

Performance of the ECMWF Forecasting System in Polar Regions

T. Jung and M. Leutbecher

*ECMWF, Shinfield Park, Reading
RG2 9AX, United Kingdom
jung@ecmwf.int*

ABSTRACT

The performance of the ECMWF forecasting system during boreal winter/austral summer in polar regions is investigated. A comparison between different analysis products suggests that synoptic-scale features in the Arctic are relatively well represented by state-of-the-art analysis systems. Furthermore, it is shown that the improvement of deterministic forecast error in polar regions since the early 1980s follows closely that reported in previous studies for the Northern and Southern hemisphere as a whole. One of the biggest “jumps” in forecast skill, both in the troposphere and stratosphere, occurred in autumn 2000 when the high-resolution system (40 km) was introduced at ECMWF. Our results suggest that 10-day forecasts of stratospheric warming events are very skilful, particularly since the introduction of the high-resolution system. Verification of the ECMWF ensemble prediction system (EPS) in the Arctic reveals substantial improvements in probabilistic predictive skill since the mid-1990s. The most pronounced increase in probabilistic forecast skill, particularly on synoptic scales, occurred in autumn 2000 when the high-resolution EPS (about 80 km) was introduced. The pronounced flow-dependence of the sensitivity of short-range forecast error in the Arctic to initial perturbations, as revealed by adjoint sensitivity computations, highlights the importance of using ensemble prediction systems for weather forecasting in the Arctic. Moreover, by diagnosing seasonal integrations of the ECMWF model, it is shown that relatively low-resolution models (T_L95) such as used to carry out seasonal and climate predictions underestimate the observed level of synoptic variability in polar regions of both hemispheres substantially. Increasing resolution to T_L511 (about 40 km) leads to much more realistic levels of synoptic activity. Finally, it is shown that the observed stratosphere-troposphere link (“downward propagation” of stratospheric anomalies) is realistically represented in seasonal integrations with the ECMWF model—even at the relatively low horizontal resolution of T_L95 .

1 Introduction

Results which are based on numerical weather forecasting (NWP) models are increasingly being used by a wide range of different users. Nowadays, weather forecasts out to 10 days and longer, for example, provide the basis for decision making (e.g. Morss et al., 2005). Moreover, analysis or reanalysis data (Kalnay et al., 1996; Uppala et al., 2005), which to various degree (depending on the parameter and the region being considered) are based on a first guess (i.e., a short-range model forecast) are nowadays routinely being used by the climate community. Given the importance of NWP products their validation is a crucial task. Traditionally, most of the validation efforts have focussed on the mid-latitudes¹ and tropics. The polar regions, on the other hand, have attracted relatively little attention. The aim of this study is to improve our understanding of the performance of one of the most widely used NWP systems—the ECMWF forecasting system—in polar regions. This seems to be a particularly timely task with the International Polar Year (IPY) starting in early 2007.

More specifically the following questions will be investigated in this study:

- How realistic is the representation of the atmospheric circulation in current data assimilation systems?

¹The model performance over the Northern Hemisphere as a whole, of which polar regions represent only a small portion, is one of the most widely studied regions.

- How predictable is the atmospheric circulation in a deterministic sense in polar regions, both in the troposphere and the stratosphere?
- How predictable are sudden stratospheric warming events?
- How does the ECMWF ensemble prediction system perform in polar regions?
- How sensitive are forecast errors in polar region to initial perturbations and, in particular, in which regions do initial perturbations potentially have the largest influence on subsequent forecasts?

While addressing the above mentioned questions we will also discuss how the performance of the ECMWF forecasting system has changed over the years.

Two further aspects will be addressed in this study using seasonal integrations with recent versions of the ECMWF model. It is well-known that low-pressure systems at high-latitudes are of relatively small-scale structure. Polar lows (see [Rasmussen and Turner, 2003](#), for an overview), for example, represent one of the most important features in this context. Given the relatively small-scale nature of high-latitude cyclones ([Simmonds, 2000](#)) it seems reasonable to assume that increased resolution will be beneficial. Moreover, it has been suggested that large anomalies of the strength of the stratospheric polar vortex have a significant impact (after some delay) on weather systems in the troposphere ([Baldwin and Dunkerton, 2001](#)). Given that monthly forecasts are operationally being produced at ECMWF since October 2004 ([Vitart, 2004](#)) and that extended-range forecasts are likely to benefit the most from the “stratosphere-troposphere link” ([Baldwin et al., 2003](#); [Jung and Barkmeijer, 2006](#)), it is crucial to investigate whether the observed stratosphere-troposphere connection is present in seasonal integrations with the ECMWF model.

Throughout this study the focus will be on the atmospheric circulation during boreal winter/austral summer.

The structure of the paper is as follows. The data and methods used will be described in section 2. This is followed by the results section, which encompasses a discussion of the quality of ECMWF analysis and reanalysis data in describing the atmospheric circulation in polar region. Then, the deterministic and probabilistic forecast performance in polar regions will be described. This is followed by a discussion of the sensitivity of forecast error in polar regions to initial conditions, the sensitivity of simulated synoptic-scale features to horizontal resolution as well as the realism of the influence of changes of the tropospheric circulation to changes in the strength of the stratospheric polar vortex. Finally, the conclusions of this study will be summarized and discussed.

2 Data and Methods

2.1 Data

In this study a variety of different data sets is used. The truth is either represented by operational analysis or ERA-40 reanalysis data ([Uppala et al., 2005](#)). Moreover, short-range and medium-range forecasts from three different sources are used: (1) operational high-resolution deterministic forecasts, (2) lower resolution control forecasts from the ECMWF ensemble prediction system (EPS, see also below), and (3) ERA-40 reforecasts. The operational forecasting system underwent substantial changes throughout the last two decades which affected the quality of all operational forecast products, including the analysis. Further details on how the operational forecasting system has changed are given elsewhere (e.g. [Simmons and Hollingsworth, 2002](#); [Jung, 2005](#); [Simmons, 2006](#)). ERA-40 reforecasts, on the other hand, were carried out using the same atmospheric model throughout the period 1958–2001 ([Uppala et al., 2005](#)). Therefore, ERA-40 reforecasts are

extremely useful for understanding the impact of changes of the observing systems and flow-dependent aspects of predictability.

ECMWF carries out ensemble forecasts using a lower-resolution version of the ECMWF model (Molteni et al., 1996) in order to give *a priori* guidance about the predictability of the atmospheric flow of the day and in order to allow users to determine the probability of occurrence of certain events. Initial perturbations are constructed using the singular vector approach (e.g. Buizza and Palmer, 1995). Since October 1998 model uncertainty is taken into account through the use of a simple stochastic physics scheme (Buizza et al., 1999). Since December 1996 the ECMWF encompasses 51 members (one control and 50 perturbed forecasts). The ECMWF EPS underwent numerous changes in recent years. A more detailed overview of the ECMWF EPS and how it changed after its implementation in December 1992 is given by Buizza (2006).

Seasonal forecasts were carried out in order to compare the model's climate with ERA-40 data. The model resolution and version used varies among the different experiments (see Section 3.3 for details). A common feature of the seasonal forecast is that observed sea surface temperature fields were used as lower boundary condition, that forecasts were started on 1 October of each the years considered (e.g., 1991–2006), and that the first month was discarded. A more detailed description of the experimental setup used to carry out the seasonal forecast runs is given elsewhere (Jung and Tompkins, 2003).

2.2 Adjoint sensitivity of forecast error to initial conditions

The adjoint technique (e.g. Errico, 1997) is used in order to study the sensitivity of forecast error in polar regions to initial perturbations. In the following, a short summary of the adjoint sensitivity technique will be given. For a more detailed introduction the reader is referred to the paper by Rabier et al. (1996).

The forecast error can be quantified by means of the diagnostic function (or cost function) $J = 1/2 \langle P(\mathbf{x}_t - \mathbf{x}_t^{ref}), P(\mathbf{x}_t - \mathbf{x}_t^{ref}) \rangle$, where \mathbf{x}_t denotes the forecast (D+2 in this study), \mathbf{x}_t^{ref} the verifying analysis and P is the projection operator (Buizza, 1994) used to confine forecast error to the Northern Hemisphere polar region (north of 70°N). Here, we are interested in changes of J with respect to changes in the initial conditions, that is, we are interested in the gradient $\nabla_0 J$. The gradient contains the sensitivity of forecast error to initial perturbations of temperature, vorticity, divergence and the logarithm of surface pressure at all grid points and vertical levels available. By construction initial perturbations, $\delta \mathbf{x}_0$, which project onto the gradient are efficient in changing J , that is, $\delta J = \langle \nabla_0 J, \delta \mathbf{x}_0 \rangle$. As shown by Rabier et al. (1996), $\nabla_0 J$ can be obtained as follows:

$$\nabla_0 J = C_0^{-1} R^* P^* C_1 P(\mathbf{x}_t - \mathbf{x}_t^{ref}), \quad (1)$$

where C_0 and C_1 are norms at initial and final time, respectively; R^* denotes the adjoint of the tangent forward propagator R ; P^* is the adjoint of the projection operator; and $\mathbf{x}_t - \mathbf{x}_t^{ref}$ is the forecast error at final time t . All adjoint computations were carried using a horizontal resolution of T_L159 with 60 levels in the vertical. In this study, the dry total energy norm (e.g. Ehrendorfer et al., 1999) is used at initial and final time, that is, humidity perturbations were not taken into account. The tangent linear model and its adjoint employ the same linearized physics as used by Mahfouf (1999) comprising vertical diffusion, large-scale condensation, long-wave radiation, deep cumulus convection, and subgrid-scale orographic effects. The adjoint computations were carried out every fifth day for the two winters (DJFM) of 2001/02 and 2004/05 (a total of 25 adjoint computations for each of the two winters).

