking episodes. Too high levels of synoptic activity in the model in these regions suggest that the frequency of blocking events (and their duration) is underestimated by the model. That this is indeed the case will be shown in Section 3.7.

3.6.3   Systematic Error Growth

In this section the question is addressed as to how systematic errors of synoptic activity grow throughout the first 20 days of the forecast. The focus is on the winters of the period 2001–03. Results are based on daily D+20 EPS control forecast. Differences of synoptic acticity between EPS control forecasts and verifying operational analyses are shown in Fig. 41 for four different forecast ranges. It is worth noting that a direct comparison for the D+5 to D+10 forecast range between EPS control ($T_L255L40$, Fig. 41b) and high-resolution deterministic forecasts ($T_L511L60$, Fig. 40c) reveals that systematic synoptic activity errors show little sensitivity to horizontal and vertical resolution. Major changes of systematic synoptic activity errors during the 20 days of the forecast embrace predominantly the magnitude of systematic errors; the horizontal structure of the errors remains virtually unchanged. Except for the North Pacific region the systematic error growth seems to be more or less completed between D+10 and D+15. In the North Pacific, however, some error growth is still evident beyond D+15.

---

[7]It is worth stressing that it is essential to perform highpass-filtering for consecutive days of single forecasts. If instead day-to-day changes of D+10 forecasts were to be considered, then (large) parts of the high-frequency variance would be due to day-to-day inconsistencies resulting from the loss of predictability throughout the forecast range.
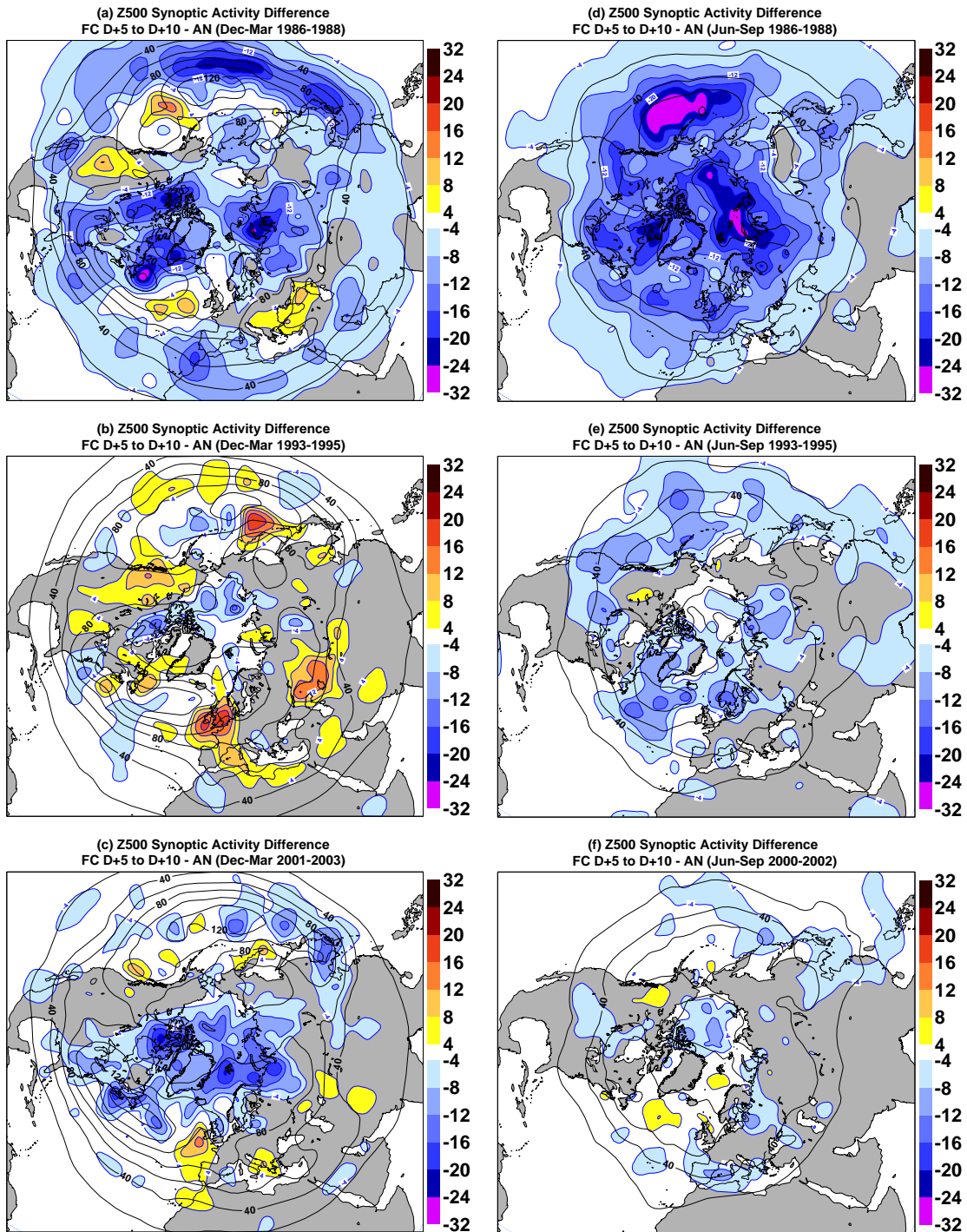
Figure 40: Synoptic Z500 activity (solid contours, in m day^{-1}) and difference in Z500 synoptic activity (colour shading, in m day^{-1}) between operational D+5 to D+10 model forecasts and verifying operational analyses for the periods (upper panels) 1986–88 (middle panels) 1993–95, and (lower panels) 2001–2003 and 2000–2002, respectively. Results are shown for the winter (left panels) and summer seasons (right panels). Synoptic activity has been determined by taking the standard deviation of day-to-day Z500 changes.
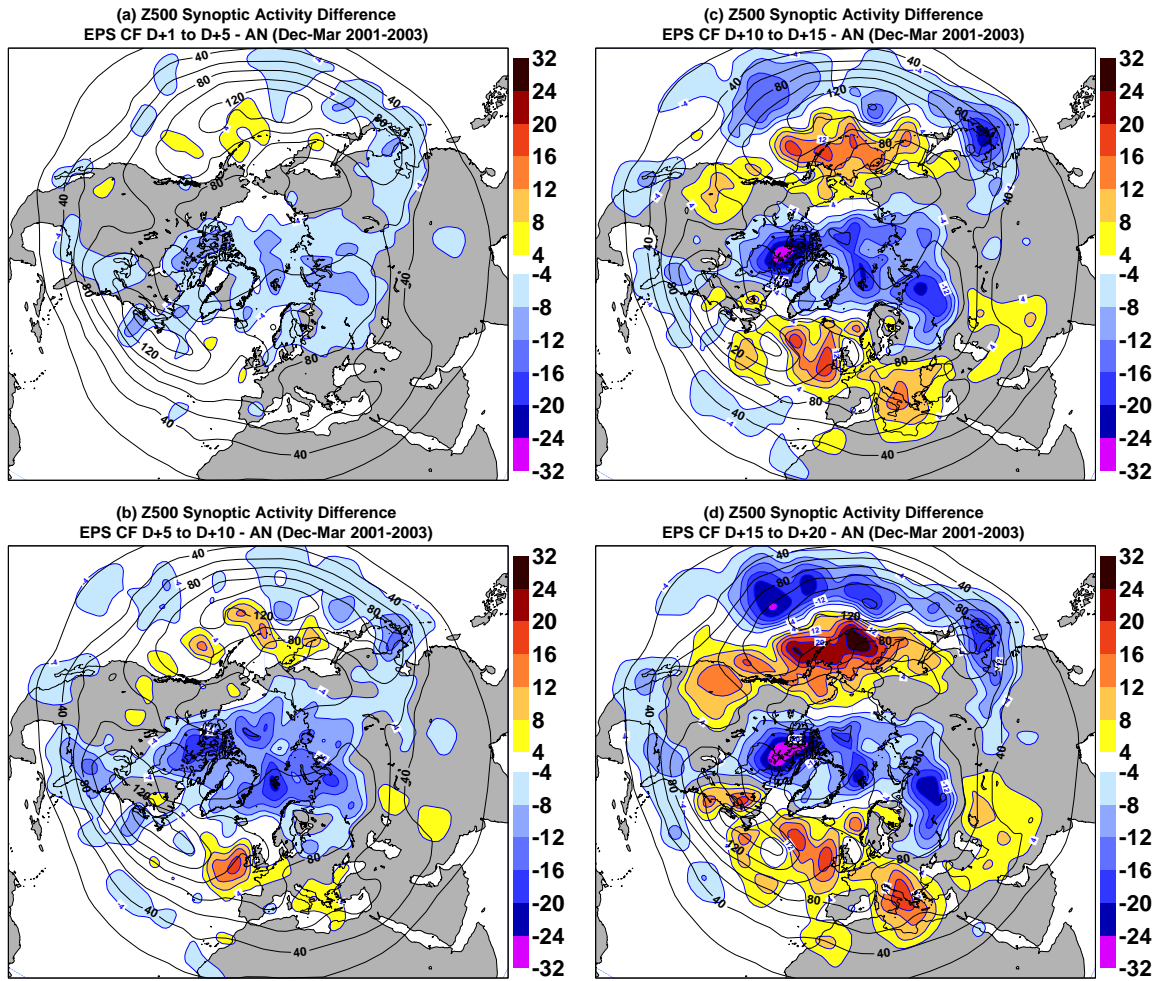
Figure 41: Same as in Fig. 40c, except for EPS control forecasts: (a) D+0 to D+5, (b) D+5 to D+10, (c) D+10 to D+15, and (d) D+15 to D+20.

3.7 Atmospheric Blocking

In this section the ability of the ECMWF model to simulate blocking activity is described. The capability of the model to capture observed blocking statistics is crucial given the importance of blocking events for local weather conditions particularly in the Euro-Atlantic region. A well-known shortcoming of many atmospheric circulation models, including the ECMWF model, is that they underestimate the frequency of occurrence of blocking (D'Andrea et al., 1998). In this section the latest ECMWF model cycle (26r1) is used to study the model's ability to simulate blocking. Furthermore, the growth of systematic blocking errors throughout the first 30 days of the integration is decribed. Finally, the question is addressed how systematic blocking errors have changed in recent decades.

In order to objectively describe atmospheric blocking the procedure described by Tibaldi and Molteni (1990) is used. First, 500 hPa fields are interpolated onto a 4° by 4° regular latitude-longitude grid. Then the geopotential height gradients GHGS and GHGN are computed for every longitude point on the grid:

$$GHGS \quad = \quad \frac{Z(\phi_o) - Z(\phi_s)}{(\phi_o - \phi_s)} \tag{8}$$

$$GHGN \quad = \quad \frac{Z(\phi_n) - Z(\phi_o)}{(\phi_n - \phi_o)}, \tag{9}$$

where $\phi_n = 80^oN + \Delta$, $\phi_o = 60^oN + \Delta$, $\phi_s = 40^oN + \Delta$, and $\Delta = -4^o, 0^o, +4^o$. A given longitude is defined as being instantaneously blocked if the following criteria are satisfied for at least one value of Δ:

$$GHGS \quad > \quad 0 \tag{10}$$
$$GHGN \quad < \quad -10 \quad \text{m/deg lat.} \tag{11}$$

Note, that for $GHGS > 0$ geostrophic winds in the midlatitudes are easterly. The second criterion eliminates most cases with large southward displacements of the westerly jets.

The above procedure is used to describe *instantaneous* blocking for every day and every longitude point on the grid. In order to identify blocking *episodes* for specific sectors the following technique is used: A blocking episode is defined if at least 3 neighbouring longitude grid points in a sector are instantaneously blocked for at least 4 consecutive days. This condition has been slightly relaxed by allowing blocking episodes to be "interrupted" by single days with less than 3 neighbouring longitude grid points blocked.

