TECHNICAL MEMORANDUM

321

Forecasting system performance in summer 1999. Part 1 - Diagnostics related to the forecast performance during spring and summer 1999

E. Klinker and L. Ferranti

Research Department

September 2000

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

Forecasting system performance in summer 1999 - Part 1 Diagnostics related to the forecast performance during spring and summer 1999

by **Ernst Klinker and Laura Ferranti**

Abstract

The poor performance of the operational forecasting system during summer 1999 motivated this diagnostic study of forecast errors. Larger than normal forecast errors in the medium range were found over the North Atlantic, the Norwegian Sea and parts of Northern Europe. A large portion of the error increase could be linked to unusual flow conditions during summer 1999. Measuring the baroclinic instability of the flow with the vertical wind shear and the static stability (Eady Index), a substantial increase of error growth could be expected from the more unstable conditions during summer 1999 compared to summer 1998. A confirmation of faster error growth in 1999 has been found from the adjoint calculations that identifies the sensitivity of forecast errors to initial conditions. Based on these sensitivity integrations it has been shown that a large part of the error increase in summer 1999 has been due to fast growing analysis errors.

By applying the Singular Value Decomposition (SVD) technique to pairs of forecast errors and analyses the flow dependent part of increased forecast errors in 1999 could be identified. Larger errors during the summer of 1999 occur mostly during episodes of a strong westerly flow, of which more episodes occurred during summer 1999 than summer 1998. Flow dependent error magnitudes show a clear picture of errors propagating from North Canada and area around Greenland in the short range forecasts into the North Atlantic and Northern Europe by day-5.

North Canada and Greenland are also areas from where flow-dependent error differences between ECMWF and UK Met. Office forecasts evolve into larger errors over Northern Europe for the ECMWF system in the medium range. Further diagnostics based on regional sensitivity again emphasize the importance of North and North-West Canada for the error evolution over Northern Europe in the medium range.

Some diagnostics have been repeated on parallel runs for the new cycle (21r4) of the forecasting system that became operational in October 1999. A noticeable improvement was found in particular for situations of a strong baroclinic westerly flow over the Atlantic. As a consequence the North Atlantic and Northern Europe are the areas where medium range forecast errors have been reduced most with the new model cycle.

Tracing the differences between forecast errors of the new and old system from the medium range back to the very short range, including the 9 hour forecast in the data assimilation, it appears that the update of orography in cycle 21r4 in the Greenland area played an important part in the reduction of forecast errors.

1. Introduction

The poor performance of the operational forecasting system during summer 1999 has been the motivation for an extensive diagnostic study of forecast errors. The problems of comparing forecast errors of seasons from different years are well known. Modifications of the forecasting system and differences in the flow conditions both have an important influence on forecast performance. We would like to separate the two effects as far as possible by diagnosing the effect of different flow anomalies but also by comparing operational forecasts with forecasts from another Centre (UK Met. Office). Further diagnostic support has been obtained by comparing operational forecasts with forecasts from an improved version (higher cycle) of the forecasting system that has been run in parallel for the same season.

It was felt that, for a number of error statistics consistent results could only obtained by extending the seasons beyond the standard summer months of June, July, and August. In particular results for the diagnosis of flow dependent errors using the Singular Value Decomposition Technique (SVD) were much improved with longer time series. Therefore the summer seasons are defined to include the months from May to September.

2. Verification scores

The performance of the forecasts during the last 4 summers measured in terms of rms-errors for the Northern Hemisphere shows quite a variation from year to year. The medium range forecasts in the two summers 1999 and 1997 have larger errors than in the summers of 1998 and 1996. As the maximum performance difference is found between the two consecutive summers 1999 and 1998 a large part of this investigation will focus on the comparison between these two summers. For the performance over the European area as defined in the ECMWF verification system (35N to 75N, 12.5W to 42.5E) only a small degradation in 1999 compared to previous years has been found (not shown). Most of the poor performance in 1999 is limited to Northern Europe as shown in the scores (Fig. 2) for that area (55N to 75N, 5E to 35E). Whereas the day-1 errors are the smallest of all in 1999 (10.3 m compared to a range of 10.5 to 12.4 m in previous summers), after day 2 the error growth is so much larger in 1999 that by day-6 the rms errors are more than 10% larger than errors in the 1998 and 1997 summers. The relatively large error growth in the medium range for 1999 is likely to be strongly influenced by error propagation from further upstream.