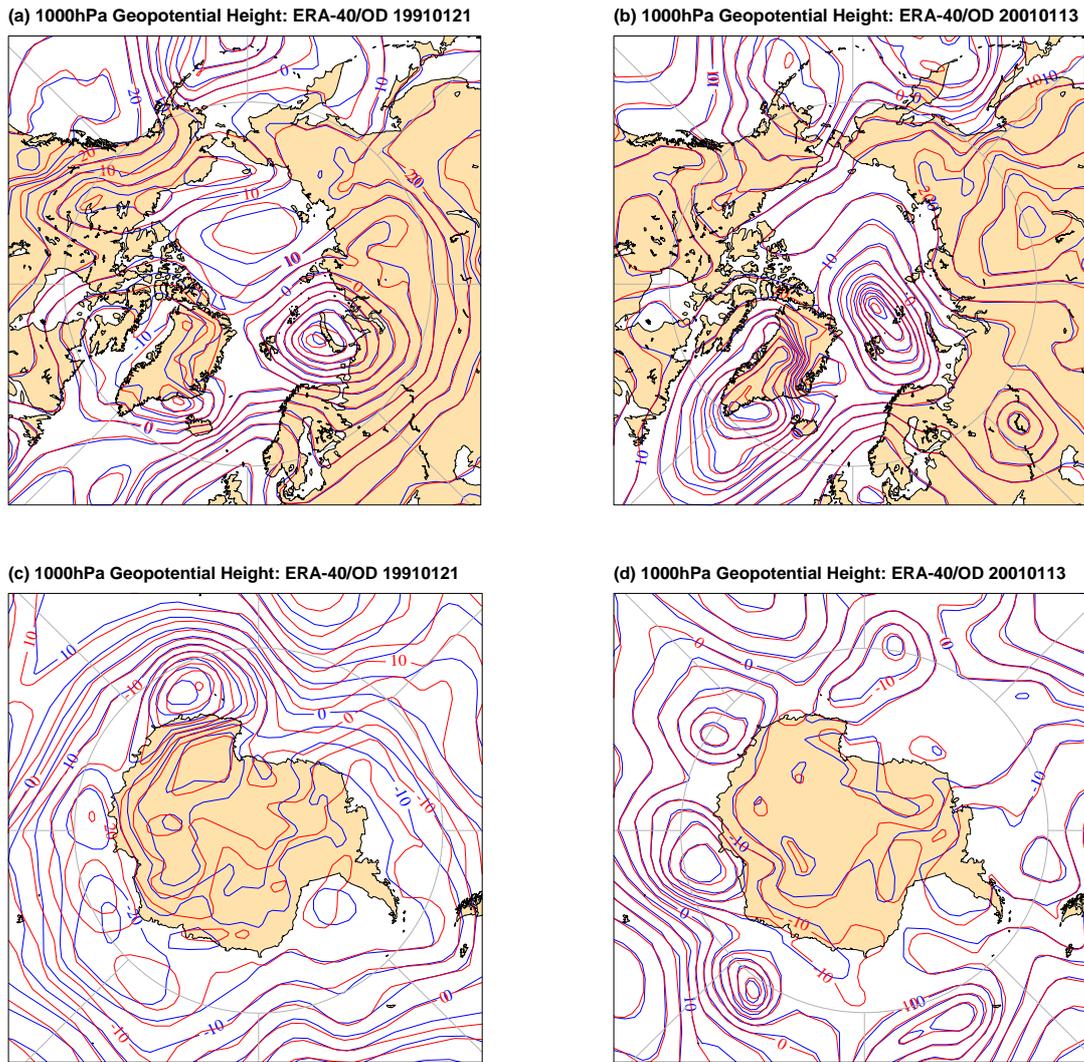


Figure 1: Z1000 fields (in m) at 12UTC on (a),(c) 21 January 1991 and (b),(d) 13 January 2001 from operational analysis (blue contours) and ERA-40 reanalysis (red contours) data. Results are shown separately for the (a)–(b) Northern Hemisphere and (c)–(d) Southern Hemisphere.

3 Results

3.1 Analysis quality

Throughout this study operational analysis and ERA-40 reanalysis data are used as truth for describing the state of the atmospheric circulation. It is important for what follows, therefore, to begin with an assessment of the validity of this assumption. To this end, analysis fields from two different analysis cycles, that is, (a) operational analysis and (b) ERA-40 reanalysis fields, are compared. The two analysis fields differ (depending on the year being considered) in the way data assimilation is carried out and how the model is formulated. If the analysis and reanalysis fields were strongly (weakly) constrained by observations, then small (large) differences between the two fields could be expected. Notice that for the extreme case of no observations used, the two analysis cycles correspond to two independent (except for the initial and lower-boundary conditions) extended-range integrations with (slightly) different atmospheric models.

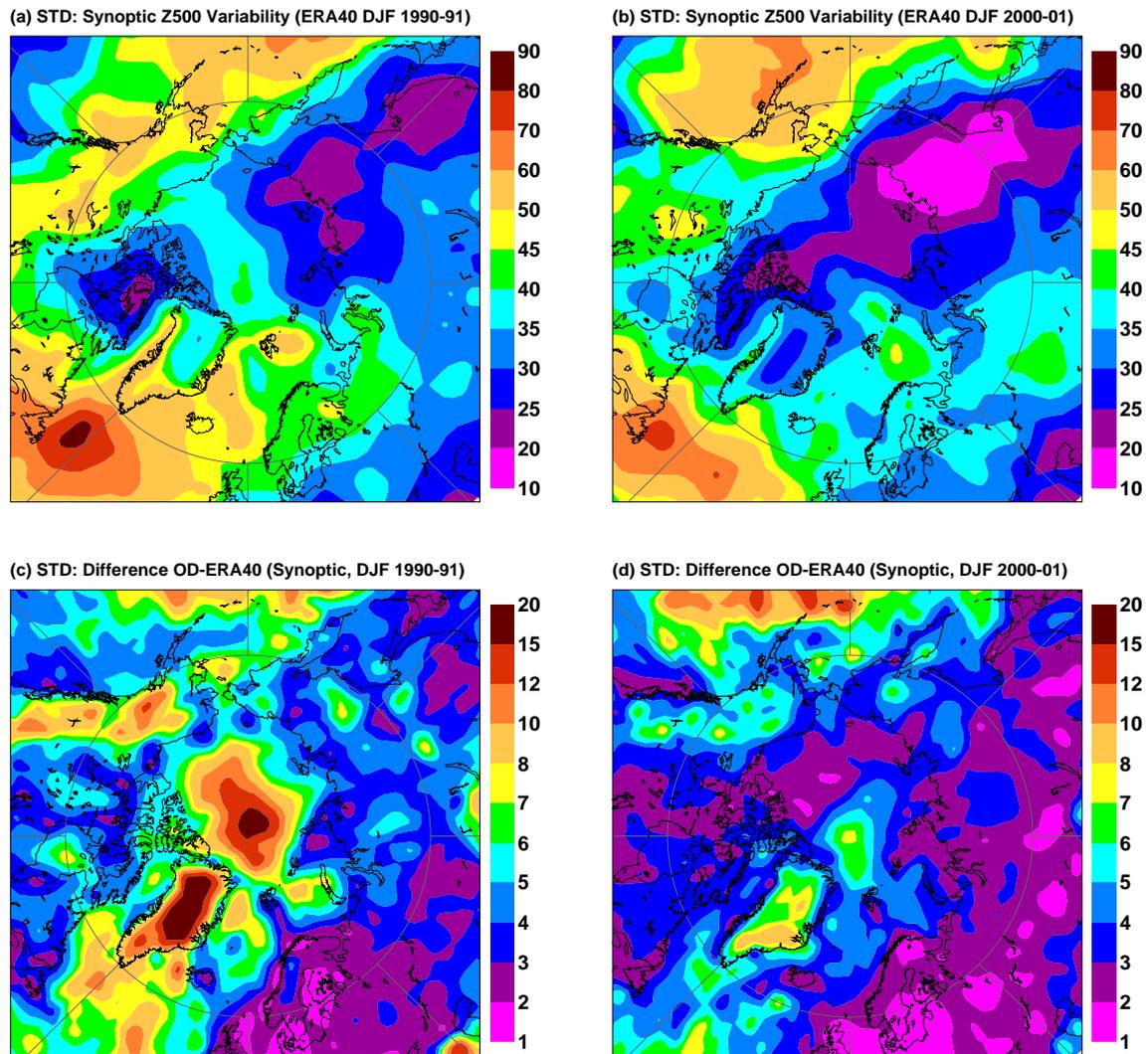


Figure 2: Standard deviation of synoptic Z500 variability (in m) in ERA-40 for the two winters (December–February) of (a) 1990/91 and (b) 2000/01. Synoptic variability has been obtained by highpass-filtering the daily fields retaining variability on time scales shorter than 8 days. Also shown is the standard deviation of the difference between highpass-filtered Z500 fields from ERA-40 and the-then operational analysis for the two winters of (c) 1990/91 and (d) 2000/01.

Snapshots of 1000 hPa geopotential height (Z1000) fields are shown in Figure 1 for 21 January 1991 (left column) and 13 January 2001 (right column) for the-then operational analysis and ERA-40 reanalysis. Over the high-latitude Northern Hemisphere (Figure 1a,b) large-scale and synoptic-scale features in the two analyses are in good agreement. This is particularly true for the January 2001 case for which a more sophisticated data assimilation system (horizontal and vertical resolution, model formulation and 4D-Var instead of optimal interpolation) was used operationally than in the winter of 1990/91. In polar regions over the Southern Hemisphere the analysed flow is also very similar in 2001 for the two assimilation systems; larger differences are apparent in 1991, particularly between 30°E and 90°E. These rather large differences are likely to be the result of the relatively simple data assimilation system (e.g., optimal interpolation) used operationally in the early 1990s (see also next section).

In order to allow more quantitative conclusions to be drawn, the standard deviation of highpass-filtered (time

scales < 8 days) 500 hPa geopotential height (Z500) fields from ERA-40 were computed for the two winters 1990/91 and 2000/01 (Figure 2a,b). Also shown in Figure 2 is the standard deviation of the Z500 difference between the-then operational analysis and ERA-40. Large values in Figure 2c,d reflect large uncertainties of (re-)analysed Z500 fields on synoptic scales. For both winters relatively large analysis uncertainties are found over the North Atlantic, North Pacific and in polar regions. Conversely, the analysis quality is relatively high in well-observed areas such as Europe and North America. Particularly over the North Atlantic region and in polar regions substantially lower uncertainties are found for the winter 2000/01 compared to 1990/91. Notice, that the smaller uncertainties for the former winter are more representative of sophisticated analysis systems such as that used to carry out the ERA-40 reanalysis. Nevertheless, in high latitudes, north of about 80°N , the uncertainty in representing synoptic scale aspects of Z500 locally may amount to as much as 20% of the day-to-day variability of high-pass filtered Z500 fields for current data assimilation systems (Figure 2b,d).

In summary, the comparison of the synoptic-scale aspects of the tropospheric circulation obtained from two (slightly) different analysis systems suggests that in relative terms uncertainties in polar regions are generally higher than those in the mid-latitude Northern Hemisphere oceans. However, the analysis quality seems high enough (i.e., individual low-pressure and high-pressure systems are well constrained by the available observations) to use (re-)analysis products as a good proxy of the truth.

A different way of assessing the quality of the analysis, that is, by focussing on the quality of short-range forecast error (Uppala et al., 2005; Simmons, 2006), will be described in the following section.