Recently, Pelly and Hoskins (2003) have proposed an improved technique to determine atmospheric blocking. This new technique, which is based on a reversal of the south-north gradient of potential temperature at the tropopause level, seems to be superior to the technique proposed by Tibaldi and Molteni (1990) in the Pacific region. This can in part be explained by allowing for longitudinal variations in blocking latitude in the formalism by Pelly and Hoskins (2003). Despite of the advantages of the Pelly and Hoskins (2003) scheme, here we use the technique by Tibaldi and Molteni (1990). The reason for this is that in contrast to potential temperature (at the PV2 level), long time series of Z500 are available for the operational model. Given the weaknesses of the Tibaldi and Molteni (1990) scheme in the North Pacific section, the focus of this study is on Euro-Atlantic blocking.
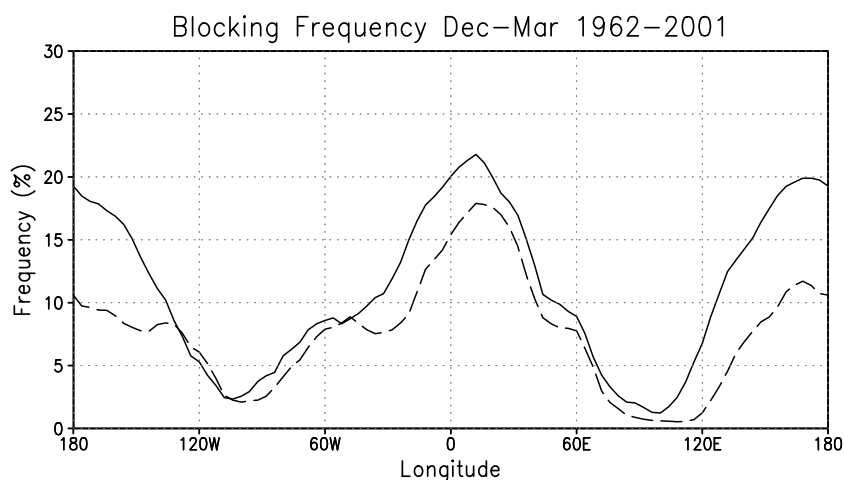
3.7.1 Systematic Blocking Errors

Figure 42: Frequency of wintertime days (Dec–Mar) that were classified as being blocked (%) for ERA-40 data (solid) and model cycle 26r1 (dashed). Results are based on the period 1962–2001. Details on the definition of instantaneous blocking are given in the text.

Table 5: *Total number of wintertime (Dec–Mar) Euro-Atlantic blocking episodes (n_{epi}) during the period 1962–2001 for ERA-40 data and the ECMWF model (cy26r1) together with the percentage of blocking episode lengths (in % of the total number of blocking episodes)] for six different length classes (4–6, 7–10, 11–15, 16–20, 21–30 and >30 days). See text for details on how blocking episodes are defined.*

	n_{epi}	4–6	7–10	11–15	16–20	21–30	>30
ERA-40	205	37.3	23.7	19.2	9.6	7.9	2.3
Model	188	34.9	31.4	15.7	9.9	7.6	0.6

Instantaneous blocking frequencies during winter (Dec–Mar) are shown in Fig. 42 for ERA-40 data (solid) and the seasonal integrations of the ECMWF model (cy26r1, dashed). In agreement with previous studies (e.g. Tibaldi and Molteni, 1990; Tibaldi et al., 1997) two areas with enhanced blocking frequency are evident from the ERA-40 data, that is, the North Pacific and the Euro-Atlantic region. In the North Pacific sector the model dramatically underestimates instantaneous blocking frequencies; the number of blocked days in the model is only half those in the ERA-40 data. In the Euro-Atlantic sector the model also underestimates instantaneous blocking frequencies. In this area, however, the underestimation is by far much less dramatic than it is in the North Pacific. The strong underestimation of instantanous blocking frequencies in the North Pacific is consistent with a zonal model bias in the high-latitude North Pacific (see Fig. 23) in which blocking is diagnosed by the Tibaldi and Molteni (1990) scheme. Given the pronounced systematic Z500 error in the North Pacific one could further expect the blocking frequency bias to be sensitively dependent on the central blocking latitude used for diagnosis.

Next, the capability of the ECMWF model to simulate blocking episodes during the winter season in the Euro-Atlantic sector is described. Here, the Euro-Atlantic sector is defined to extend from 40^oW to 52^oE. Some episode statistics are summarized in Tab. 5. For the ERA-40 data a total of 205 Euro-Atlantic blocking episodes have been identified during the forty year period 1962–2001. For the ECWMF model the number amounts to 188, that is, the number of simulated blocking episodes is lower than observed values by about 8%. The length distribution of wintertime Euro-Atlantic blocking episodes is relatively well simulated. In particular the model appears to be capable of producing persistent episodes of Euro-Atlantic blocking. Largest discrepancies in the frequency of occurrence are found for episodes lasting 7–10 days and more than 30 days.
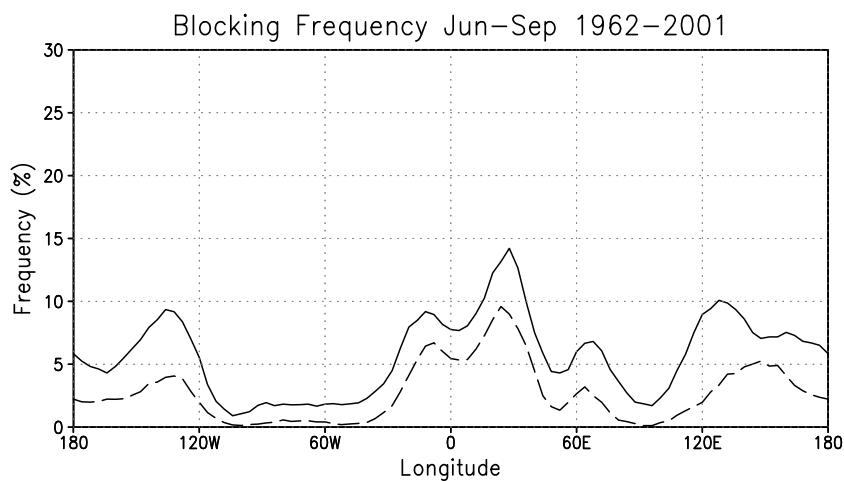

Figure 43: *Same as in Fig. 42, except for the summer season (June-September).*

	n_{epi}	4–6	7–10	11–15	16–20	21–30	>30
ERA-40	154	46.7	32.9	15.3	2.9	2.2	0.0
Model	93	51.9	29.9	11.7	5.2	1.3	0.0

Instantaneous blocking frequencies for the summer season (June–September) are shown in Fig. 43 for ERA-40 data (solid) and the ECMWF model (dashed). Consistent with previous studies (e.g. D'Andrea et al., 1998, their Fig. 2), the blocking frequency in summer is lower than that during winter, that is, the model captures the observed annual cycle of blocking activity. Compared to the reanalysis, however, the simulated blocking frequency is underestimated at all longitudes during the summer season. Again, this is consistent with summertime systematic Z500 errors (Fig. 23b) suggesting a too zonal flow at high-latitudes for amost all longitudes.

The number of blocking episodes is considerably underestimated by the ECMWF during summer (Tab. 6). Whereas a total of 154 blocking episodes were observed during the last 40 years, the model simulates only 93 blocking episodes for the same 40-year period. Moreover, the model underestimates the frequency of occurrence of long-lasting Euro-Atlantic blocking episodes in summer.

3.7.2 Growth of Blocking Errors

In the previous section systematic blocking errors have been described. In this section the focus is on how errors in blocking frequency grow during the first 30 days of the integration. To this end, hindcast integration with the ECMWF Monthly Forecasting system for the period 1990–2002 were diagnosed. For both, the summer and winter season, a total of 108 30-days forecasts (each forecast comprising one control forecast and four ensemble members) have been evaluated. ERA-40 data have been used for verification.

The growth of the model's systematic error in simulating instantaneous wintertime blocking from D+1 to D+30 is shown in Fig. 44a (blue means that the model underestimates the blocking frequency). In the Euro-Atlantic region most of the systematic blocking error develops during the first 10 days of the integration. There is some indication that systematic errors grow a little faster over Europe than they do over the North Atlantic (the significance of this difference is questionable, though). In the North Pacific region systematic blocking errors grow rapidly during the 5 days or so. Thereafter, the growth rate is slightly reduced (note the nonlinear contour scale). That is, systematic model errors continue to grow well beyond the medium-range.

During summertime the growth of blocking differences between the model and ERA-40 during the first 30 days of the integration are less conclusive (Fig. 44b). Despite of the relatively large sample size differences are rather noisy. As a result no clear growth signature can be inferred from the differences between simulated and observed blocking frequencies. It is apparent, though, that differences appear to develop quite quickly throughout the first 5 days or so and remain more or less constant thereafter.

3.7.3 Evolution of Blocking Errors

Instantaneous local blocking frequencies for the three periods 1986–88, 1993–95, and 2001–03 are shown in Fig. 45 for D+10 forecasts and verifying operational analyses. The operational model cycles in the mid-1980s underestimated the blocking frequency at D+10 considerably (Fig. 45a), particularly in the North Pacific region. The Euro-Atlantic blocking maximum is also underestimated and shifted towards the east. This "eastward shift"
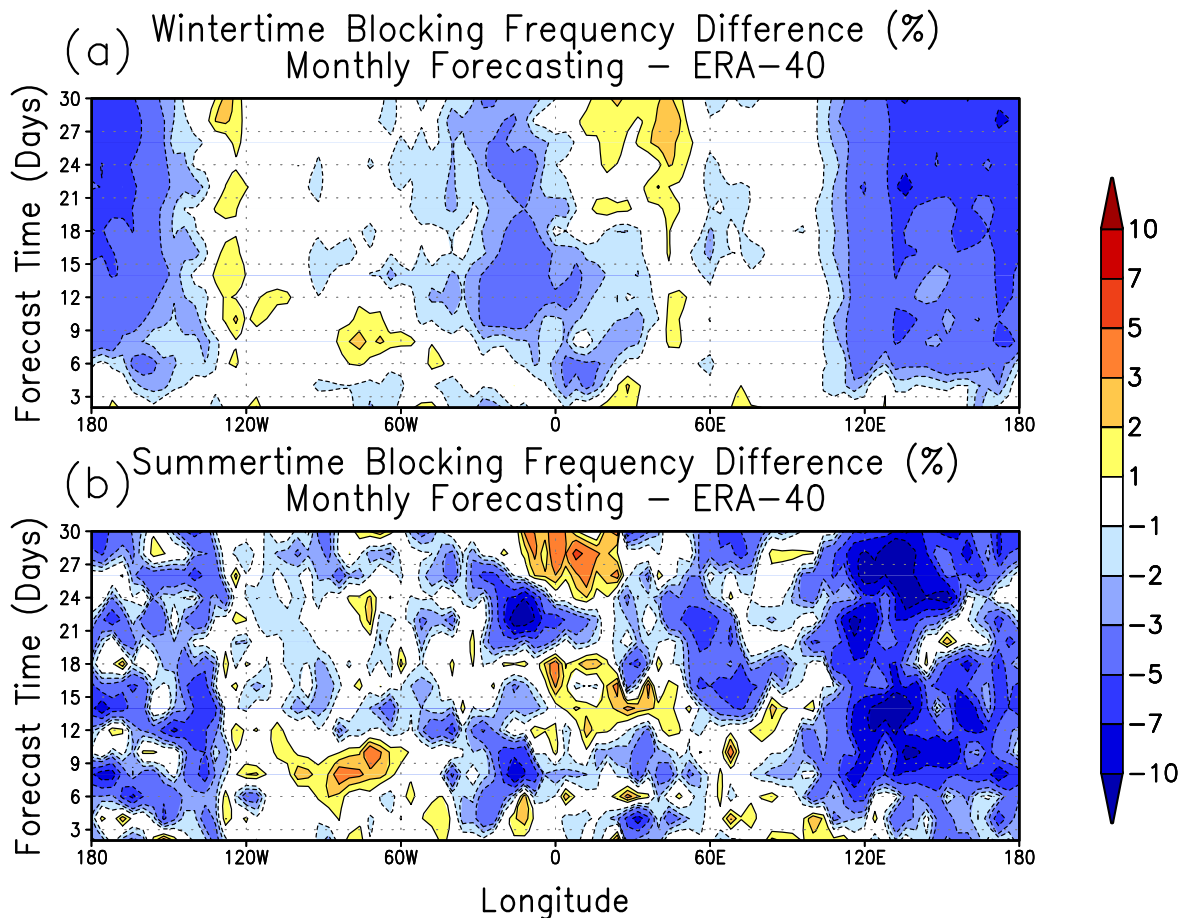
Figure 44: Difference (a) wintertime and (b) summertime instantaneous blocking frequency (%) between monthly forecasts and ERA-40 reanalysis data. Negative values indicate that the model underestimates blocking events. Each of the above figures is based on 108 monthly forecasts.

is largely due to underestimation of the blocking frequency over the eastern North Atlantic, a feature that is also seen in model cycle 26r1 (cf. Fig. 42). In the early 1990s the model's capability to simulate North Pacific and Euro-Atlantic blocking has been improved significantly (Fig. 45b); the observed pronounced two-peak structure of atmospheric blocking is captured. Systematic blocking errors for the period 2001–03 seem to be larger than those seen during the mid-1990s.