3. Geographical distribution of forecast errors

In this section rms errors are calculated differently than in the operational verification. Whereas the rms errors in Fig. 1 and 2 represent a time average of spatial Northern Hemisphere rms errors, the rms error maps in Fig. 3 are constructed from time series of errors at each grid point. However, the geographical distribution of rms error differences between summer 1999 and 1998 (shown in Fig. 3 for forecast ranges of 24, 48, 72 and 120 hours) confirms the result from the score graphs (Fig. 1 and Fig. 2). For all ranges shown the forecasts in summer 1999 have generally larger errors than in summer 1998 at middle and high latitudes with the exception of Central Europe. In the Northern Hemisphere statistics these areas of larger errors dominate over low latitude regions where the errors in 1999 are smaller.

The evolution of European forecast error differences is of particular interest. Over most parts of Northern Europe short range forecasts at day-1 are noticeably better in 1999 than in 1998. It seems that an eastward propagation of large errors in later forecasts pushes the zero line between larger and smaller errors into Northern Europe in subsequent forecasts. By day-5 the largest error differences in the Northern Hemisphere are found over the Norwegian Sea which explains most of the the additional error growths in 1999 shown in Fig. 2.

In the comparison of the two summers it is important to establish to what extent the error differences could be due to different flow conditions. A spectral analysis of the 500 hPa height field in the analysis shows that there is a higher level of baroclinic wave activity in 1999 than in 1998. A suitable measure to link eddy activity to the baroclinicity of the flow is given by the Eady index that is proportional to the vertical wind-shear and the static stability of the flow. Typical values during the summer suggest maximum growth rates of baroclinic waves of 0.5 per day (corresponding to an e-folding time of 2 days) in mid-latitudes.

The maximum baroclinic growth rates based on the Eady index have been calculated for the two summer seasons. Fig. 4 shows the percentage increase (decrease) in summer 1999 with respect to summer 1998 growth rates with orange to red colours showing an increase of baroclinic instability. From the result we notice that the flow pattern in summer 1999 is characterized by more unstable conditions in a mid-latitude band compared to summer 1998 as shown by a percentage increase of the summer 1999 values. In particular

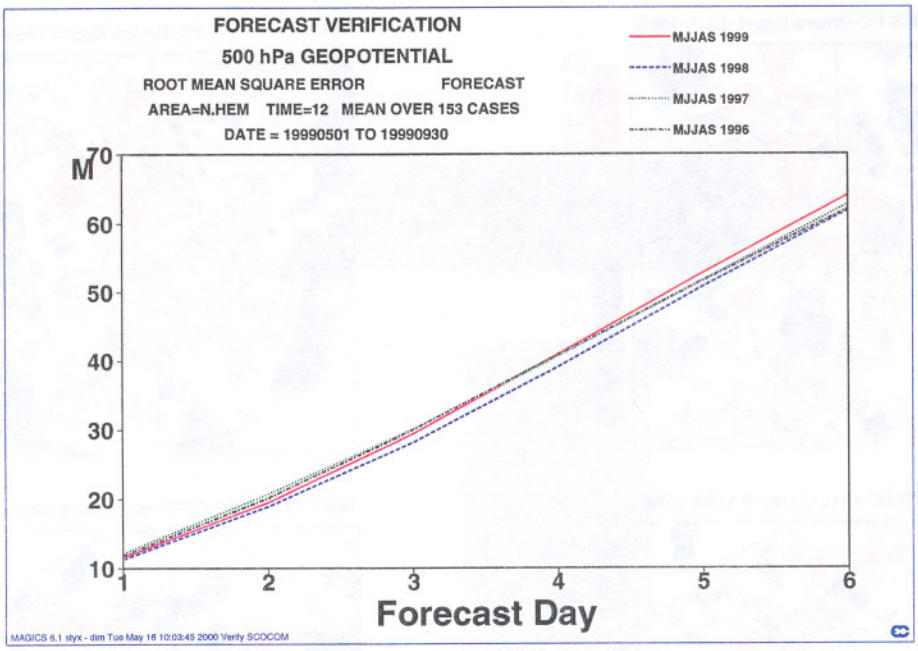

Figure 1: Root mean square forecast errors for 4 summer seasons over the Northern Hemisphere. Parameter: 500 hPa geopotential height. Units: m