3.2 Short-range and medium-range forecasts

3.2.1 Deterministic forecast skill

Time series of the mean spatial standard deviation of Z500 forecast errors at $D+2^2$ and $D+5$ are shown in Figure 3 for polar regions of the Northern Hemisphere (north of 70°N) and Southern Hemisphere (south of 70°S) and various different deterministic forecasts products (i.e., ERA-40 reforecasts as well as operational high-resolution deterministic and lower-resolution control forecasts of the ECMWF EPS). Also shown are time series of the spatial standard deviation of Z500 fields from operational analyses and ERA-40. If the forecast error is smaller than day-to-day variations then there is more deterministic forecast skill than what could be expected from a climatological forecast. Over the Northern Hemisphere (Figure 3a) useful forecast skill is found at $D+2$ and $D+5$ for all different forecast products throughout the whole period. Over the Southern Hemisphere, on the other hand, it turns out that $D+5$ reforecasts from ERA-40 are useful since the late 1970s (i.e., when satellite data became available), which highlights the importance of satellite data over the Southern Hemisphere for the quality of deterministic Z500 forecasts (see also Uppala et al., 2005).

One interesting feature revealed by the $D+2$ and $D+5$ forecasts shown in Figure 3 is that secular as well as abrupt changes such as from 2000 to 2001 show up at both forecast ranges. The signal-to-noise-ratio, however, is considerably higher for $D+2$ forecasts. In the following, therefore, the focus will be on $D+2$ forecasts.

The influence of improvements in the data assimilation scheme and the forecast model used can most readily be understood by comparing operational forecasts with ERA-40 reforecasts (frozen data assimilation scheme and atmospheric model). Differences between operational and ERA-40 forecasts were quite pronounced in both hemispheres during the early 1980s. This gap gradually becomes smaller and smaller from 1982 to 2001. The improvement in operational forecast skill is mostly likely a result of increases in resolution (e.g., resolution increase to T213L31 in September 1991), more realistic representation of physical processes (e.g., introduction of a prognostic cloud scheme in April 1995) and changes to the data assimilation algorithm used (introduction of 3D-Var and 4D-Var in January 1996 and November 1997, respectively). A more detailed description of

²As is common practice in the NWP community the abbreviation $D+n$ is to denote n -day forecasts.

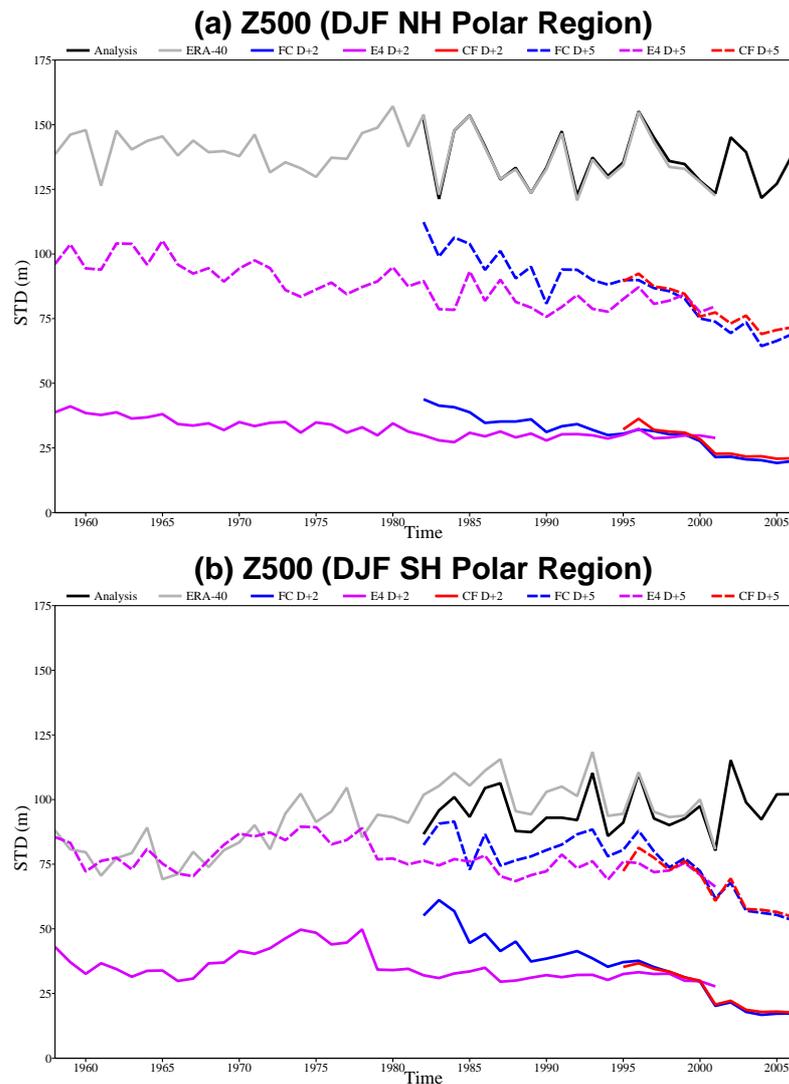


Figure 3: Time series of the temporal mean spatial standard deviation of daily Z500 forecast error at D+2 (solid lines) and D+5 (dashed lines) for the (a) Northern and (b) Southern Hemisphere polar region (poleward of 70°N and 70°S). Three different forecast sets are used: operational deterministic forecasts (blue), EPS control forecasts (red) and ERA-40 reforecasts (purple). Also shown are time series of the mean spatial standard deviation of Z500 fields from operational analyses (black, solid) and ERA-40 reanalysis (grey, solid).

major changes of the operational ECMWF forecasting system can be found elsewhere (see, e.g. Jung, 2005; Simmons, 2006).

The largest improvement of operational D+2 (and D+5) forecasts in polar regions of the Northern Hemisphere occurred in autumn 2000 and was the result of an increase in resolution to T_L511 with 60 levels in the vertical. This interpretation is further substantiated by the fact that ERA-40 reforecasts, for which model resolution is fixed, did not perform significantly better in winter 2000/01 compared to previous winters. The beneficial impact of increased horizontal resolution is also visible in polar regions of the Southern Hemisphere. There is some indication for further improvements of operational forecasts in polar regions in more recent years, particularly at D+5. However, this signal is less clear than the ‘jump’ in autumn 2000, given that large year-to-year fluctuations in forecast skill can occur as a result of changes in the atmospheric flow (Ferranti et al.,

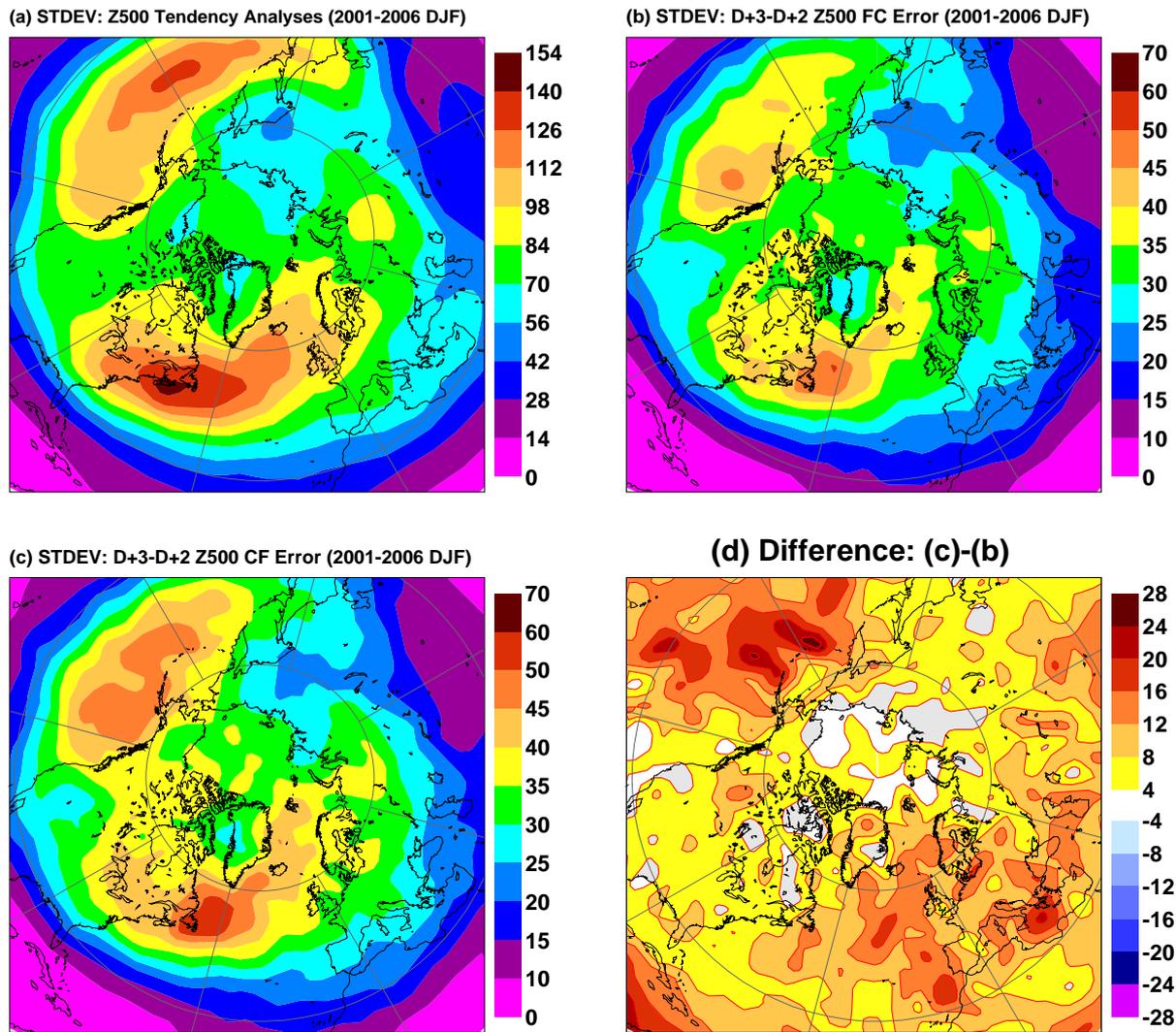


Figure 4: (a) Standard deviation of day-to-day changes (tendency) of Z500 fields (in $m\text{day}^{-1}$) from operational analysis data for winters of the period 2001–06. Also shown is the standard deviation of the Z500 tendency from D+2 to D+3 for operational (b) high-resolution deterministic forecasts and (c) lower-resolution control forecasts of the ECMWF EPS for the same period. (d) Difference between (c) and (b) (in %).

2002). The presence of such “noise” can readily be seen from the year-to-year variations in skill of ERA-40 reforecasts.