Differences between the two periods 1993–95 and 2001–03 have to be interpreted carefully because the observed blocking frequency is subject to pronounced interannual variations. During the mid-1990s Euro-Atlantic blocking was most pronounced over the European continent, that is, in an area where most models are more skilful in simulating atmospheric blocking than over the North Atlantic (D'Andrea et al., 1998). In contrast the period 2001–03 was marked by a high frequency of occurrence of North Atlantic blocking events. In summary, the above results show that the model performance in simulating atmospheric blocking has been improved notably after the early 1980s, particularly in the North Pacific region. It is less clear, however, whether model changes during the last 10 years have led to significant improvements of the ECMWF model's simulation of atmospheric blocking.
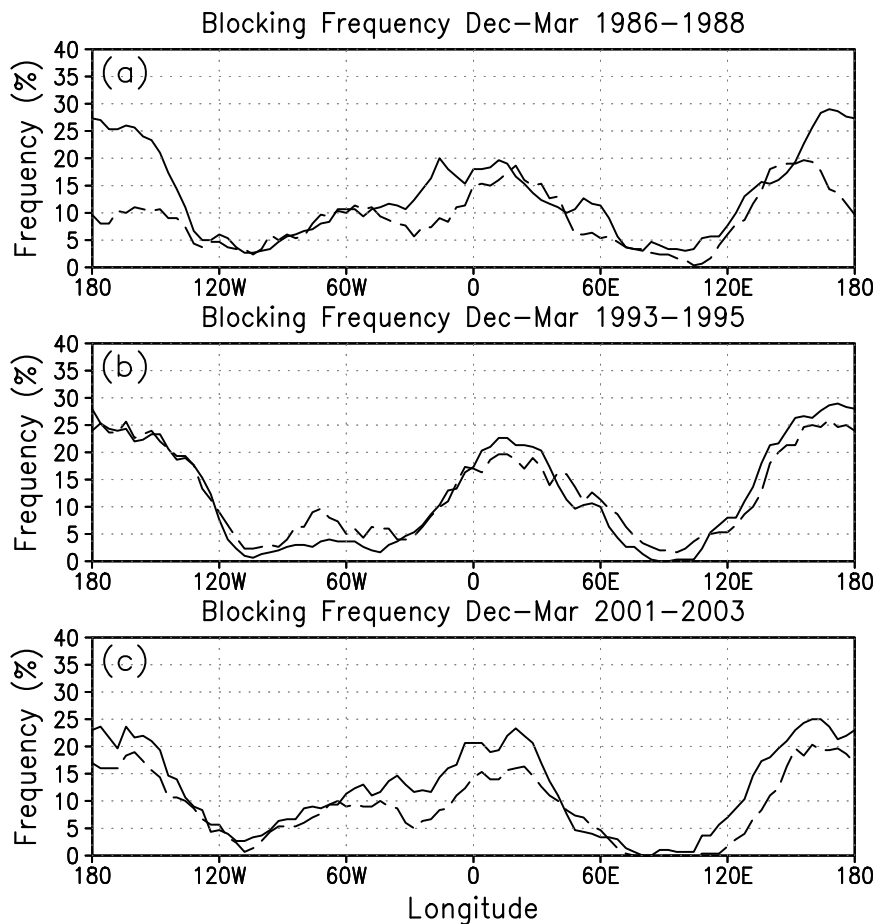
Figure 45: Instantaneous blocking frequencies (%) during winter (Dec–Mar) for the-then operational analysis (solid) and D+10 forecast (dashed): (a) 1985/86–1987/88, (b) 1992/93–1994/95, and (c) 2000/01–2002/03.

3.8 Madden-and-Julian Oscillation

The performance of the ECMWF forecast model in simulating the MJO is investigated using six month integrations starting on 1 October (extended winter season) and 1 April (extended summer season) for the period 1962–2001. Notice that, in contrast to previous sections, the whole 6 months of the integration are used for diagnosis. Whereas this increases the length of the time series and therefore the significance of the results, it should be mentioned that at least the first days of these integrations are still affected by the initial conditions. Subsequent results are based on model cycle 26r1 with a resolution of T_L95L60. To assess the model performance ERA-40 reanalysis data are used for comparison. Given that the dominant time scale of the MJO is 30–60 days, forty years of 6 months integrations should contain roughly 180 realizations of the MJO[8].

In this study the velocity potential at pressure level 200hPa (VPOT, hereafter) is used to characterized the MJO. At this level VPOT, describes the large-scale divergent part of the upper-tropospheric flow. The focus is on the Tropics. To this end VPOT fields have been averaged from 5^oS to 5^oN. Daily data were smoothed by forming 4 day non-overlapping averages. Thus, each extended season consists of a total of 46 4-day mean fields. Finally,

---

[8]However, since the strength of MJO activity undergoes low-frequency changes, as shown below, this estimate is likely to be too optimistic.
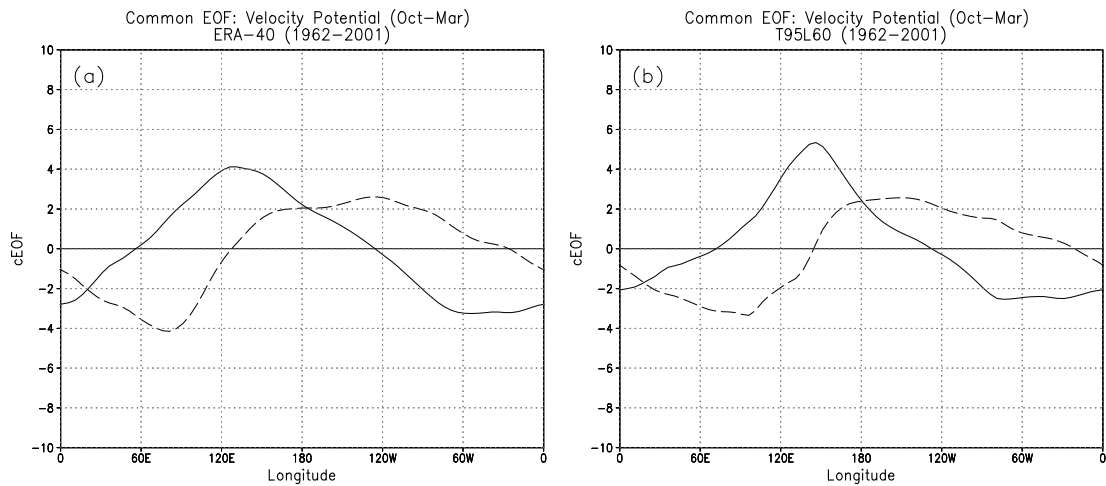
Figure 46: First (solid) and second (dashed) EOF of wintertime (Oct–Mar) tropical velocity potential anomalies ($m^2 s^{-1}$) based on the period 1962–2001: (a) ERA-40 reanalysis data and (b) $T_L 95L60$ integrations based on model cycle 26r1.

prior to diagnosis, the mean annual cycle has been removed.

In the first part of this section the focus is on the extended winter season and in the second part the model's performance in simulating the MJO during the extended summer season is described. Then, the sensitivity of MJO variance to horizontal resolution is described.

3.8.1 Extended Winter Season

The first two Empirical Orthogonal Functions (EOFs) of wintertime tropical VPOT anomalies are shown in Fig. 46 for ERA-40 and the ECMWF model. The first and second EOFs explain 37.8% and 28.8% of the total variance for the ERA-40 data, respectively, and 33.4% and 24.1% for the ECMWF model. The first two EOFs show wavenumber one type structures which are reminiscent of the MJO. The first EOF describes out-of-phase anomalies between the warm pool and the tropical Atlantic, whereas the second EOF shows anomalies of opposite polarity over the Indian and Pacific ocean. The first two EOFs are in quadrature. The most important result in the context of the present study is, though, that the structure of the leading two EOFs is very similar for both datasets. This suggests that the ECMWF model captures the dominant modes of intraseasonal tropical VPOT variability.

In order to further investigate the characteristics of the leading two eigenmodes of tropical VPOT anomalies and to allow for a direct comparison between the model results and ERA-40 data, EOF analysis has been performed for the combined data set. That is, the first $46 \cdot 40$ vectors of the combined dataset are ERA-40 anomalies and the last $46 \cdot 40$ vectors are formed by simulated anomalies. By performing EOF analysis for the combined data set principal axes can be found that are *common* to both datasets. For this reason we shall call this technique common EOF analysis (cEOF analysis) hereafter (see Barnett et al., 1999, for details). The first two cEOFs are shown in Fig. 47. They are very similar to the EOFs obtained by applying conventional EOF analysis to both datasets separately.

As noted above, the MJO can be characterized by tropical VPOT anomalies of wavenumber one. Moreover, these anomalies tend to propagate eastward. Therefore, for the MJO to be efficiently described by the leading two EOFs, the first two principal components (PCs, hereafter) should be in quadrature—with PC1 leading
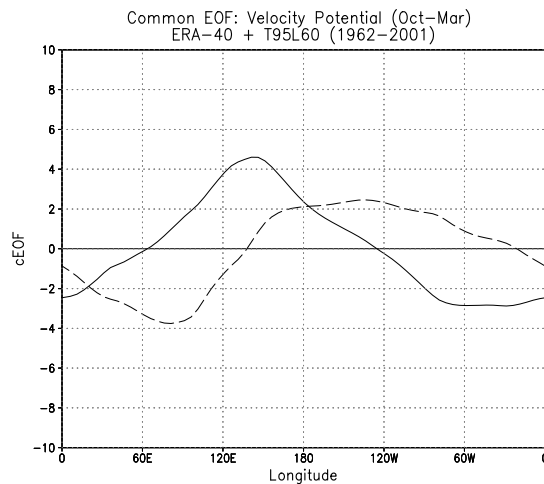
Figure 47: Same as in Fig. 46, except for the combined ERA-40 reanalysis and $T_L 95L60$ model dataset.

PC2 by 90°. The squared coherency and phase spectrum between the first two cPCs are shown in Fig. 48 for the ERA-40 (solid) and model (dashed) data. The cPCs were obtained by separately projecting observed and modelled anomalies onto the leading two cEOFs. Significant squared coherency between the leading two cPCs is found for ERA-40 data on timescales from 20–60 days. Furthermore, cPC1 leads cPC2 by about 90° confirming that the leading two cEOFs describe MJO variability. For the ECWMF model, however, significant squared coherency is evident on timescales from 20–30 days only, that is, the leading two cPCs are much less coherent than the observations on time scales longer than 30 days. This confirms previous (subjective) experience with the ECMWF model that the model underestimates MJO-related *slowly* eastward propagating VPOT anomalies.