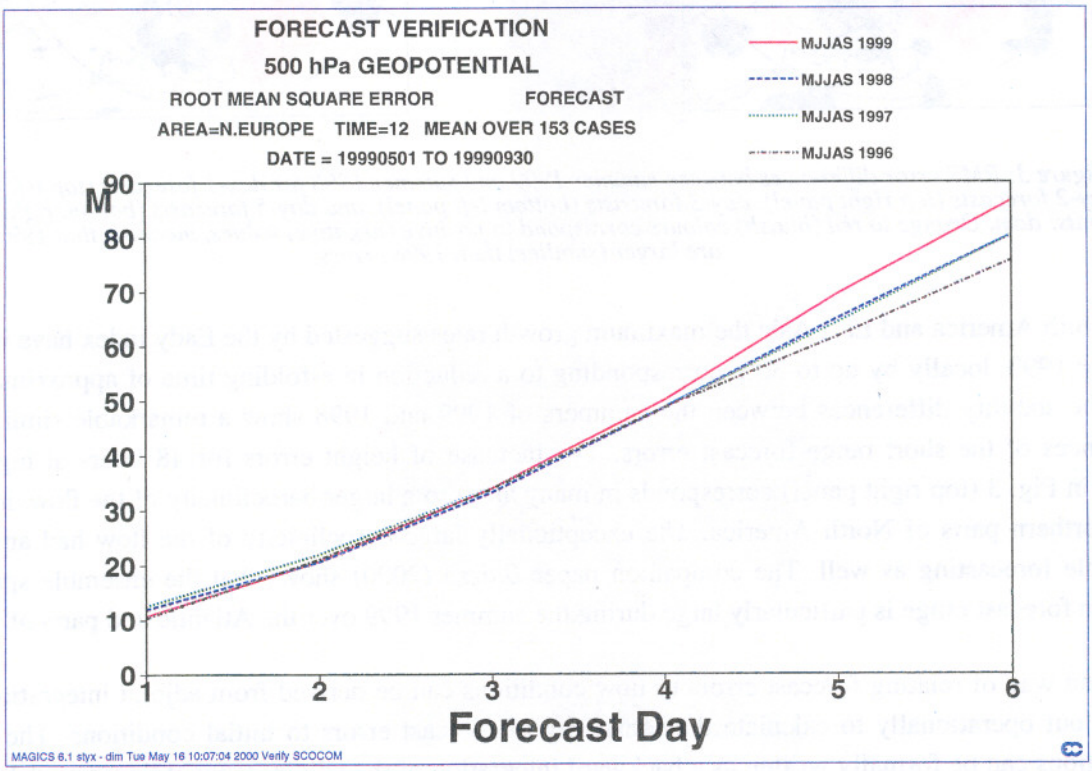

Figure 2: Root mean square forecast errors for 4 summer seasons over Northern Europe. Parameter: 500 hPa geopotential height. Units: m