A comparison of operational high-resolution deterministic forecasts with lower-resolution control forecasts of the ECMWF EPS for winters of the period 2001–06 is useful for understanding the influence that resolution has. The skill scores shown in Figure 3, however, are representative for the polar regions as a whole. In order to understand whether there are regional differences in the influence that resolution has on forecast skill, high-resolution deterministic and lower-resolution control forecasts have been assessed on a grid point basis (Figure 4). The scores have been computed for Z500 tendencies from one day to the other (for the verifying analysis) and D+2 to D+3 (forecasts) in order to highlight synoptic-scale aspects. More detail on the spectral characteristics of the tendency filter are given by Jung (2005). The efficiency of the filter can be inferred from Figure 4a, which shows the main two storm tracks over the North Pacific and North Atlantic. Independent

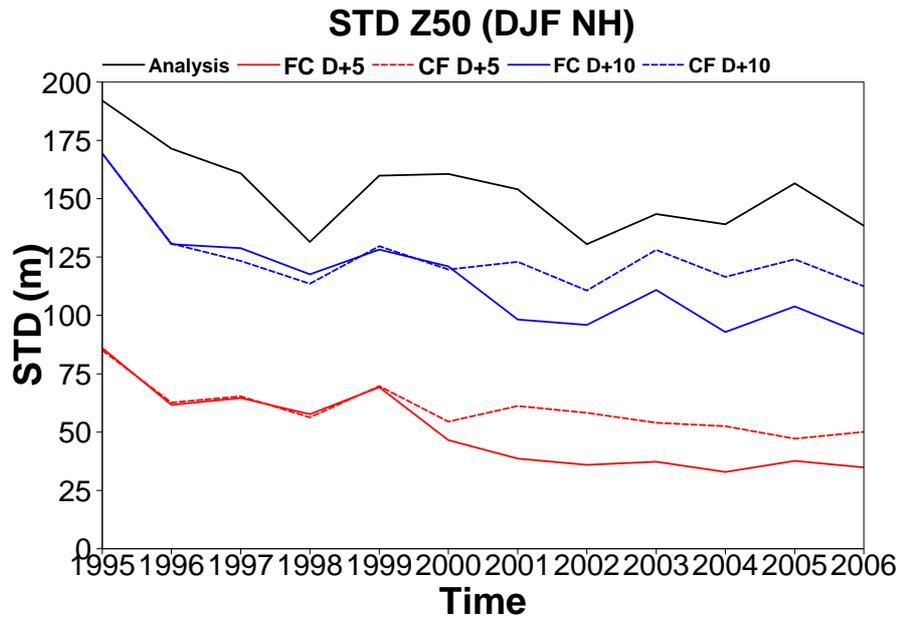


Figure 5: Time series of the temporal mean spatial standard deviation of daily Z50 forecast error over the Northern Hemisphere at D+5 (red lines) and D+10 (blue lines). Two different forecast sets are shown: operational high-resolution deterministic forecasts (solid) and lower-resolution EPS control forecasts (dashed). Also shown is the time series of the mean spatial standard deviation of Z500 fields from operational analyses (black, solid).

of resolution the largest errors in forecasting Z500 tendencies from D+2 to D+3 are found in regions of large synoptic activity. This is particularly true for the North Atlantic. Generally, forecast errors are significantly smaller than observed Z500 tendencies suggesting that there is predictability of synoptic-scale features in the short-range (up to D+3). This conclusion also holds for polar regions. Interestingly, the influence of resolution is most dramatic downstream of the major storm tracks (Figure 4) suggesting that increased resolution is particularly beneficial in the eastern North Atlantic, Europe and the Greenland/Icelandic/Norwegian sea region. Over the Arctic—although beneficial—resolution seems to play a less important role, at least as far as short-range forecasts of synoptic-scale aspects are concerned³. In the medium-range, differences in forecast skill between the high-resolution deterministic and the lower-resolution control forecasts are much more evenly distributed over the Northern Hemisphere (not shown).

So far, the focus has been on the tropospheric circulation. However, pronounced variations can also occur in the high-latitude stratosphere during boreal winter. The most dramatic events are associated with a complete breakdown of the stratospheric polar vortex (see, e.g., Labitzke, 1999, for an overview) and there is evidence that changes in the strength of the stratospheric polar vortex can influence weather in the underlying troposphere (e.g. Baldwin and Dunkerton, 1999; Jung and Barkmeijer, 2006). Therefore, the question arises as to how predictable the stratospheric flow is.

Time series of D+5 and D+10 forecast error of 50 hPa geopotential height fields (Z50) over the Northern Hemisphere during boreal winters of the period 1995–2006 are shown in Figure 5 for the high-resolution deterministic and the lower-resolution EPS control forecast. Until the winter 2000/2001 the performance of both forecasts was comparable. Thereafter, the high-resolution system clearly outperformed the EPS control forecast. Again, the increase in horizontal resolution in autumn 2000 appears to be the key-factor. The influence of

³It should be kept in mind, however, that the lower-resolution T_{L255} model indirectly benefits from the higher resolution. This is because the ensemble is started from the (truncated) high-resolution analysis.

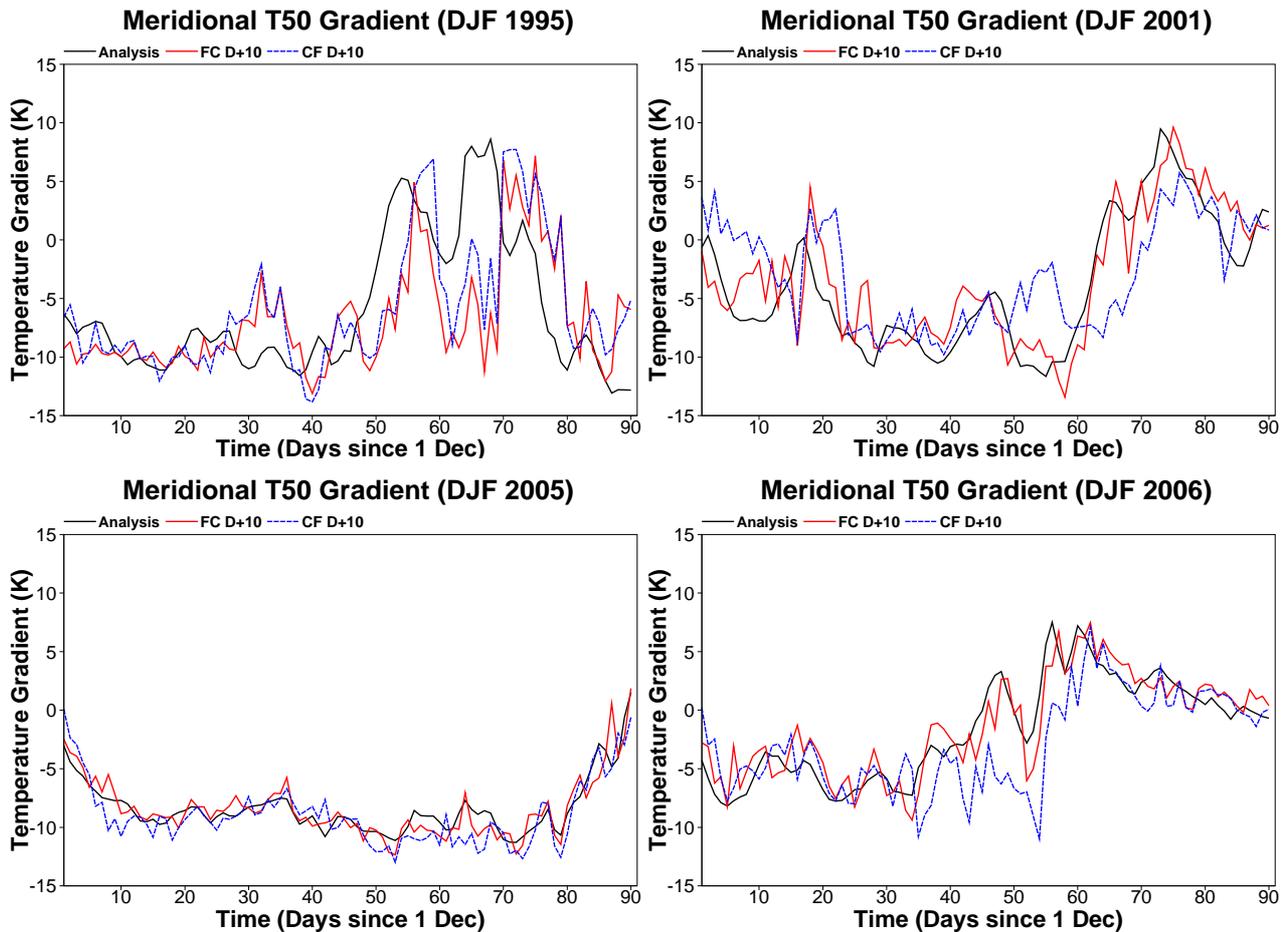


Figure 6: Time series of the difference of zonally-averaged D+10 temperature forecasts at the 50 hPa level between the polar cap (75°N–90°N) and the mid-latitudes (50°N–65°N) for four different winters: (a) 1994/95, (b) 2001/02, (c) 2004/05 and (d) 2005/06. Results are shown for high-resolution deterministic (red) and lower-resolution EPS control forecasts. Also shown is the verifying operational analysis.

the increase in vertical resolution in autumn 1999, although beneficial (see D+5 forecast in winter 1999/2000), seems to be less dramatic.

In order to get a better understanding of the predictability of stratospheric warming events, D+10 forecast of the temperature differences at the 50 hPa level (ΔT_{50}) between high-latitudes and low-latitudes have been computed for selected winters of the period 1995–2006 (Figure 6). Positive values for ΔT_{50} are indicative of a breakdown of the stratospheric polar vortex (“stratospheric warming events”). In 1994/95 both models perform relatively poorly in predicting the warming events 10 days ahead. In 2000/01 and 2005/06, on the other hand, the high-resolution model clearly outperforms the lower-resolution model. In fact, the high-resolution system shows considerable skill at D+10 in predicting stratospheric warming events in more recent years, which is mostly likely due to the increase in resolution that took place in autumn 2000 (see also Figure 5). For periods in which the temperature difference is close to its “normal” value (e.g., Figure 6c) there is relatively little variability and resolution seems to play a minor role.