Next, the model's capability to simulate the strength of the MJO has been investigated. To this end partial eigenvalues (pEVs) have been computed for each of the two datasets and every winter during the period 1962–2001 (Fig. 49). The pEVs were obtained by projecting VPOT anomalies for every winter onto the normalized leading
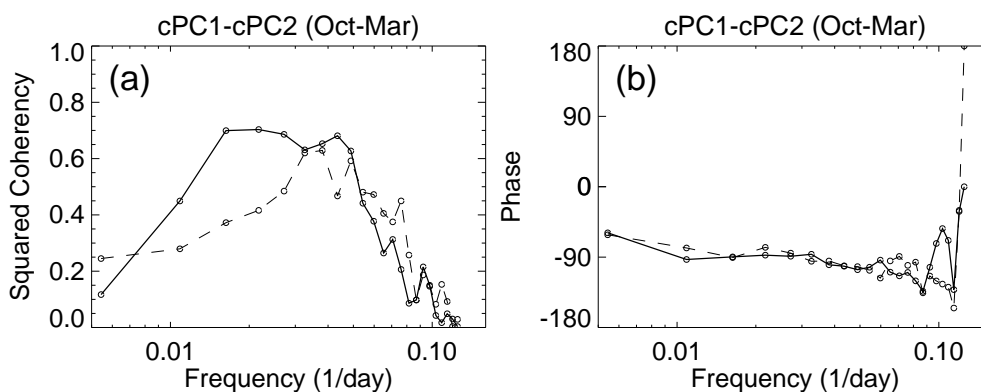

Figure 48: (a) Squared coherency and (b) phase spectrum for the first two common PCs of wintertime tropical velocity potential anomalies from ERA-40 data. Smoothing has been performed by averaging over all 40 winters (i.e., equivalent degrees of freedom amount to 80). The squared coherency spectrum shows how different frequencies contribute to the squared correlation coefficient. The first PC leads the second PC in (b) for negative phases. Methodological details can be found, for example, in Jenkins and Watts (1968).
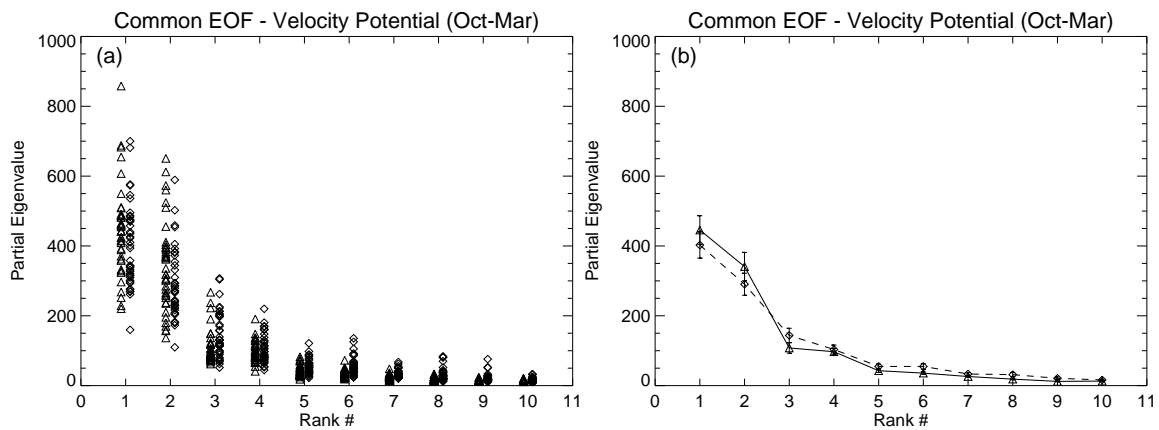
Figure 49: (a) Partial eigenvalues for ERA-40 reanalysis data (triangles) and $T_L 95L60$ integrations based on model cycle 26r1 (diamonds) for the leading 10 common EOF modes of wintertime (Oct–Mar) tropical velocity potential anomalies (m^2/s) based on the period 1962–2001. For each winter one triangle (diamond) is shown. (b) Average and 95% confidence intervals for the partial eigenvalues shown in (a).

cEOFs and by computing the variance of the projection coefficients (cPCs) for each winter separately. Thus, high (low) values suggest high (low) activity along the direction of the respective cEOF. Firstly, it is evident that MJO activity, as given by the pEVs for the first two cEOFs, shows pronounced interannual variability. In fact individual pEVs for the first two eigenmodes differ by as much as a factor of 4 between different years. This clearly shows that relatively long time series are necessary in order to draw meaningful conclusions about the model's capability to simulate MJO activity. Average pEVs are shown in Fig. 49b along with 95% confidence levels. From these statistics it is evident that MJO activity in the ECMWF model is underestimated by about 10% compared to the analyses, and that the model overestimates activity along the direction of cEOF mode 2–10 (These modes describe VPOT variability at higher wavenumbers).

Finally, the spectral characteristic of the observed and modelled MJO are studied. Average power spectra and 95% confidence intervals are shown in Fig. 50 for the first two cPCs based on ERA-40 data (solid) and the 6 month model integrations (dashed). The observed MJO clearly shows a broad-band peak on intraseasonal time scales (30–60 days). In contrast, the model spectra show no broad-band peaks; they rather resemble those
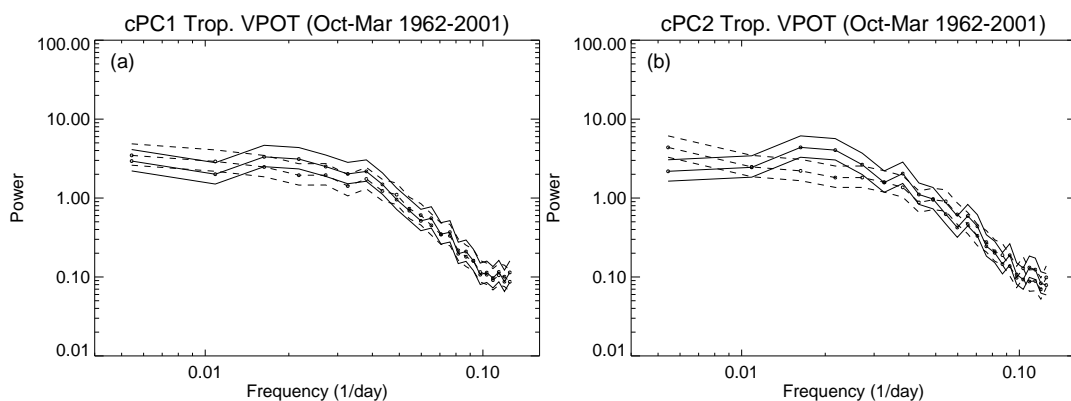

Figure 50: Smoothed power spectrum of the (a) first and (b) second PC of wintertime tropical velocity potential anomalies based on ERA-40 data (solid) and $T_L 95L60$ integrations based on model cycle 26r1 (dashed). Thin lines denote 95% confidence levels. Details on the method are given, for example, by von Storch and Zwiers (1999).
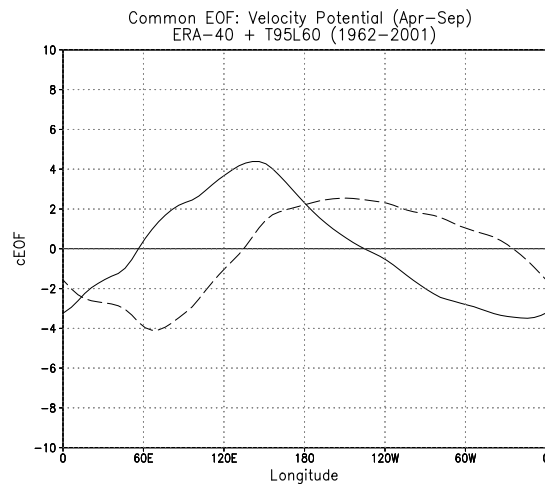
Figure 51: Same as in Fig. 47, except for the extended summer season (April–September).

of first-order auto-regressive processes. The underestimation of MJO variability of the model is particularly pronounced (and statistically significant) on longer timescales (i.e., around 60 days.)

3.8.2   Extended Summer Season

In this section the focus is on the model's performance to simulate the MJO during the extended summer period April–September.

First, common EOF analysis has been applied to combined topical VPOT anomalies from ERA-40 and the model (Fig. 51). The leading two cEOFs, which describe MJO variability and explain 34.9% and 29.0%, respectively, of the total variance, are very similar to those for the extended winter season (Fig. 47).

Cross-spectral characteristics between the leading two cPCs for ERA-40 (solid) and the ECWMF model (dashed) are shown in Fig. 52. As for the winter season the model shows realistic eastward propagating MJO-type VPOT anomalies on timescales from 20 to 30 days. The coherence of slowly (30–60 days) propagating features in the model, however, is strongly underestimated.
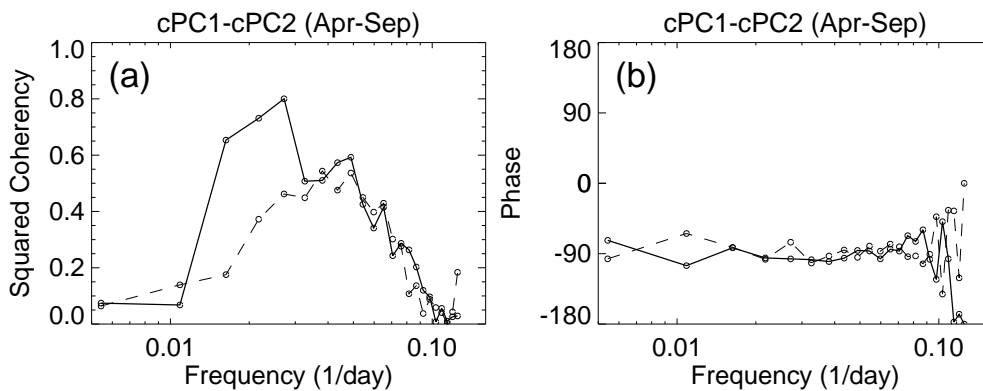

Figure 52: Same as in Fig. 48, except for the extended summer season April–September.
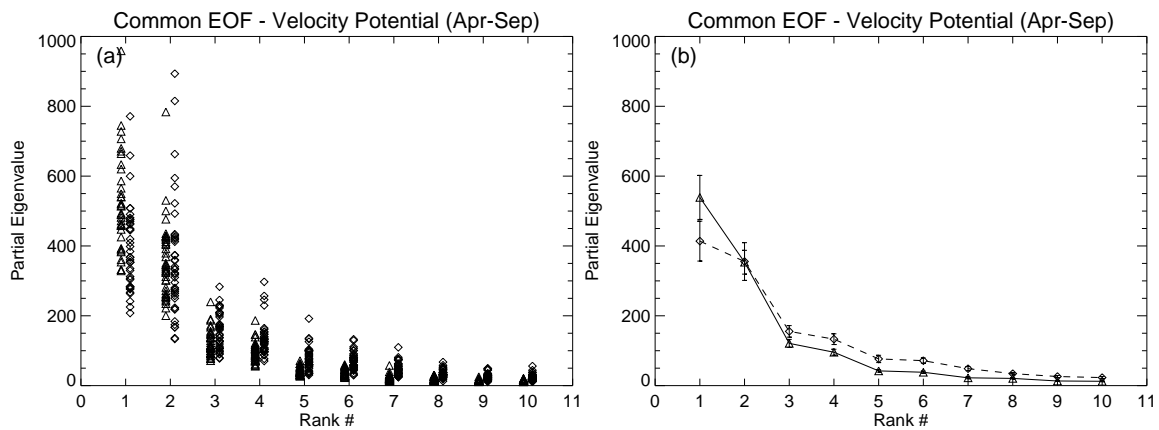
Figure 53: Same as in Fig. 49, except for the extended summer season April–September.

MJO activity during the extended summer season, expressed in terms of pEVs, in ERA-40 and the ECMWF model is summarized in Fig. 53. The model significantly underestimates MJO variability along the direction of cEOF1 which captures large-scale divergence anomalies over the Maritime Continent.

Estimated power spectra for the first two cPCs are shown in Fig. 54 for ERA-40 and the ECMWF model. As for the winter season, the observed broad-band spectra peak around 30–60 days are absent in the ECMWF model, suggesting that there is no quasi-oscillatory behaviour of simulated MJO.