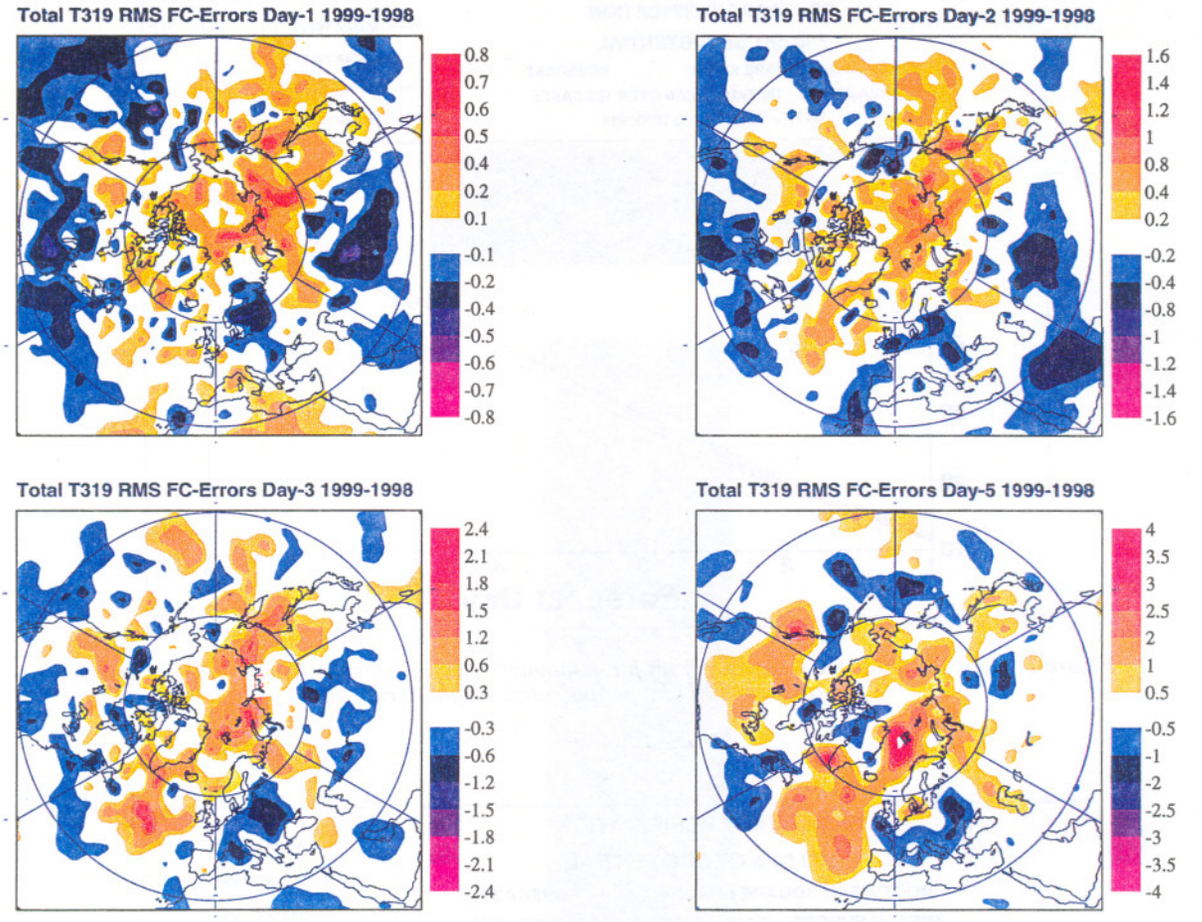

Figure 3: RMS error differences between summer 1999 and summer 1998 for day-1 forecasts (top left panel), day-2 forecasts (top right panel), day-3 forecasts (bottom left panel), and day-5 forecasts (bottom right panel). Units: dam. Orange to red (bluish) colours correspond to positive (negative) values, meaning that 1999 errors are larger (smaller) than 1998 errors.

over North America and East Asia the maximum growth rates suggested by the Eady index have increased in summer 1999, locally by up to 30% corresponding to a reduction in e-folding time of approximately half a day. The stability differences between the summers of 1999 and 1998 show a remarkable similarity to the differences of the short range forecast errors. The increase of height errors for 48 hours at high latitudes shown in Fig. 3 (top right panel) corresponds in many areas to a larger baroclinicity of the flow, in particular over northern parts of North America. The exceptionally large baroclinicity of the flow had an impact on ensemble forecasting as well. The companion paper *Buizza* (2000) shows that the ensemble spread in the medium forecast range is particularly large during the summer 1999 over the Atlantic and parts of Europe.

A second way of relating forecast errors to flow conditions can be derived from adjoint integrations that are carried out operationally to calculate the sensitivity of forecast errors to initial conditions. The diagnostic calculations can be formally written as a backward integration of the adjoint tangent linear model with day-2 errors as starting conditions. The result represents a projection of the forecast errors to the initial conditions and identifies the fast growing part of analysis errors (Rabier et al, 1996). These diagnostic integrations are based on simplified physics in the tangent linear and adjoint models. A refined estimate of analysis errors is

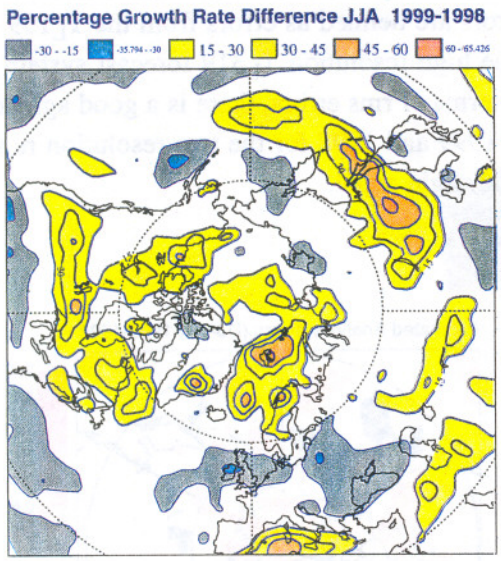

Figure 4: Percentage increase/decrease of the summer 1999 low level Eady-index compared to summer 1998. Positive values (yellow to red colours) mean increase in maximum growth rate.