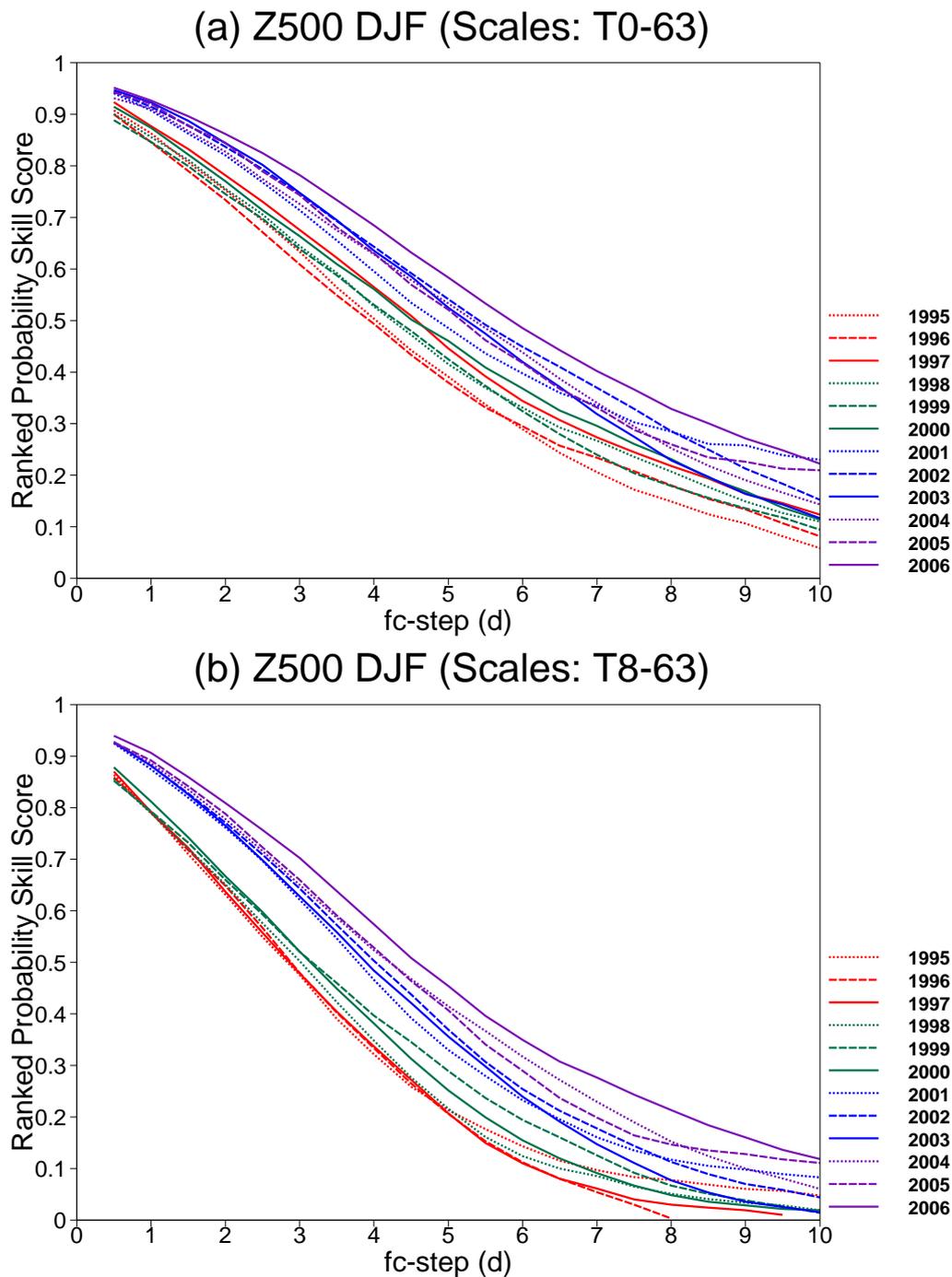


Figure 7: Ranked probability skill score as a function of forecast step (in days) for operational ECMWF ensemble forecasts of polar regions (north of 65°N). Results are based on 5 categories separated by the 20th, 40th, 60th and 80th percentile. Verification has been carried out for (a) all total wavenumbers from 0–63 and (b) total wavenumber from 8–63 only (synoptic and sub-synoptic features). Scores are shown separately for individual winters (DJF) from 1995 to 2006 (see legends). All results are based on a total of 50 ensemble members from 1997 and 31 members for 1995–1996..

3.2.2 Probabilistic forecast skill

As mentioned in the Introduction, ensemble forecasts are carried out at ECMWF since the end of 1992. The ECMWF EPS makes use of the singular vector technique in order to perturb the initial conditions. It is well known that these singular vectors tend to occur primarily in the major baroclinic zones in the *mid-latitudes*. The question, therefore, arises how well the ECMWF EPS performs in polar regions. Ranked probability skill scores (RPSS) for probabilistic Z500 forecasts (5 different categories were taken into account) of the ECMWF EPS in polar regions (north of 65°N) are separately shown in Figure 7 for all winters of the period 1995–2006⁴. Moreover, a distinction is made between probabilistic Z500 forecasts taking into account all spatial scales up to a total wavenumber of 63 (Figure 7a) and those for synoptic and sub-synoptic spatial scales only (total wavenumber 8–63, Figure 7b). The most important conclusion, which can be drawn from Figure 7, is that the EPS shows substantial probabilistic forecast skill in polar regions. Even on synoptic and sub-synoptic scales, recent versions of the EPS produce skillful probabilistic Z500 forecasts well into the far medium-range. Moreover, it turns out that the skill of the EPS has increased substantially during the 12-year period from 1995 to 2006. This improvement seems to be even more pronounced if the focus is on synoptic and sub-synoptic spatial scales. The same skill of D+4 forecasts of synoptic and sub-synoptic Z500 features in 1995, for example, is nowadays achieved at D+7—a gain in predictability of about 3 days! The largest improvement took place in late 2000 with the introduction of a higher-resolution model (T_L255) used to carry out the nonlinear integrations. It appears that individual changes made to the way the initial conditions are computed, albeit beneficial, had a smaller influence on the probabilistic forecast skill in polar regions. The fact that the skill did not increase monotonically from year to year highlights the importance of flow-dependent aspects of probabilistic forecast skill. Finally, it is worth mentioning that the forecast skill was highest during the last winter (2005/06). It is likely that this is due to a further increase in resolution that was introduced in early February 2006 (T_L399 with 62 levels in the vertical).

3.2.3 Sensitivity of forecast error to initial conditions

Adjoint sensitivity gradients of D+2 forecast errors north of 70°N to initial perturbations have been computed every fifth day for the winters of 2001/02 and 2004/05 (see Methods section for details). Mean values of vertically integrated (taking into account the thickness of the layers) gradients with respect to relative vorticity are shown in Figure 8a,b. In areas where large values occur, initial vorticity perturbations potentially have a large impact on subsequent forecasts in polar regions (provided that they project onto the gradients). Perhaps not too surprisingly the gradients have a strong local component, that is, large gradients are found in polar regions. However, there is also considerable sensitivity of D+2 forecast error in polar regions to small initial perturbations in mid-latitudes, particularly in the North Atlantic region. Although there is also some sensitivity to initial perturbations in the North Pacific region, on average this sensitivity is much lower than that found in the North Atlantic region (notice the nonlinear contour scale in Figure 8). It is likely that this is a result of the more zonal flow in the North Pacific compared to the North Atlantic region (Figure 8c,d). Furthermore, there are pronounced interannual variations in the sensitivity of D+2 forecast error in polar regions to both local and remote initial perturbations. The sensitivity gradients in the eastern North Atlantic, for example, were much higher in 2001/02 than those in 2004/05.

Vertically integrated absolute values of the sensitivity gradient of D+2 forecast error north of 70°N to tropospheric initial vorticity perturbations for different cases in winter 2001/02 are shown in Figure 9. The atmospheric flow on 29 November 2001, for example, was associated with extremely high sensitivity of the subsequent forecast in polar regions to vorticity perturbations in the eastern North Atlantic (Figure 9a). The

⁴The reference forecast to define the skill scores is a probabilistic forecast based on a climatological distribution which is determined from ERA-40 analyses (1979–2001). The climatological distribution is computed for every grid point and verification day.

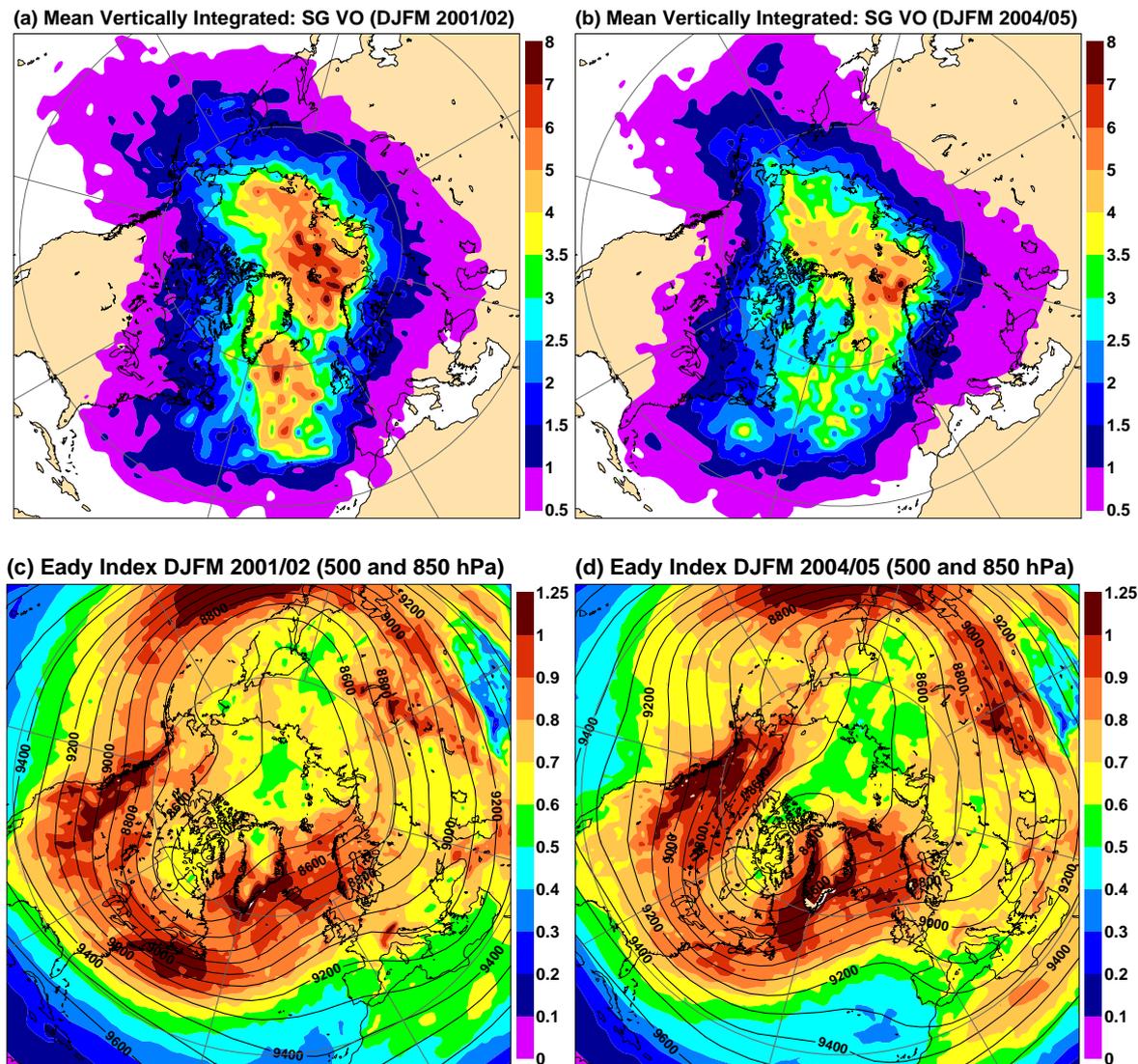


Figure 8: Mean of vertically integrated absolute values of the sensitivity gradient of D+2 forecast error north of 70°N to initial vorticity perturbations throughout the troposphere for two winters (December–March): (a) 2001/02 and (b) 2004/05. Also shown are winter mean values of the Eady index (in day⁻¹, colour shading) and 300 hPa geopotential height fields (contour interval is 100 m) for the two winters (c) 2001/02 and (d) 2004/05. The Eady index has been computed using the two levels 500 and 850 hPa.