In summary, the main model deficiencies in terms of MJO variability are very much alike in winter and summer. They encompass the failure to *slowly* propagate MJO-related VPOT anomalies eastward. While there is some eastward propagation of wavenumber-one VPOT anomalies in the model, it is only at the high-frequency end of the MJO spectrum (20–30 days) that the model is consistent with the ERA-40 data. Another significant model deficiency is the absence of any quasi-oscillatory behaviour of the modelled MJO.
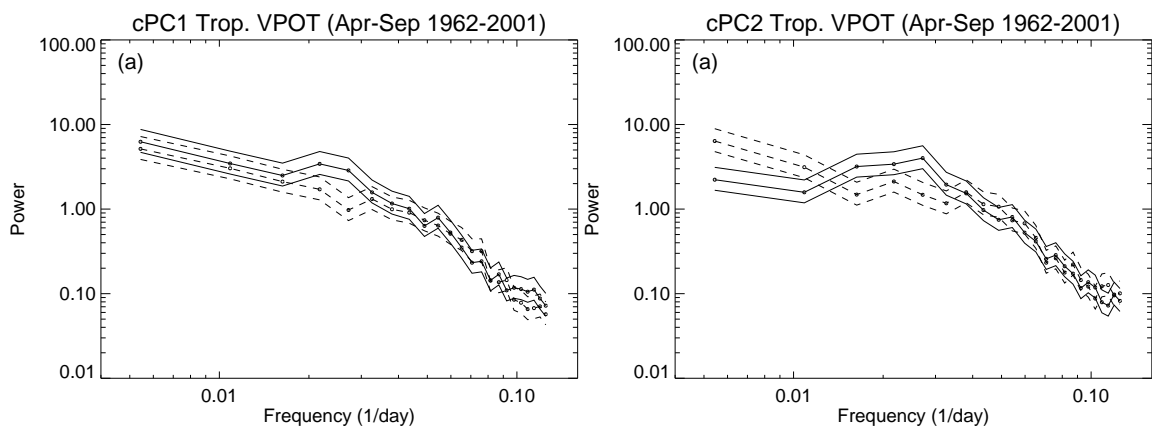

Figure 54: Same as in Fig. 50, except for the extended summer season April–September.
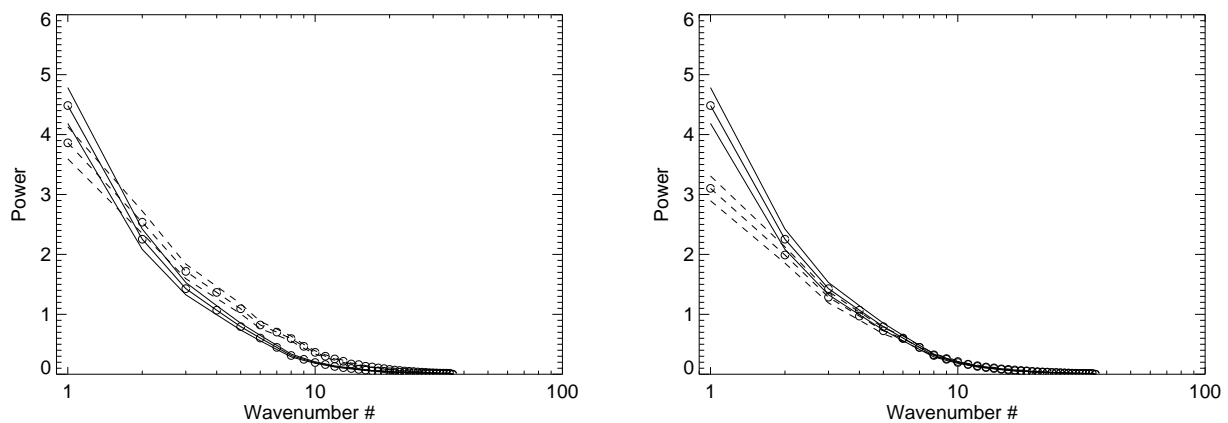
Figure 55: Analysed (solid) and simulated (dashed) wavenumber spectra for summertime tropical VPOT anomalies: (left) T_L95L60 and (right) $T_L159L60$. Also shown are 95% confidence intervals. Results are based on 6 months integrations over the period 1991–98 based on model cycle 25r1.

3.8.3 Influence of Horizontal Resolution

The capability of the ECMWF model to simulate the MJO during the first 30 days of the integration is described by Vitart (2003). There it is found that MJO variance decreases during the first 10 days by as much as 50%. This is contrast to the model integrations described above, where the loss in variance amounts to only about 10%. There are several possible reasons for these differences. Monthly forecasts are based on a coupled atmosphere-ocean model and the drift of the coupled model system might deteriorate the MJO. Furthermore, most of the monthly forecasts were performed at a different horizontal (T_L159) and vertical (L40) resolution using older model cycles. Therefore, it might well be possible that the MJO is sensitive to coupled model drift, resolution or model formulation (model cycles).

In order to test the sensitivity to horizontal resolution, six-month integrations for the summer and winter seasons of the years 1991–1998 have been performed using model cycle 25r1[9] at two different resolutions (T_L95L60 versus $T_L159L60$). Average wavenumber spectra of summertime tropical VPOT anomalies from ERA-40 and the ECMWF model are shown in Fig. 55 for the T_L95L60 (left) and the $T_L159L60$ (right) model. Results for the winter season are very similar (not shown). Consistent with the results for the Monthly Forecasting system the relatively high-resolution ($T_L159L60$) model underestimates MJO variance by about 30%. The lower-resolution model (T_L95L60), on the other hand, shows much more realistic values of variance at wavenumber one. This result clearly indicates a large sensitivity of MJO variance with respect to horizontal resolution.

3.8.4 Medium-Range Forecasts

As shown in the previous section the $T_L159L60$ model shows a larger loss of MJO variance than the T_L95L60 (30% versus 10%). It is interesting to address the question whether models at horizontal resolutions higher than T_L159 show a similar behaviour as the $T_L159L60$ model. To this end operational medium-range forecasts at $T_L255L60$ (EPS control) and $T_L511L60$ (deterministic forecast) of tropical VPOT anomalies were diagnosed. Average wavenumber spectra of tropical VPOT anomalies for the period 1 June 2002 to 21 August 2002 are shown in Fig. 56 for the operational analysis (solid) and deterministic (left) as well as EPS control (right)

---

[9]The integrations with model cycle 25r1 were originally carried out to decide which horizontal resolution to use for the long runs with the later model cycle 26r1 over the whole ERA-40 period.
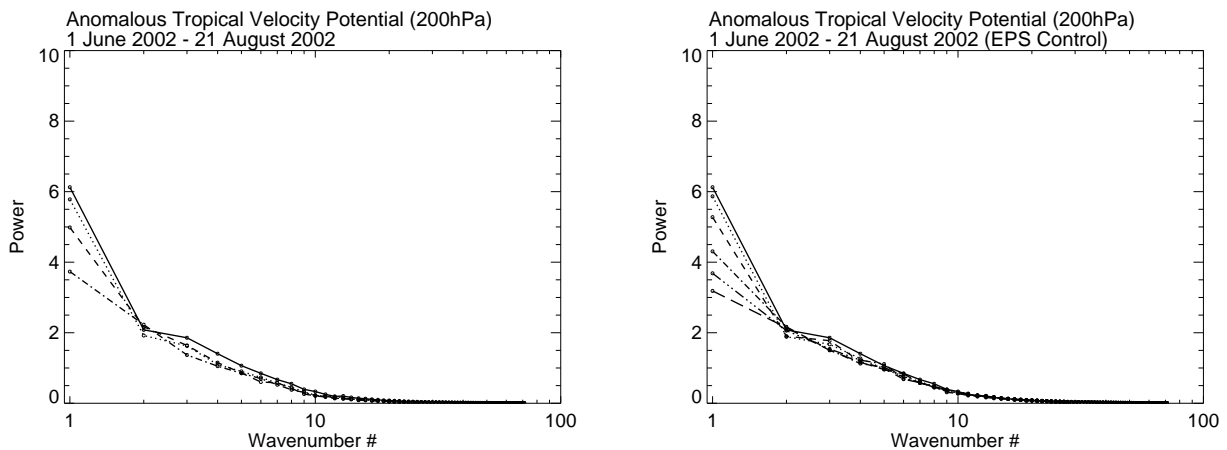
Anomalous Tropical Velocity Potential (200hPa)
1 June 2002 - 21 August 2002

Anomalous Tropical Velocity Potential (200hPa)
1 June 2002 - 21 August 2002 (EPS Control)

Figure 56: As in Fig. 55, except for operational (left) high-resolution deterministic ($T_L 511$) and (right) EPS control forecasts ($T_L 255$). Shown estimates based on D+0 (solid, analysis), D+2 (dotted), D+5 (dashed), D+10 (dash-dotted), D+15 (dash-dot-dot-dot), and D+20 (long-dashed) forecasts. Results are for the summer period 1 June 2002 to 21 August 2002.

forecasts. At both resolutions a large loss of MJO-type variance can be found during the first 10 days of the integration (about 30 %). The EPS control forecasts show that this loss of variance continues into the near extended-range (10–20 days) where the simulated variance amounts to about 50% of the values for the analyses.

4 Summary and Discussion

Systematic errors in the ECMWF model have been described using a wide range of operational forecast products and a series of research experiments. It has been shown that for most parameters asymptotic systematic model errors are still substantial and their spatial structure has changed rather little over the years. For most of the parameters considered in this study the continuous improvement of the model in recent years is primarily reflected in lower values in the short-range and reduced rates at which systematic errors tend to grow. As a consequence most of the systematic errors in the medium-range are much smaller than they used to be several years ago.

Another feature of systematic errors in the ECMWF model found in this study is that for some parameters and regions the spatial structure of systematic error undergoes substantial changes during the course of their growth in the medium-range and extended-range. The simplest explanation for this characteristics is presumably that systematic errors in the medium-range and extended-range are partly the response to remote model deficits.

Atmospheric Circulation The most pronounced systematic error of the atmospheric circulation occurs during wintertime in the central North Pacific region and shows up in form of an anticyclonic bias. The reason for this bias is not yet fully understood. There is some indication, though, that at least part of its origin is over the Maritime Continent and the Indian ocean where the model tends to underestimate upper tropospheric large-scale divergence and rainfall. Support for the hypothesis that the origin lies in the Tropics comes from asymptotic systematic errors of the latest ECMWF model in simulation the wintertime 200 hPa streamfunction field (Fig. 57). The North Pacific bias seems to be part of a wave train emanating from the western Tropical Pacific and Maritime Continent (see Sardeshmukh and Hoskins, 1988; Neale and Slingo, 2003, for a physi-

(a) 200hPa Streamfunction ERA40 DJFM 1962-2001

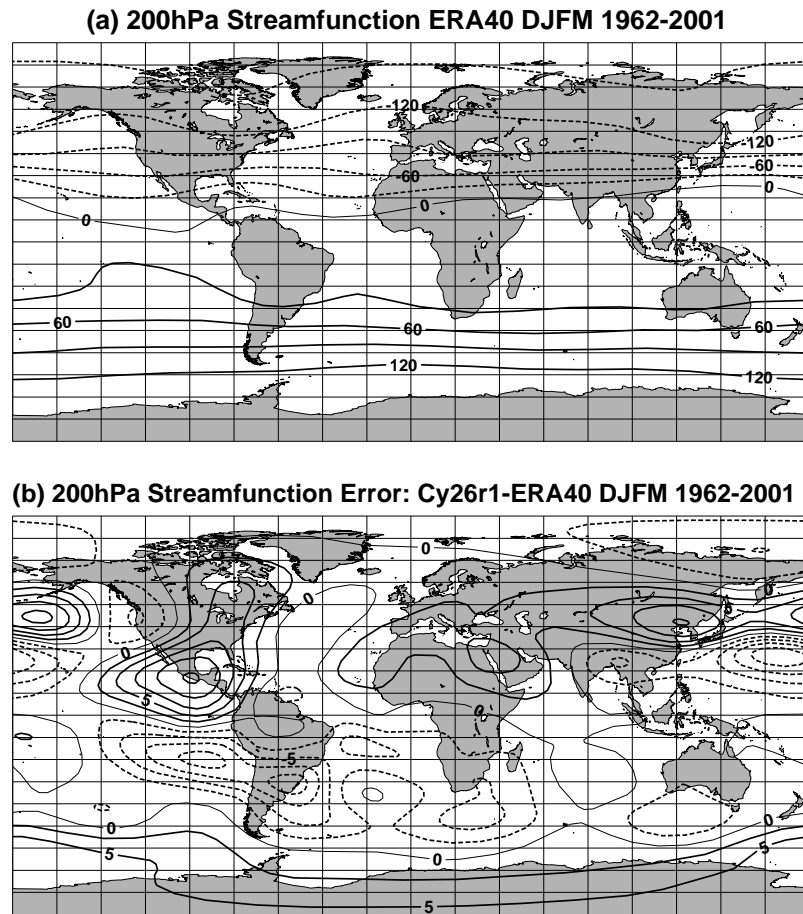### (b) 200hPa Streamfunction Error: Cy26r1-ERA40 DJFM 1962-2001

Figure 57:

cal background). The fact that biases in the North Pacific flow are almost orthogonal in the short-range and extended-range suggests indeed that the origin is remote rather than local. One way to shed further light on the nature of the North Pacific circulation biases would be to perform seasonal integrations with the Tropics being relaxed to the analysis (see Ferranti et al., 1990, for details on the method). If the error would still develop despite of the Tropics being relaxed to the analysis then this would be an indication that the error is locally forced.