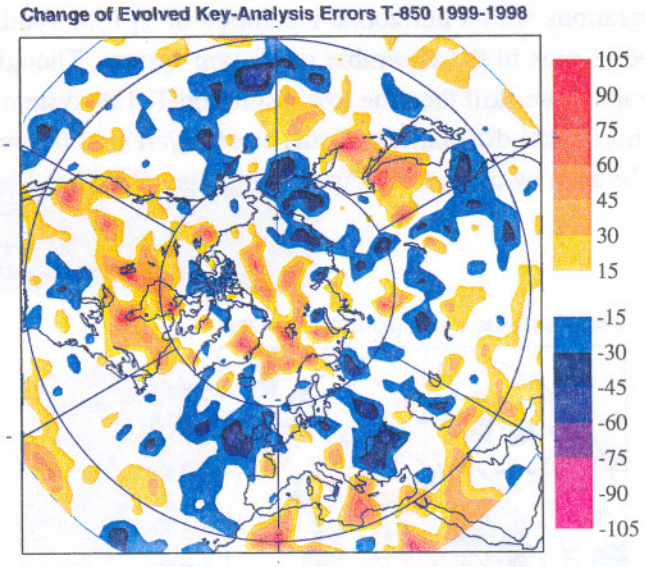

Figure 5: Percentage increase/decrease of the summer 1999 key analysis errors (evolved for 12 hours) compared to summer 1998. Parameter: temperature at 850 hPa. Positive values (orange to red colours) mean increase of key analysis errors in 1999.

obtained by using these adjoint integrations to minimize further the day-2 forecast errors (Klinker et all, 1998). The perturbations that minimize the day-2 forecast errors are called key analysis errors.

As these adjoint estimates of analysis errors project predominantly on to the unstable subspace it is not surprising that they are linked to the stability of the flow. The key analysis errors are characterized by small scale structures that have a vertical tilt similar to fast growing baroclinic waves. Key analysis errors evolved for a relatively short period of time (12 to 24 hours) provide a good estimate of the size of fast growing analysis errors for later forecast steps. Fig. 5 shows the differences between evolved key analysis errors for 1999 and 1998 in terms of percentage increase relative to summer 1998. The evolution time is 12 hours and the parameter is temperature at 850 hPa. Comparing these differences of key analysis errors with the percentage difference in the Eady index (Fig. 4) it is obvious that the baroclinicity of the flow has a strong influence on fast growing analysis errors. In most areas the increase of short range evolved key analysis errors can be explained by a change in the stability of the flow. The same is true for areas where the stability has been larger and the forecast errors have been smaller in 1999 compared to 1998. In particular larger key analysis errors over the Norwegian Sea correspond to larger estimates of maximum baroclinic growth rates.

4. Separation of "analysis errors" and "model errors"

Given that one can isolate fast growing analysis errors by using sensitivity integrations, it seems possible to separate to a certain extent "analysis errors" and "model errors". When "analysis errors" are compensated by correcting the initial conditions proportional to the diagnosed key analysis errors, the remaining forecast errors will be dominated by "model errors" if they are larger than slow-growing "analysis errors".

As the adjoint sensitivity integrations have been run with a reduced resolution, the error separation is based on integrations with a horizontal resolution of T_L159. Total errors are defined as errors from the T_L159 control forecast runs in the ensemble prediction system. Though the high resolution T_L319 forecast system has on average more skill then the low resolution T_L159 system in terms of rms errors, there is a good agreement in the horizontal distribution of rms error *differences* between 1999 and 1998 for the two resolution runs - this can be seen by comparing Fig. 6 bottom panel to Fig. 3 bottom right panel.