high sensitivity in this region for this particular case can be explained by strong baroclinicity and the fact that during the following 2-day period a very strong northwesterly flow developed over the northeastern North Atlantic, which extended all the way into the Arctic (not shown). A case of very low sensitivity (of forecasts in polar regions), for example, occurred on 28 January 2002 (Figure 9e). The low sensitivity to vorticity perturbations in the North Atlantic region for this case is consistent with the anomalously southerly position of the polar jet stream. Another interesting feature revealed by the adjoint sensitivity computations is that the sensitivity gradients in the mid-latitudes occur highly intermittent. In fact, the high mean values of the sensitivity gradients in the eastern North Atlantic (Figure 8a) are primarily the result of two strongly sensitive cases (29 November and 9 December 2001). Sensitivity gradients in the Arctic, on the other hand, tend to have smaller values and occur more frequently (Figure 9). In summary, it turns out that the sensitivity of forecasts in polar

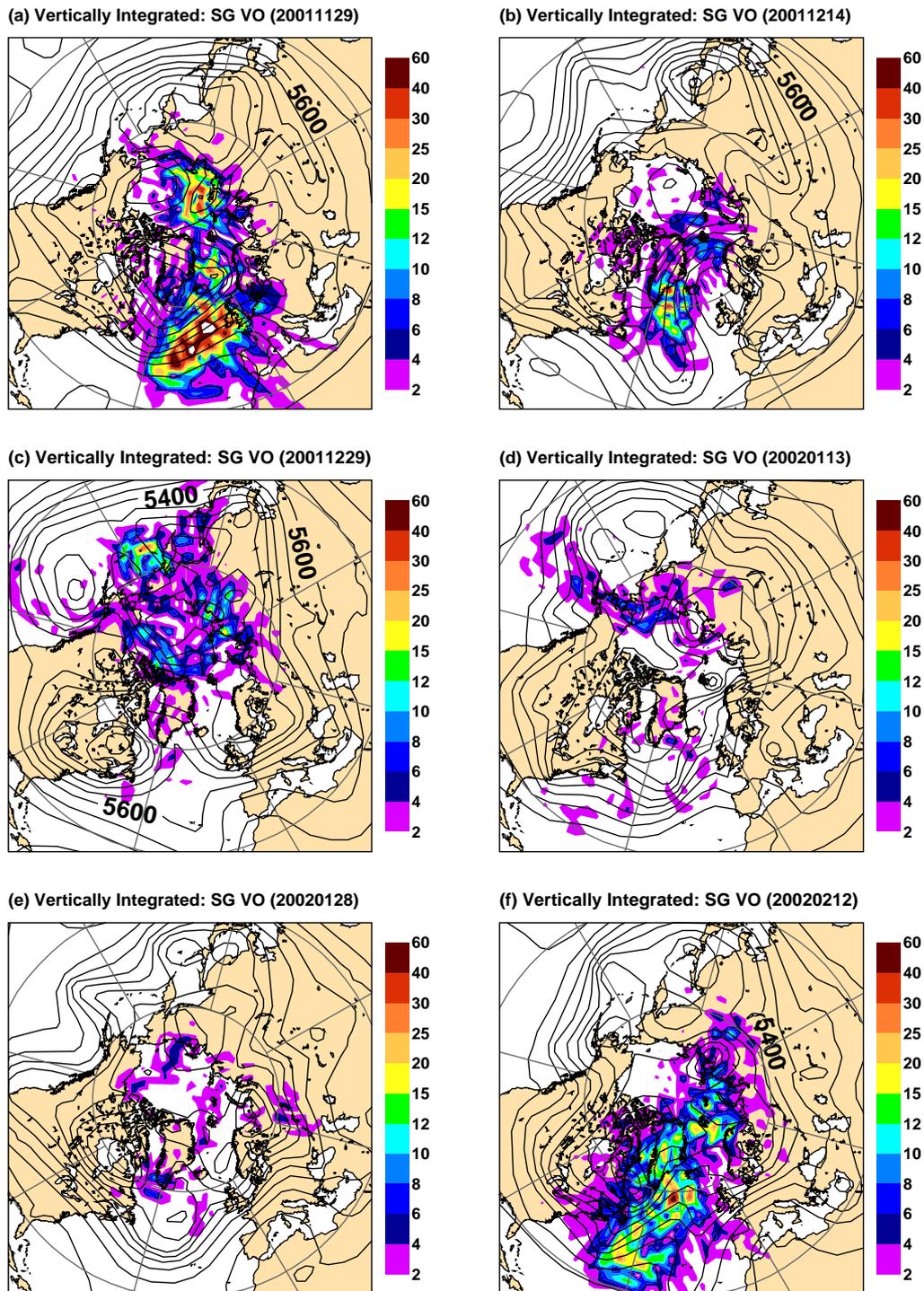


Figure 9: Vertically integrated absolute values of the sensitivity gradient of D+2 forecast error north of 70°N to initial vorticity perturbations throughout the troposphere for different dates in winter 2001/02: (a) 29 November 2001, (b) 12 December 2001, (c) 29 December 2001, (d) 13 January 2002, (e) 28 January 2002, and (f) 12 February 2002. Also shown are 500 hPa geopotential height fields (contour interval is 100 m) from the operational ECMWF analysis. Notice the nonlinear contouring for the integrated sensitivity gradient.

regions to initial perturbations, particularly to those in the North Atlantic region, is highly flow-dependent. The results from the adjoint sensitivity computations, therefore, highlight the importance of using ensemble prediction systems for weather forecasting in polar regions.

3.3 Seasonal integrations

3.3.1 Synoptic activity and horizontal resolution

Seasonal integrations (see Data and Methods section) were carried out with model version 31R1 (used operationally at ECMWF since 12 September 2006) at two different horizontal resolutions and with 91 levels in the vertical. The first horizontal resolution used is T_L95 (about 180 km) and corresponds to what is typically used in climate research. The second resolution considered, T_L511 (about 40 km), corresponds to what is typically employed by the NWP community. For the two model versions, 16 winters (1991–2006) were simulated starting on 1 November for each of the years. The results have been diagnosed for the period December to March only in order to remove transient effects.

In the following, synoptic activity is defined as the standard deviation of of highpass-filtered (retaining variability on time scales shorter than 8 days) Z500 time series. The difference in synoptic activity between the T_L95 model and ERA-40 is shown in Figure 10a. Evidently, the ECMWF model—if run at a horizontal resolution typically used in climate research—underestimates the observed synoptic activity by about 15% in the high-latitudes of both hemisphere (see also Jung et al., 2006). The level of synoptic activity becomes much more realistic if the horizontal resolution is increased to T_L511 (Figure 10a,b). The most likely explanation is that the realistic simulation of relatively small-scale low-pressure systems found at high latitudes (e.g. Simmonds, 2000) requires relative high resolution; a model run at a resolution of T_L95 might be simply too dissipative on spatial scales relevant for the life cycle of high-latitude cyclonic systems.

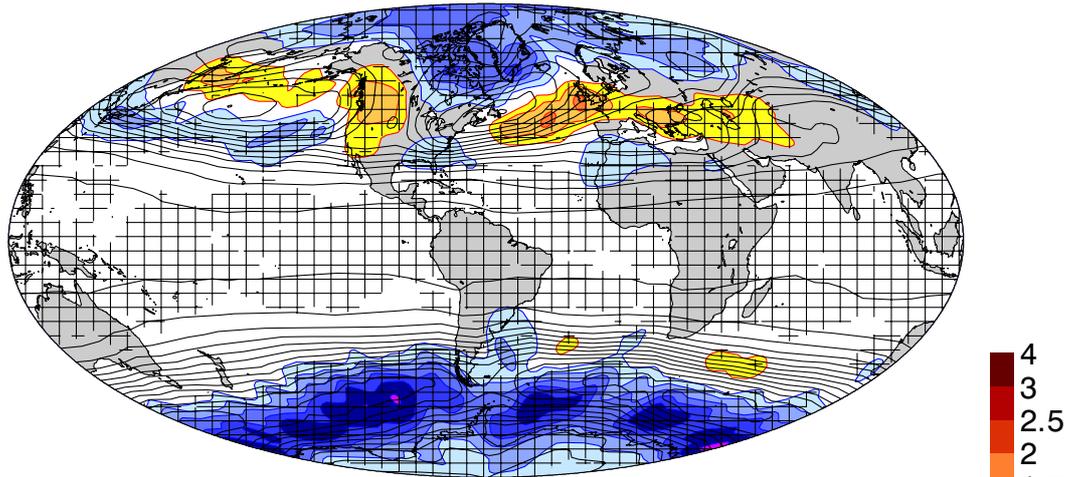
3.3.2 Stratosphere-troposphere link

In the following, the stratosphere-troposphere link—a potentially important source of extended-range predictability in the Euro-Atlantic region (e.g. Baldwin and Dunkerton, 2001; Baldwin et al., 2003)—will be investigated statistically using ERA-40 data and seasonal integrations with prescribed SST fields. All forecast experiments were carried out using model version 29R1 (in operational use from 5 April to 28 June 2005) with a horizontal resolution of T_L95 and with 60 levels in the vertical. The upper-most model level is located at 0.1 hPa. Results will be shown for 40 winters (December through March) of the period 1962–2001. The mean annual cycle has been removed prior to the computations.