One possible way to diagnose systematic errors of the atmospheric circulation (and other parameters) in further detail is to use the adjoint technique. Preliminary work along this line has been carried out by applying the new forcing sensitivity suite to D+2 forecast errors (Barkmeijer et al., 2003). The forcing sensitivity method enables the determination of those model tendency perturbations which, when held constant over the optimization period, reduce D+2 forecast errors in some optimal sense. By adding the averaged optimal tendency perturbations during nonlinear integrations of the ECMWF it has been shown that systematic Z500 errors can be reduced by a factor of two or so. Therefore, by further diagnosing the tendency perturbations obtained with the adjoint technique it might be possible to obtain a deeper understanding of their possible origin. Moreover, the averaged tendency perturbations might be used in an empirical way to reduce systematic model errors. Further research will be carried out along this line.

It is worth mentioning that Northern Hemisphere asymptotic systematic Z500 errors as simulated by the

ECMWF model bear resemblance with those evident for a wide range of atmospheric models. This can be seen by comparing Fig. 30 with the multi-model ensemble mean error for all models investigated in the Atmospheric Model Intercomparison Project (Gates et al., 1999, their Fig. 1).

Specific Humidity   The latest cycles of the ECMWF model show a tendency to be drier than the analysis in the Tropics. A comparison with total water vapour content estimates from SSM/I retrievals reveals that the model indeed has a dry bias in the tropical troposphere. The adjustment of the model is associated with excessive amounts of convective precipitation during the first couple of hours of the integration. This problem is a consequence of the high sensitivity of the ECMWF model to small humidity changes in lower parts of the tropical atmosphere. A reduction of the amount of boundary layer humidity by less than 2% largely eliminates the excessive amounts of convective rain during the first few hours of the integration (E. Holm, personal communication). Given that the accuracy of the observational mean values is likely to be less than 2%, it is difficult to constrain the humidity analysis to errors well below 2%. One pragmatic way, which may be followed at ECMWF, is to accept that the current observational accuracy is not high enough to constrain the humidity analysis in the Tropics sufficiently and to correct the humidity analysis in some well-balanced manner. Such a "correction" would at least avoid the occurrence of transient biases during the first few hours that are associated with the model's adjustment such as the warm bias in the tropical middle atmosphere that is associated with the excessive release of latent heat.

Moreover, it has been shown in this study that the Tropics in ERA-40 are slightly too moist compared to SSM/I data. This bias might be important when medium-range forecasts are started using ERA-40 as initial conditions and when ERA-40 data are used to study the hydrological cycle.

Cloud Related Parameters   A cursory examination of cloud and convection related systematic biases in the ECMWF climatology has been carried out using various retrieved fields from ISCCP, GPCP, ERBE and SSM/I for validation. The analysis was based on two ensembles of 4 month integrations of the IFS code at T_L95L60 resolution.

This experimental setup and the use of these data sets is justifiably open to criticism. The experiment is restricted to a mere two seasons, raising concerns about generality. Many of the retrieval products have been superceded. The ERBE products have been replaced with CERES-TRMM (e.g. Chevallier and Morcrette, 2000) and MODIS products. GPCP rainfall is now complemented by other algorithms from TRMM active radar and improved SSM/I retrievals. The C2 ISCCP product is superceded by a superior D2 product. The reason this experimental framework has persisted is that its numerical efficiency has allowed a useful database to be established, containing climate simulations for a number of major model cycles over the previous 7 years. The results here show that this simple framework is capable of revealing a number of systematic biases that were also identified in higher resolution short-term investigations. Work is now underway at ECMWF to organise a new climate-experiment framework with a modernized observational suite for future model validation and development.

The climatology of the reference model cycle 23r4, used for both ERA40 and the current operational seasonal forecast system, was quantified. Overall, the model was found to reproduce many features of the observations well, with the means of surface winds and total column water vapour closely reproducing the observations. Nevertheless, a number of systematic departures from observations were noted, which were found to agree with biases in the high resolution short-term operational forecasts documented in previous investigations. Briefly these were: too much liquid water in the subtropical clouds, a lack of stratocumulus, low cloud cover and too little ice in midlatitude winters, too much high cloud associated with deep convection which was optically too thick (inferred from the TOA net flux biases), and too little clouds over the summer Eurasian continent.

The improvements between cycle 23r4 and the current cycle 26r1 were then documented. While a number of the systematic biases remain, the encouraging result was the significant reduction of the subtropical overestimation of LWP and improved deep convective high cloud characteristics. These were accompanied by distinct enhancement in the representation of the TOA shortwave net fluxes. This was set in a longer term context by examining a series of experiments conducted at T63 resolution, using various model cycles from the past 7 years. These showed that overall, the latest model cycle possesses the best climate. They also revealed that in terms of the cloud LWP, the ERA40 cycle 23r4 used for the seasonal forecasts was furthest from observations. An attempt was made to quantify the improvement in model climatology due to modifications in physics, by comparing this to the gain achieved by increasing horizontal resolution from T63 to T255. This revealed surprisingly little impact of resolution on the climate statistics for the seasons examined. This lack of impact from horizontal resolution agrees with the other findings of this document. However, it should be remembered that each experiment was conducted using the "default mode" for each resolution, which implied different timesteps and convective closure timescales.

Finally, despite the significant improvements, the cycle 26r1 retains a number of problematic aspects in its climatology, not least the continued lack of stratocumulus clouds off the west coasts of the major continents. Additionally, cloud cover, and presumably ice content (inferred from the radiation budget) are still lacking in the midlatitude winters. What are the prospects therefore for tackling these biases in future? In terms of the stratocumulus regimes, a replacement for the current local mixing, dry boundary scheme is under development, based on a moist, non-local parcel model concept. Preliminary results with the new scheme are promising, and appear to reproduce observed stratocumulus regimes well. The model cloud scheme is also under development. The aim is twofold; to introduce a simple statistical framework to represent the subgrid fluctuations of humidity and their effect on cloud cover, and to combine this with a separate ice variable. The ability to represent the subgrid distributions of total water in different cloud regimes combined with an improved ice microphysics will hopefully allow the midlatitude climate to be improved without negatively impacting the representation of high ice clouds in the Tropics; an objective that has to date proved elusive.

Budget Diagnostics   A valuable technique to identify the origin of (short-range) systematic errors by investigating initial tendency errors has been suggested by Klinker and Sardeshmukh (1992). Using this technique, Klinker and Sardeshmukh (1992) were successful to identify deficiencies in the model's mechanical and thermal forcing. The same method has been used by Klinker and Ferranti (2000) to diagnose systematic model errors of a later model cycle (they diagnosed January 1999). Given that most of the systematic error structure remained more or less unchanged during the last couple of year it can be expected that most of the results described in Klinker and Ferranti (2000) are still valid. A natural extension of the above studies, however, would be take into account systematic humidity errors and systematic errors in the stratosphere. The code to carry out budget diagnostics has been recently implemented in the latest model cycle. Preliminary results for January 2003 (not shown) suggest, for example, that the positive temperature bias in the upper stratosphere (Fig. 2) is due to problems in the short-wave radiation code.

Stochastic Physics   One conclusion of the present study is that for most of the parameters the rate at which systematic errors grow has been reduced during the last couple of years by model improvements. The spatial structure of the systematic errors, however, has changed rather little. As pointed out by Palmer (2001), one way to explain the above characteristics could be that the methodology used to approximate the equations of motion (i.e., neglecting the variability of unresolved scaled) is itself a source of large-scale systematic errors, for example, in leading to a systematic over-population of dominant circulation regimes. From the above reasoning Palmer (2001) proposes the use of stochastic-dynamical parameterization schemes. In a very simple form stochastic physics is used in the Ensemble Prediction System (Buizza et al., 1999) to account for the

influence of model errors. Research experiments using the Seasonal Forecasting system have been run with and without stochastic physics for several years. It has been found that systematic errors of rainfall over South America were largely reduced using stochastic physics (R. Hagedorn, personal communication). However, other systematic errors like, for example, the underestimation of the frequency of occurrence of Euro-Atlantic blocking events, show little improvement in the experiment with stochastic physics switched on (not shown). Development of a revised stochastic physics scheme based on cellular automata is in progress.

Synoptic Activity   It has been shown that the model underestimates synoptic activity by up to 25% over the high-latitude Northern Hemisphere. From the diagnostics applied so far, however, it is not clear what the causes of these error are. One way to shed further light on the nature of these errors would be to determine cyclone and anticyclone statistics using feature-based tracking schemes (see Hoskins and Hodges, 2002, for an overview) instead of applying an Eulerian scheme as used in this study.

Another approach to try to unravel the causes for the underestimation of synoptic variability in the ECWMF model would be study the influence of diabatic processes on the development of cyclones taking a Lagrangian perspective (e.g. Wernli and Davis, 1997; Wernli et al., 2002). The Lagrangian Analysis Tool (LAGRANTO) has been recently implemented at ECMWF. By applying LAGRANTO to both the analysis and the model, a direct comparison could reveal possible model deficiencies in simulation cyclone related diabatic processes. Moreover, the skill of the ECMWF model to produce realistic intrusion of high potential vorticity (PV) air parcels from the stratosphere could be quantified. Research along this line is planned to be carried out in the near future.

Atmospheric Blocking   It has been shown that the latest ECMWF model still shows the tendency to underestimate the instantaneous blocking frequency in the Euro-Atlantic sector. From the diagnostics presented in this study there is an indication that this is primarily due to an underestimation of the number of occurrences of blocking episodes rather than due to an underestimation of their duration. However, it should be mentioned that the number of blocking episodes and the distribution of episode durations are somewhat sensitive to the details of how they are defined in the automatic scheme. Given that these ambiguities occur only for a limited number of blocking episodes it might be beneficial and practical to use semi-automatic schemes to investigate systematic model errors. Such a scheme could be automatic for clearly defined blocking episodes (i.e., for most of the cases) and an experienced user could make decisions for equivocal cases by considering the time evolution of the full flow. In this way it would be possible to obtain a reliable dataset (for ERA-40 and the model) of blocking episode onsets and durations from which more reliable conclusions could be drawn.

A shortcoming of the present study is that systematic blocking errors in the North Pacific region have been merely mentioned in passing. This is because the Tibaldi and Molteni (1990) scheme is not very reliable in the North Pacific (Pelly and Hoskins, 2003). This is primarily associated with the fact that the same blocking latitude in used for the Euro-Atlantic region and the North Pacific in the Tibaldi and Molteni (1990) scheme. Where possible, future studies will be based on the latter scheme[10] which has also the advantage of being dynamically more relevant.

It is also planned to further investigate the role of different dynamical and physical processes for atmospheric blocking. This sort of study is partly motivated by a recent finding at ETH Zuerich (Conny Schwierz, personal communication) that diabatic processes play a non-negligible role for the onset and maintenance of atmospheric blocks. First, the role of diabatic processes in the vicinity of blocks will be estimated for the ECMWF model using LAGRANTO, that is, changes of PV, potential temperature and specific humidity along trajectories will

---

[10]Potential temperature on the PV2 surface, which is employed by the Pelly and Hoskins (2003) scheme, however, has not been archived at ECMWF for operational analyses and forecasts prior to July 2000.

be evaluated. Then the model results will be compared to those based on analysis data. Finally, the influence of different physical processes will be estimated by studying the contributions from different physical parameterizations. This might help to study the sensitivity of blocking to physical processes.

Madden-and-Julian Oscillation   Despite intense research on the MJO over the last decade, it is still unclear which physical processes (or process feedbacks) are central to the phenomenon, and therefore essential to represent well in global circulation models. Earlier studies examined the role that instability theories such as Frictional Wave-CISK and Wind Induced Surface Heat Exchange (WISHE) could possibly play (e.g. Blade and Hartmann, 1993; Hendon and Salby, 1994). Additionally a number of studies have focussed on interaction with sea surface temperatures. For example, Hendon and Glick (1997) documented the cycle of sea surface temperatures that accompany the passage of an MJO event, but from observations it is not always easy to identify how important these are for amplifying and maintaining MJO wave activity. The enhancement of surface temperature ahead of the convective zone could equally be a diagnostic feature of the MJO. Indeed, while investigators have found that their representation of MJO activity improved when coupling an atmospheric model to a responsive ocean (Kemball-Cook et al., 2002; Inness and Slingo, 2003), the fact that atmosphere only models can represent MJO-type activity (Slingo et al., 1996) indicates that ocean feedbacks play a refining role, but are not central to its existence.

In addition to the above mechanisms, recent studies have started to question whether other thermodynamic feedbacks form an essential aspect of intraseasonal variability. Atmosphere-only mechanisms that have been suggested include feedbacks with surface fluxes through convectively induced low level circulations, which Maloney and Hartmann (1998) indicated could play a role in MJO activity. Certainly Maloney (2002) documented a central role for low level moisture convergence in his investigations. However, it is possible that atmospheric water vapour plays a much more general role in the MJO. It is already clear from observations (Parsons et al., 2000) and modelling studies (Redelsperger et al., 2002; Tompkins, 2001a,b) that a feedback between water vapour and convection could exist. They show that a dry mid to upper troposphere can significantly suppress convection, and that convection can locally moisten its environment, preconditioning it for future convective activity. Further studies have documented significant modulations of mid and upper troposphere humidity associated with the MJO (Mote et al., 2000; Sassi et al., 2002). In recent idealized global simulations, Grabowski (2003) claims that the strong MJO-like signal results directly from feedback with water vapour. Blade and Hartmann (1993) also gave evidence of a discharge-recharge mechanism involving water vapour. In addition to the direct thermodynamic effect of humidity anomalies, Raymond (2001) has also suggested radiative feedbacks with the convectively generated clouds. Many numerical investigations have documented a role for radiative feedbacks in determining the tropical circulation (Slingo and Slingo, 1988; Sherwood et al., 1994), and thus it is likely that these will also have some effect in MJO convective organisation.

These thermodynamic feedbacks involving water raise interesting issues concerning the ability of the ECMWF model to simulate correctly intraseasonal variability. For example, if the model under-represents the effect that dry mid-troposheres have on suppressing deep convection, then it is likely that convection will propagate too quickly. In this case oscillation phase speeds are likely to approach those relevant for dry atmospheric Kelvin waves. Indeed, the analysis contained elsewhere in this report concerning MJO phase speeds demonstrated that this is, like for almost all other models, a fault of the ECMWF model. Preliminary studies indicate that a number of the above mechanisms do indeed play a significant role in the ECMWF model. The aim is to conduct a further series of long integrations, with each experiment attempting to isolate one feedback mechanism in turn. For example, the tropospheric water vapour-deep convection feedback can be prevented by artificially forcing the convective scheme to entrain zonal mean humidity profiles, in the Tropics and above the boundary layer. In this way, local variations in humidity can not suppress or enhance deep convective activity.

A pronounced sensitivity of the variance of the MJO to the horizontal resolution has been found, with the

relatively low-resolution model (T$_L$95) showing more realistic values of MJO variance than higher-resolution models (T$_L$159, T$_L$255 and T$_L$511). The reason for this sensitivity is not yet understood. What is clear from preliminary diagnostics, though, is that ITCZ tends to excessively split over the Indian ocean, the Maritime Continent, and the western tropical Pacific at T$_L$159. This tendency to split is much less pronounced in the T$_L$95 integration. Moreover, the amount of large-scale precipitation produced by the model in the tropical belt is highly sensitive to horizontal resolution, the T$_L$95 model showing much more large-scale precipitation over the eastern Indian ocean and western tropical Pacific ocean than the T$_L$159 model. Regarding the MJO, one possible explanation is that the PV forcing from large-scale precipitation is more efficient in triggering MJO-type anomalies than the forcing from convective precipitation.

Flow-Dependent Systematic Errors   A topic which has not been covered in the present study is the possible flow-dependence of systematic errors. In fact, Eqn. 1, which has been used exclusively in this study, is only of limited value to identify these sort of errors. Flow-dependent systematic errors have been studied by Ferranti et al. (2002). They applied Singular Value Decomposition (SVD) to anomalies of Z500 analyses (o'_t) and Z500 forecast errors (e'_t). The dominant mode of flow-dependent systematic error in the North Atlantic region involves the North Atlantic Oscillation (NAO); during high NAO episodes the model tends to underestimate the NAO at D+7 and vice versa. However, as pointed out by Ferranti et al. (2002), there are some inherent difficulties associated with this method, that is, the dominant modes of flow-dependent systematic errors identified by the SVD technique can largely be explained by the loss of predictability. Although, the SVD technique can pick-up some of the flow-dependence, it is not straightforward to separate this contribution from those due to the loss of predictability. This problem could be circumvented by studying systematic errors beyond the limit of predictability (i.e., by employing Eqn. 1). Alternatively, the influence from the loss of predictability could be estimated and taken into account when applying the SVD technique.

Extratropical Weather Regimes   Another topic that has not been covered in the present study are possible systematic errors of the ECMWF model in simulating the observed characteristics of extratropical weather regimes. The realistic simulation of weather regimes is crucial for medium-range forecasts as well as for Monthly Forecasting and Seasonal Forecasting. Note in this context that the kinetic energy of the transient eddies is also considerably underestimated by the ECMWF model on spatial scales larger than those typically associated with cyclones and anticyclones. In fact, weather regimes are a suitable framework to interpret systematic errors of the extratropical atmospheric flow (Palmer et al., 2000). In fact, the pronounced systematic Z500 error in the North Pacific (Fig. 23) resembles "Cluster C" described in Corti et al. (1999) suggesting that this cluster is overpopulated in the latest ECMWF model. Moreover, it has been mentioned that the westerly wind bias in the North Atlantic region, which has been prominent through all model cycles of the 1980s and 1990s (Palmer et al., 2000; Brankovic et al., 2002), is absent in the latest ECMWF model. This suggests that recent model improvements have lead to a more realistic simulation of frequency of occurrence of "Cluster D" (positive phase of the NAO) in the ECMWF model.

It is planned to perform a detailed analysis of the observed regime structure based on ERA-40 data using different linear and nonlinear diagnostic techniques. The results will not only provide the reference against which to compare model results; they are also expected to shed new light on the regime structure (including regime transitions) of the troposphere *as well as* the stratosphere. It is expected that improvements of the vertical representation in ERA-40 compared to previous reanalyses will be of particular importance for diagnosis of the stratosphere. The same sort of diagnostics will be applied to the 6 month integrations based on the latest model cycle 26r1, which have been used extensively within this study. In this way it will also be possible to characterize and assess the regime structure of the latest ECMWF model.

Moreover, the adjoint technique will be applied to determine the sensitivity of particular weather regimes—

identified using ERA-40 data—to model tendency perturbations. It is expected that this approach, which is basically an extension of the studies by Oortwijn and Barkmeijer (1995) and Barkmeijer et al. (2003), reveals the basic instabilities underlying extratropical weather regimes. Apart from being of general interest, the diagnostic outlined above might form the fundament to tackle systematic errors of the extratropical flow.

Extreme Events   Finally, possible systematic errors of the ECWMF model in simulation of extreme events have not been considered in this study. An assessment of extreme events is crucial, however, given the increasing interest in severe weather predictions. Whereas the focus in this study has been on the mean and the width of the probability density function (PDF) of different parameters, it would be important to know how well the ECMWF model simulates the tails of the PDF and possible departures from the Gaussian distribution. Given the availability of ERA-40 data and long integrations of the latest ECWMF model such an undertaking has become practicable and will be carried out in the near future.

References

Adler, R. F., C. Kidd, G. Petty, M. Morissey, and H. M. Goodman, 2001: Intercomparison of global precipitation products: The third precipitation intercomparison project (PIP-3). *Bull. Amer. Meteor. Soc.*, **82**, 1377–1396.

Anderson, D., T. Stockdale, M. Balmaseda, L. Ferranti, F. Vitart, P. Doblas-Reyes, R. Hagedorn, T. Jung, A. Vidard, and T. Palmer, 2002: Comparison of the ECMWF seasonal forecast System 1 and 2, including the relative performance for the 1997/98 El Nino. Technical Report 404, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK.

Arkin, A. and P. P. Xie, 1994: The global precipitation climatology project: First algorithm intercomparison project. *Bull. Amer. Meteor. Soc.*, **75**, 401–419.

Barkmeijer, J., T. Iversen, and T. Palmer, 2003: Forcing singular vectors and other sensitive model structures. *Q. J. R. Meteorol. Soc.*, **129**, 2401–2423.

Barnett, T. P., K. Hasselmann, M. Chelliah, T. Delworth, G. Hegerl, P. Jones, E. Rasmussen, E. Roeckner, C. Ropelski, B. Santer, and S. Tett, 1999: Detection and attribution of recent climate change: A status report. *Bull. Amer. Meteor. Soc.*, **80**, 2631–2659.

Blade, I. and D. L. Hartmann, 1993: Tropical intraseasonal oscillations in a simple nonlinear model. *J. Atmos. Sci.*, **50**, 2922–2939.

Brankovic, C., C. Jakob, M. Miller, A. Untch, and N. Wedi, 2002: Climate diagnostics of the ECMWF AMIP-2 simulations. Technical Report 360, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK.

Buizza, R., M. Miller, and T. Palmer, 1999: Stochastic representation of model uncertainities in the ECWMF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **125**, 2887–2908.

Chevallier, F. and P. Bauer, 2003: Model rain and clouds over oceans: Comparison with SSM/I observations. *Mon. Wea. Rev.*, **131**, 1240–1255.

Chevallier, F., P. Bauer, G. Kelly, C. Jakob, and T. McNally, 2001: Model clouds over oceans as seen from space: Comparison with HIRS/2 and MSU radiances. *J. Climate*, **14**, 4216–4229.

Chevallier, F. and G. Kelly, 2002: Model clouds as seen from space: Comparison with geostationary imagery in the 11-micron window channel. *Mon. Wea. Rev.*, **130**, 712–722.

Chevallier, F. and J.-J. Morcrette, 2000: Comparison of model fluxes with surface and top-of-the-atmosphere observations. *Mon. Wea. Rev.*, **128**, 3839–3852.

Ciesielski, P. E., R. H. Johnson, P. T. Haertel, and J. H. Wang, 2003: Corrected TOGA COARE sounding humidity data: Impact on diagnosed properties of convection and climate over the warm pool. *J. Climate*, **16**, 2370–2384.

Corti, S., F. Molteni, and T. N. Palmer, 1999: Signature of recent climate change in the frequency of natural atmospheric circulation regimes. *Nature*, **398**, 799–802.

D'Andrea, F., S. Tibaldi, M. Blackburn, G. Boer, M. Deque, M. R. Dix, B. Dugas, L. Ferranti, T. Iwasaki, A. Kitoh, D. Pope, V. Randall, E. Roeckner, D. Straus, W. Stern, H. Van den Dool, and D. Williamson, 1998: Northern Hemisphere atmospheric blocking as simulated by 15 atmospheric general circulation models in the period 1979–1988. *Clim. Dyn.*, **14**, 385–407.

Dubuisson, P., J. C. Buriez, and Y. Fouquart, 1996: High spectral resolution solar radiative transfer in absorbing and scattering media: Application to the satellite simulation. *J. Quant. Spectrosc. radiat. transfer*, **55**, 103–126.

Ferranti, L., E. Klinker, A. Hollingsworth, and B. J. Hoskins, 2002: Diagnosis of systematic forecast errors dependent on flow pattern. *Q. J. R. Meteorol. Soc.*, **128**, 1623–1640.

Ferranti, L., F. Molteni, C. Brankovic, and T. N. Palmer, 1994: Diagnosis of extratropical variability in seasonal integrations of the ECMWF model. *J. Climate*, **7**, 849–868.