To begin with, empirical orthogonal function (EOF) analysis has been applied separately to analyzed and simulated daily anomalies of 10 hPa geopotential height fields (Z10, hereafter). The first EOFs from reanalysis and model data are very similar (not shown) and reflect changes in the strength of the stratospheric polar vortex. Next, the first principal components (PCs) were used to identify weak vortex cases. The *onset* of weak vortex cases is defined as that day at which PC1 exceeds one standard deviation (positive values for PC1 correspond to an anomalous weak vortex). To qualify as a day of weak vortex onset it is further required that PC1 remained below (above) this threshold for the previous (following) 10 days.

Figure 11 shows composites of Z10 (upper row) and Z1000 (lower row) anomalies, 1–20 (left column) and 21–40 days (right column) after the onset of a weak stratospheric polar vortex for ERA-40 data. Consistent with previous observational studies (Baldwin and Dunkerton, 1999, 2001) breakdowns of the stratospheric polar vortex in ERA-40 tend to be followed (after some delay) by a significant near-surface tropospheric response resembling the negative phase of the Arctic Oscillation/North Atlantic Oscillation (Figure 11).

(a) Synoptic Z500 Activity: Difference T95-ERA40 (12-3 1991-2006)



(b) Synoptic Z500 Activity: Difference T511-T95 (12-3 1991-2006)

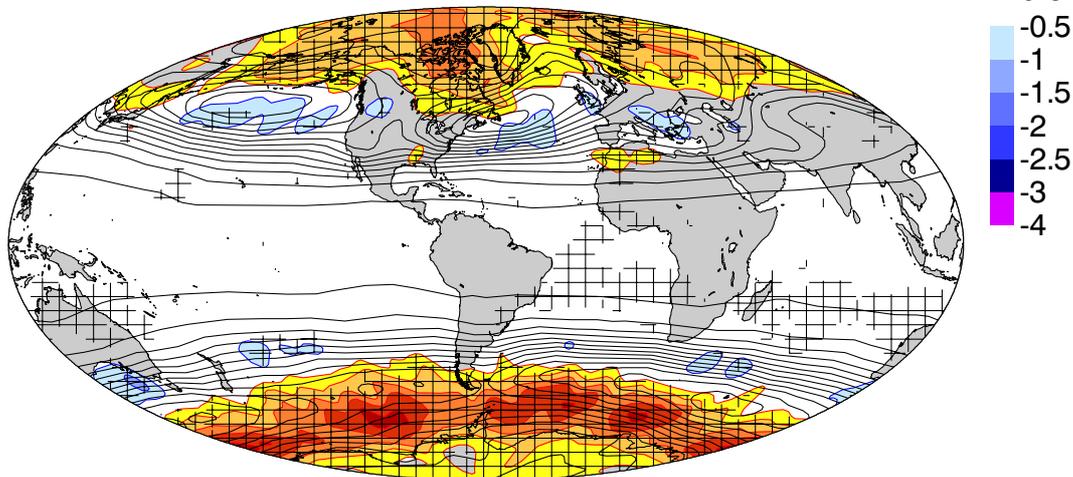


Figure 10: Difference in the standard deviation (shading in dam) of high-pass filtered Z500 fields between (a) the TL95L91 version of the ECMWF model and ERA-40 data and the (b) T_L511L91 and T95_L91 version of the ECMWF model for winters (December through March) of the period 1991–2006. Also shown is the long-term mean standard deviation of highpass-filtered Z500 (contour interval is 1 dam) from (a) ERA-40 and (b) the T_L95L91 version of the ECMWF model. Differences statistically significant at the 95% levels are hatched.

It turns out that even at a relatively coarse horizontal resolution of T_L95, the ECMWF model is capable of reproducing the observed stratosphere-troposphere link (Figure 12). This finding is consistent with previous modelling studies. It should be pointed out, however, that unlike in previous studies (e.g. Boville, 1984; Polvani and Kushner, 2002; Jung and Barkmeijer, 2006), here no model forcing (apart from the prescribed observed SST fields) is applied during the course of the integration, that is, the stratosphere-troposphere link appears naturally in free seasonal integrations of the ECMWF model.

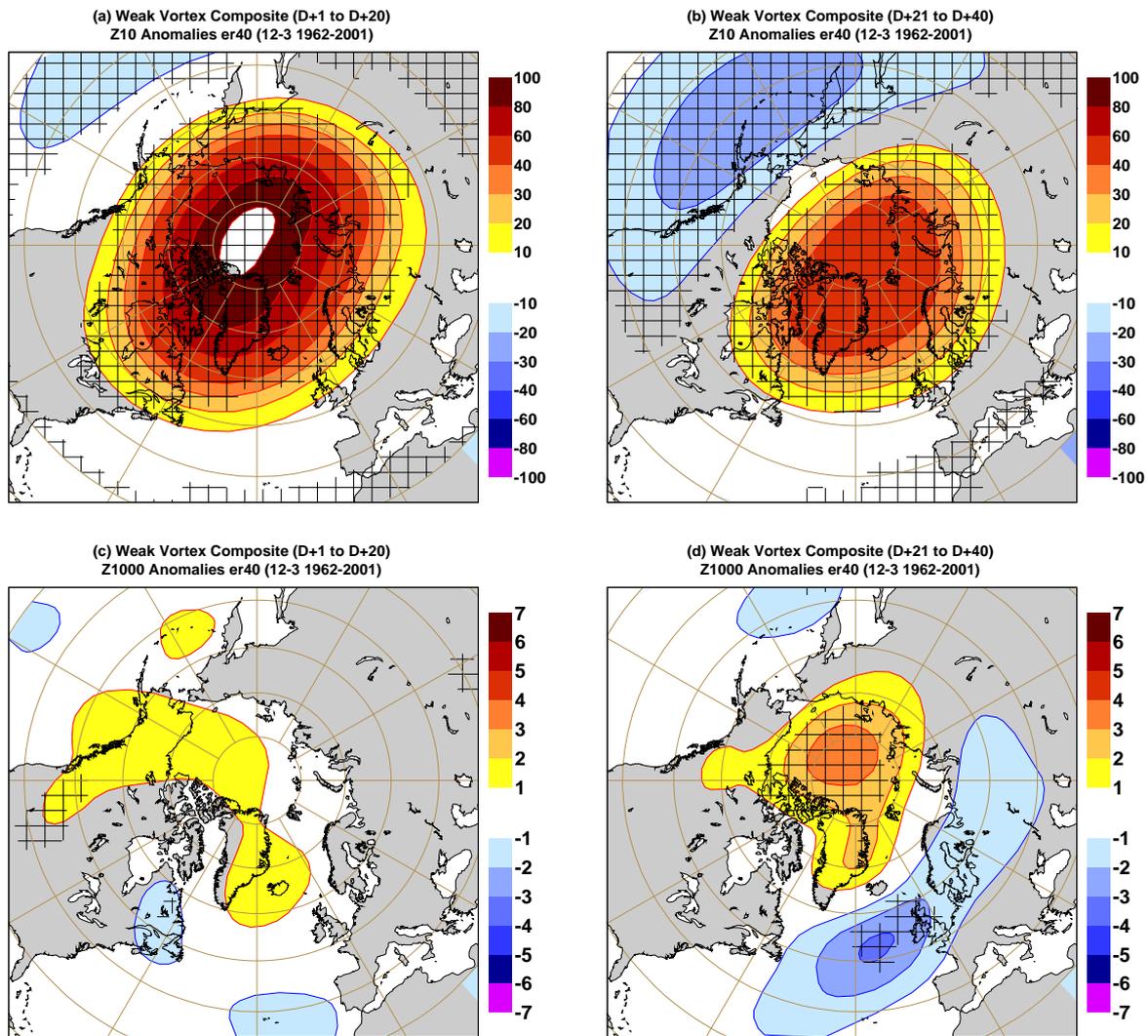


Figure 11: Mean (a)–(b) Z10 and (c)–(d) Z1000 anomalies (in m) following the onset of weak stratospheric polar vortex events (see text for details) for ERA-40 data: 1–20 days (left column) and 21–40 days (right column) after the onset. Statistically significant (at the 95% confidence level using a two-sided Student’s *t*-test) anomalies are hatched. Results are based on 40 winters of the period 1962–2001.

4 Discussion

The realism of analysed synoptic-scale features in geopotential height fields in polar regions during boreal winter (austral summer) has been investigated. From the similarity of operational and reanalysed Z1000 fields in winter 2001/02 it is argued that synoptic-scale features are realistically represented by state-of-the-art data assimilation systems (using an observing system similar to that available in 2001/02). This is particularly true for the Northern Hemisphere for which more conventional observations are available. Our results suggest that geopotential height fields in the Arctic from ERA-40 are of sufficient quality to obtain realistic climatologies of synoptic-scale, polar cyclone characteristics (e.g. [Serreze and Barry, 1988](#); [Jung et al., 2006](#))

The evolution of forecast error in polar regions of both hemispheres is very similar to changes found in the Northern and Southern Hemisphere as a whole (see [Simmons and Hollingsworth, 2002](#); [Simmons, 2006](#), for

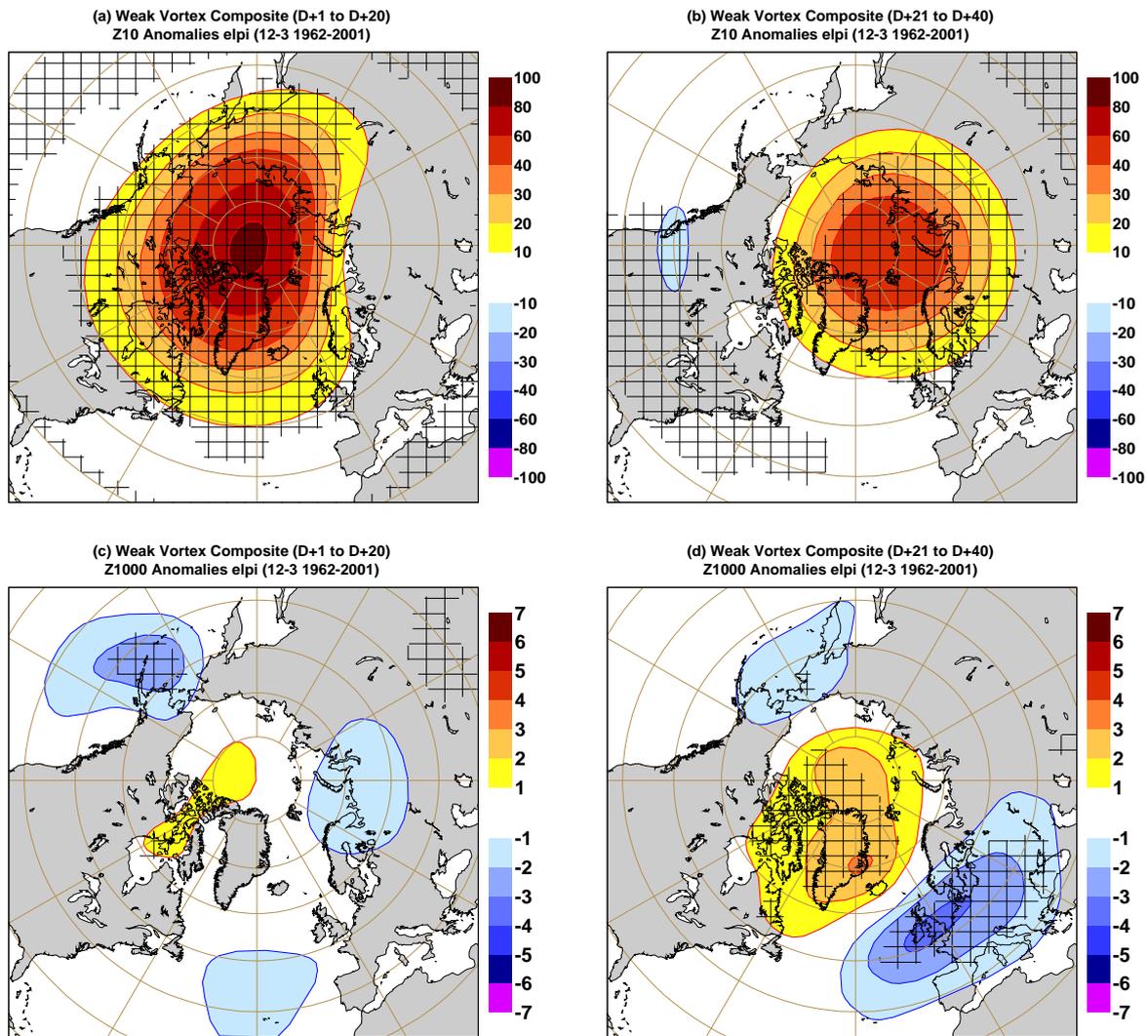


Figure 12: Same as in Figure 11, but for seasonal forecasts carried out with the ECMWF model (T_L95 with 60 levels in the vertical).

hemispheric scores) suggesting that the beneficial impact of model changes (e.g. increased model resolution in autumn 2000) in polar regions is similar to that found in mid-latitudes⁵. The increase in horizontal resolution of the analysis and deterministic forecasting system to T_L511 in autumn 2000 was also beneficial in terms of medium-range forecast skill of stratospheric warming events. In fact, our results suggest that stratospheric warming events are predictable well beyond D+10.

Based on results obtained by the adjoint technique it has been shown that the sensitivity of forecast error in polar region to initial perturbations, particularly to those in the North Atlantic region, is highly flow-dependent. From this it has been argued that the use ensemble prediction systems in polar regions is crucial. An assessment of the quality of the ECMWF EPS model reveals considerable probabilistic medium-range forecast skill in polar regions of the Northern Hemisphere. This is even true for synoptic-scale aspects of the flow. The increase in probabilistic predictability in polar regions during the the last ten years amount to as much as 2 days. The increase in horizontal resolution to T_L255 (about 80 km) in autumn 2000 was the single most beneficial model

⁵Notice, that hemispheric scores are dominated by forecast error in mid-latitudes.

change, particularly for the quality of probabilistic ensemble forecasts of synoptic-scale features. The best ever probabilistic forecast performance in both the short-range and medium-range was achieved last winter (2005/06), suggesting that the increase in resolution (from T_L255L40 to T_L399L62) in early February 2006 has been beneficial.

Seasonal integrations with the ECMWF model have been carried out in order to investigate the impact that horizontal resolution has on the level synoptic activity (transient eddies) in polar regions (see also Jung et al., 2006). It has been found that the T_L95-version of the model (“climate resolution”) substantially underestimates synoptic activity in polar regions. The T_L511-version of the model (“NWP resolution”), on the other hand, produces very realistic levels of synoptic activity in polar regions. In the mid-latitudes, the sensitivity of transient wave activity to resolution is much less pronounced (at least for the two resolution considered). The above results are consistent with the decrease of the size of extratropical cyclones with increasing latitude (Simmonds, 2000), that is, the simulation of smaller scale systems in polar regions requires higher horizontal resolution. It would be interesting to extend the investigation presented in this study to polar lows, which represent particularly small-scale and high impact phenomena in high latitudes (Rasmussen and Turner, 2003). However, the development of polar low tracking software (Harold et al., 1999; Bracegirdle, 2006) is still in its infancy, and it is questionable how well relatively low-resolution reanalyses (e.g., NCEP-NCAR and ERA-40) capture observed polar lows. More work is clearly necessary before any conclusive assessment of the realism of high-resolution atmospheric models in simulating polar lows can be carried out.

Finally, it has been shown that the observed “downward propagation” of stratospheric polar vortex anomalies into the troposphere (Baldwin and Dunkerton, 1999, 2001) is well captured in seasonal integrations with the T_L95L60 version of the ECMWF model. The fact that the stratosphere-troposphere link exists in “free” (i.e., no artificial forcing is applied) seasonal integrations shows that operational seasonal and in particular monthly forecasts (Vitart, 2004) with the ECMWF model may benefit from the implied extended-range predictability (see Baldwin et al., 2003; Charlton et al., 2003, for details).

Acknowledgements

We are grateful to numerous colleagues of the research department at ECMWF for useful discussions on various aspects of the topics presented in this study. We acknowledge Tim Palmer’s and Martin Miller’s support in carrying out the very high-resolution seasonal integrations. Rob Hine helped improving the quality of some of the figures.

References

- Baldwin, M. P. and T. J. Dunkerton (1999). Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *J. Geophys. Res.* *104*, 30937–30946.
- Baldwin, M. P. and T. J. Dunkerton (2001). Stratospheric harbingers of anomalous weather regimes. *Science* *294*, 581–584.
- Baldwin, M. P., D. B. Stephenson, D. W. J. Thompson, T. J. Dunkerton, A. J. Charlton, and A. O’Neill (2003). Stratospheric memory and skill of extended-range weather forecasts. *Science* *301*, 636–640.
- Boville, B. A. (1984). The influence of the polar night jet in the tropospheric circulation in a GCM. *J. Atmos. Sci.* *41*, 1132–1142.

- Bracegirdle, T. (2006). *The role of convection in the intensification of polar lows*. Ph. D. thesis, Department of Meteorology, The University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK.
- Buizza, R. (1994). Localization of optimal perturbations using a projection operator. *Quart. J. Roy. Meteor. Soc.* 120, 1647–1682.
- Buizza, R. (2006). The ECMWF Ensemble Prediction System. In T. N. Palmer and R. Hagedorn (Eds.), *Predictability of Weather and Climate*, pp. 459–488. Cambridge University Press.
- Buizza, R., M. Miller, and T. N. Palmer (1999). Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.* 125, 2887–2908.
- Buizza, R. and T. Palmer (1995). The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.* 52, 1434–1456.
- Charlton, A. J., A. O. O’Neill, D. B. Stephenson, W. A. Lahoz, and M. P. Baldwin (2003). Can knowledge of the state of the stratosphere be used to improve statistical forecasts of the troposphere? *Quart. J. Roy. Meteor. Soc.* 129, 3205–3225.
- Ehrendorfer, M., R. Errico, and K. Raeder (1999). Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.* 56, 1627–1648.
- Errico, R. M. (1997). What is an adjoint model? *Bull. Amer. Meteor. Soc.* 78, 2577–2591.
- Ferranti, L., E. Klinker, A. Hollingsworth, and B. J. Hoskins (2002). Diagnosis of systematic forecast errors dependent on flow pattern. *Quart. J. Roy. Meteor. Soc.* 128, 1623–1640.
- Harold, J. M., G. R. Bigg, and J. Turner (1999). Mesocyclone activity over the North-East Atlantic. Part 1: vortex distribution and variability. *Int. J. Climatol.* 19, 1187–1204.
- Jung, T. (2005). Systematic errors of the atmospheric circulation in the ECMWF forecasting system. *Quart. J. Roy. Meteor. Soc.* 131, 1045–1073.
- Jung, T. and J. Barkmeijer (2006). Sensitivity of the tropospheric circulation to changes in the strength of the stratospheric polar vortex. *Mon. Wea. Rev.* 134, 2191–2207.
- Jung, T., S. K. Gulev, I. Rudeva, and V. Soloviov (2006). Sensitivity of extratropical cyclone characteristics to horizontal resolution in the ECMWF model. *Quart. J. Roy. Meteor. Soc.* 132, 1839–1857.
- Jung, T. and A. M. Tompkins (2003). Systematic errors in the ECMWF forecasting system. Technical Report 422, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK. Available at <http://www.ecmwf.int/publications/>.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph (1996). The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.* 77, 437–471.
- Labitzke, K. (1999). *The Stratosphere: Phenomena, History, and Relevance*. Springer Verlag. 179 pp.
- Mahfouf, J.-F. (1999). Influence of physical processes on the tangent-linear approximation. *Tellus 51 A*, 147–166.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliaigis (1996). The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.* 122, 73–119.

- Morss, R. E., O. V. Wilhelmi, M. W. Downton, and E. Grunfest (2005). Flood risk, uncertainty, and scientific information for decision making. *Bull. Amer. Meteor. Soc.* 86, 1593–1601.
- Polvani, L. M. and P. J. Kushner (2002). Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophys. Res. Lett.* 29, 10.1029/2001GL014284.
- Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth (1996). Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.* 122, 121–150.
- Rasmussen, E. A. and J. Turner (2003). *Polar Lows*. Cambridge University Press.
- Serreze, M. and R. Barry (1988). Synoptic Activity in the Arctic Basin, 1979–85. *J. Climate* 1, 1276–1295.
- Simmonds, I. (2000). Size changes over the life of sea level cyclones in the NCEP reanalysis. *Mon. Wea. Rev.* 128, 4118–4125.
- Simmons, A. J. (2006). Observations, assimilation and the improvement of global weather prediction—some results from operational forecasting and ERA-40. In T. N. Palmer and R. Hagedorn (Eds.), *Predictability of Weather and Climate*, pp. 428–458. Cambridge University Press.
- Simmons, A. J. and A. Hollingsworth (2002). Some aspects of the improvement of skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.* 128, 647–677.
- Uppala, S., P. W. Kallberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. McNally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen (2005). The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.* 131, 2961–3012.
- Vitart, F. (2004). Monthly forecasting at ECMWF. *Mon. Wea. Rev.* 132, 2761–2779.