Ferranti, L., T. N. Palmer, F. Molteni, and E. Klinker, 1990: Tropical-extratropical interaction associated with the 30–60 day oscillation and its impact on medium and extended range prediction. *J. Atmos. Sci.*, **47**, 2177–2199.

Gates, W. L., J. S. Boyle, C. Covey, C. G. Dease, C. M. Doutriaux, R. S. Drach, M. Fiorino, P. J. Gleckler, J. Hnilo, S. M. Marlais, T. J. Phillips, G. L. Potter, B. D. Santer, K. R. Sperber, K. E. Taylor, and D. N. Williams, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, **80**, 29–55.

Grabowski, W. W., 2003: MJO-like coherent structures: Sensitivity simulations using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.*, **60**, 847–864.

Gregory, D., J. J. Morcrette, C. Jakob, A. C. M. Beljaars, and T. Stockdale, 2000: Revision of convection, radiation and cloud schemes in the ECMWF Integrated Forecasting System. *Q. J. R. Meteorol. Soc.*, **126**, 1685–1710.

Guichard, F., D. Parsons, and E. Miller, 2000: Thermodynamic and radiative impact of the correction of sounding humidity bias in the tropics. *J. Climate*, **13**, 3611–3624.

Hendon, H. H. and J. Glick, 1997: Intraseasonal air-sea interaction in the tropical Indian and Pacific Oceans. *J. Climate*, **10**, 647–661.

Hendon, H. H. and M. L. Salby, 1994: The life-cycle of the Madden-Julian oscillation. *J. Atmos. Sci.*, **51**, 2225–2237.

Heymsfield, A. J. and L. J. Donner, 1990: A scheme for parameterizing ice-cloud water content in general circulation models. *J. Atmos. Sci.*, **47**, 1865–1877.

Hogan, R. J., C. Jakob, and A. J. Illingworth, 2001: Comparison of ECMWF winter-season cloud fraction with radar-derived values. *J. Appl. Meteor.*, **40**, 513–525.

Hoskins, B. J. and K. I. Hodges, 2002: New perspectives on the Northern Hemisphere winter storm tracks. *J. Atmos. Sci.*, **59**, 1041–1061.

Hoskins, B. J. and P. J. Valdes, 1990: On the existence of storm tracks. *J. Atmos. Sci.*, **47**, 1854–1864.

Inness, P. M. and J. M. Slingo, 2003: Simulation of the Madden-Julian oscillation in a coupled general circulation model. Part II: The role of the basic state. *J. Climate*, **16**, 345–364.

Jakob, C., 1999: Cloud cover in the ECMWF reanalysis. *J. Climate*, **12**, 947–959.

Jakob, C., 2002: Ice clouds in numerical weather prediction models: Progress, problems and prospects. In: D. K. Lynch, K. Sassen, D. Starr, and G. Stephens, eds., *Cirrus*, pp. 327–345. Oxford University Press.

Jenkins, G. M. and D. G. Watts, 1968: *Spectral analysis and its application*. Holden-Day. 525 pp.

Kemball-Cook, S., B. Wang, and X. H. Fu, 2002: Simulation of the intraseasonal oscillation in the ECHAM-4 model: The impact of coupling with an ocean model. *J. Atmos. Sci.*, **49**, 1433–1453.

Klein, S. A. and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.*, **127**, 2514–2531.

Klinker, E. and L. Ferranti, 2000: Use of tendency-balance to diagnose the performance of models and data assimilation systems. In: *ECMWF Workshop on Diagnosis of models and data assimilation systems*, pp. 333–346. ECMWF, Shinfield Park, Reading RG2 9AX, UK.

Klinker, E. and P. D. Sardeshmukh, 1992: The diagnosis of mechanical dissipation in the atmosphere from large-scale balance requirements. *J. Atmos. Sci.*, **49**, 608–627.

Lalaurette, F., L. Ferranti, A. Ghelli, Ø. Saetra, and H. Böttger, 2003: Verification statistics and evaluations of ECMWF forecasts in 2001-2002. Technical Report 414, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, U.K.

Mace, G. G., C. Jakob, and K. P. Moran, 1998: Validation of hydrometeor occurrence predicted by the ECMWF model using millimeter wave radar data. *Geophys. Res. Lett.*, **25**, 1645–1648.

Maloney, E. D., 2002: An intraseasonal oscillation composite life cycle in the NCAR CCM3.6 with modified convection. *J. Climate*, **15**, 964–982.

Maloney, E. D. and D. L. Hartmann, 1998: Frictional moisture convergence in a composite life cycle of the Madden-Julian oscillation. *J. Climate*, **11**, 2387–2403.

Miller, M. J., A. C. M. Beljaars, and T. N. Palmer, 1992: The sensitivity of the ECMWF model to the parameterization of evaporation from the tropical oceans. *J. Climate*, **5**, 418–434.

Miller, S. D., G. L. Stephens, and A. C. M. Beljaars, 1999: A validation survey of the ECMWF prognostic cloud scheme using LITE. *Geophys. Res. Lett.*, **26**, 1417–1420.

Morcrette, J. J., 2002: Assessment of the ECMWF model cloudiness and surface radiation fields at the ARM SGP site. *Mon. Wea. Rev.*, **130**, 257–277.

Mote, P. W., H. L. Clark, T. J. Dunkerton, R. S. Harwood, and H. C. Pumphrey, 2000: Intraseasonal variations of water vapor in the tropical upper troposphere and tropopause region. *J. Geophys. Res.*, **105**, 17457–17470.

Neale, R. and J. Slingo, 2003: The Maritime Continent and its role in global climate: a GCM study. *J. Climate*, **16**, 834–848.

Oortwijn, J. and J. Barkmeijer, 1995: Perturbations that optimally trigger weather regimes. *J. Atmos. Sci.*, **52**, 3932–3944.

Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parameterization in weather and climate prediction models. *Q. J. R. Meteorol. Soc.*, **127**, 279–304.

Palmer, T. N., C. Brankovic, R. Buizza, P. Chessa, L. Ferranti, B. J. Hoskins, and A. J. Simmons, 2000: A review of predictability and ECMWF forecast performance, with emphasis on Europe. Technical Report 326, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK.

Pandithurai, G., R. T. Pinker, O. Dubovik, B. N. Holben, and T. O. Aro, 2001: Remote sensing of aerosol optical characteristics in sub-Sahel, West Africa. *J. Geophys. Res.*, **106**, 28347–28356.

Parsons, D. B., K. Yoneyama, and J.-L. Redelsperger, 2000: The evolution of the tropical western Pacific atmosphere-ocean system following the arrival of a dry intrusion. *Q. J. R. Meteorol. Soc.*, **126**, 517–548.

Peixoto, J. P. and A. H. Oort, 1992: *Physics of Climate*. American Institute of Physics, New York.

Pelly, J. L. and B. J. Hoskins, 2003: A new perspective on blocking. *J. Atmos. Sci.*, **60**, 743–755.

Ramanathan, V., 1987: The role of Earth Radiation Budget studies in climate and general-circulation research. *J Geophys. Res-Atmos*, **D4**, 4075–4095.

Raymond, D. J., 2001: A new model of the Madden Julian Oscillaton. *J. Atmos. Sci.*, **58**, 2807–2819.

Redelsperger, J. L., D. B. Parsons, and F. Guichard, 2002: Recovery processes and factors limiting cloud-top height following the arrival of a dry intrusion observed during TOGA COARE. *J. Atmos. Sci.*, **59**, 2438–2457.

Rossow, W. B. and L. C. Garder, 1993: Cloud detection using satellite measurements of infrared and visible radiances for ISCCP. *J. Climate*, **6**, 2341–2369.

Rossow, W. B. and R. A. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.*, **72**, 2–20.

Rossow, W. B. and R. A. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.*, **80**, 2261–2287.

Sardeshmukh, P. D. and B. J. Hoskins, 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228–1251.

Sassi, F., M. Salby, H. C. Pumphrey, and W. G. Read, 2002: Influence of the Madden-Julian Oscillation on upper tropospheric humidity. *J. Geophys. Res.*, **107**, 10.1029/2001JD001331.

Sherwood, S. C., V. Ramanathan, T. P. Barnett, M. K. Tyree, and E. Roeckner, 1994: Response of an atmospheric general-circulation model to radiative forcing of tropical clouds. *J. Geophy. Res.*, **99**, 20829–20845.

Simmons, A. and A. Hollingsworth, 2001: Some aspects of the improvement of skill of numerical weather prediction. Technical Report 342, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK.

Slingo, A. and J. M. Slingo, 1988: The response of a general-circulation model to cloud longwave radiative forcing .1. Introduction and initial experiments. *Q. J. R. Meteorol. Soc.*, **114**, 1027–1062.

Slingo, J. M., 1998: The 1997–98 El Nino. *Weather*, **53**, 274–281.

Slingo, J. M., K. R. Sperber, J. S. Boyle, J. P. Ceron, M. Dix, B. Dugas, W. Ebisuzaki, J. Fyfe, D. Gregory, J. F. Gueremy, J. Hack, A. Harzallah, P. Inness, A. Kitoh, W. K. M. Lau, B. McAvaney, R. Madden, A. Matthews, T. N. Palmer, C. K. Park, D. Randall, and N. Renno, 1996: Intraseasonal oscillations in 15 atmospheric general circulation models: Results from an AMIP diagnostic subproject. *Clim. Dyn.*, **12**, 325–357.

Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 1641–1657.

Teixeira, J. and T. F. Hogan, 2002: Boundary layer clouds in a global atmospheric model: Simple cloud cover parameterizations. *J. Climate*, **15**, 1261–1276.

Tibaldi, S., F. D'Andrea, E. Tosi, and E. Roeckner, 1997: Climatology of Northern Hemisphere blocking in the ECHAM model. *Clim. Dyn.*, **13**, 649–666.

Tibaldi, S. and F. Molteni, 1990: On the operational predictability of blocking. *Tellus*, **42A**, 343–365.

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Mon. Wea. Rev.*, **121**, 3040–3061.

Tompkins, A. M., 2001a: On the relationship between tropical convection and sea surface temperature. *J. Climate*, **14**, 633–637.

Tompkins, A. M., 2001b: Organization of tropical convection in low vertical wind shears: The role of water vapor. *J. Atmos. Sci.*, **58**, 529–545.

Tompkins, A. M., 2002: A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *J. Atmos. Sci.*, **59**, 1917–1942.

Tselioudis, G. and C. Jakob, 2002: Evaluation of midlatitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. *J. Geophys. Res.*, **107**, 10.1029/2002JD002259.

Turner, D. D., B. M. Lesht, S. A. Clough, J. C. Liljegren, H. E. Revercomb, and D. C. Tobin, 2003: Dry bias and variability in Vaisala RS80-H radiosondes: The ARM experience. *J. Atmos. Ocean. Tech.*, **20**, 117–132.

Uppala, S., 2002: ECMWF ReAnalysis 1957–2001. In: *3. Workshop on Re-analysis*. ECMWF, Shinfield Park, Reading RG2 9AX, UK.

Vitart, F., 2003: Monthly Forecasting System. Technical Report 342, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK.

von Storch, H. and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press. 484 pp.

Wang, J. H., H. L. Cole, D. J. Carlson, E. R. Miller, K. Beierle, A. Paukkunen, and T. K. Laine, 2002: Corrections of humidity measurement errors from the Vaisala RS80 radiosonde - Application to TOGA COARE data. *J. Atmos. Ocean. Tech.*, **19**, 981–1002.

Webb, M., C. Senior, S. Bony, and J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Clim. Dyn.*, **17**, 905–922.

Wentz, F. J., 1997: A well-calibrated ocean algorithm for SSM/I. *J. Geophy. Res.*, **102**, 8703–8718.

Wernli, H. and H. C. Davis, 1997: A Lagrangian-based analysis of extratropical cyclones. I: The method and some applications. *Q. J. R. Meteorol. Soc.*, **123**, 467–489.

Wernli, H., S. Dirren, M. A. Liniger, and M. Zillig, 2002: Dynamical aspects of the life cycle of the winter storm 'Lothar' (24–26 december 1999). *Q. J. R. Meteorol. Soc.*, **128**, 405–429.

Xie, P. and P. A. Arkin, 1997: Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558.