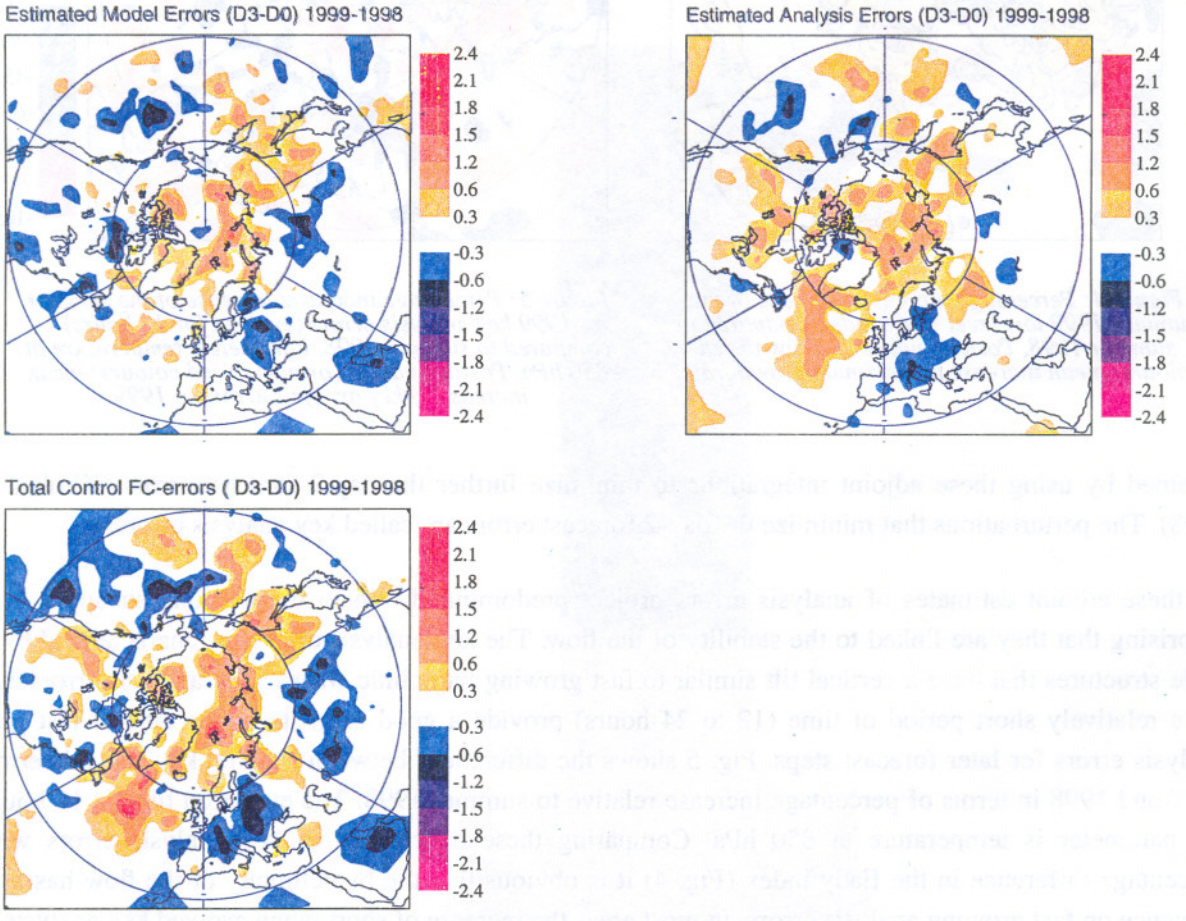

Figure 6: RMS error differences between summer 1999 and summer 1998 for the day-3 forecast. Total error (bottom left panel), evolved key analysis errors (top right panel), estimated model errors (top left panel). Units: dm. Orange to red (bluish) colours correspond to positive (negative) values, meaning that 1999 errors are larger (smaller) than 1998 errors.

The total rms errors for the T_L159 control forecast run are now crudely separated into "analysis errors" (control forecasts minus sensitivity integrations) and "model errors" (sensitivity integration minus analysis). At this point it has to be stressed that the separation is by no means unique. As the sensitivity calculations are only identifying a part of analysis errors - though the fast growing analysis errors form a vital part - the remaining errors after compensating for the key analysis errors are likely to represent an overestimate of the "model error" part. Additionally, "model errors" are by no means only slowly growing errors as their evolution is affected by the stability of the flow as well.

Fig. 6 shows rms error difference maps for day-3 forecasts. The comparison is done between total errors (Control forecast minus analysis, bottom left panel), evolved key analysis errors (Control forecast minus sensitivity integration, top right panel) and "model errors" (sensitivity integration minus analysis, top left panel). The much larger increase of key analysis errors compared to the increase in estimated model errors over northern parts of North America suggests that the error increase between 1999 and 1998 in this region is mostly due to fast growing analysis errors in the initial conditions. Over the Atlantic, North Pacific and polar areas both effects add to larger errors in 1999, but mostly with a larger contribution from the evolved key analysis errors. Over Central Europe the error *reduction* is coming from similar contributions from both parts.

The results obtained from the sensitivity integrations suggest that a large part of the error increases in 1999 compared to 1998 can be explained by fast growing analysis errors. These large errors can in principle arise from a poorer analysis or from different flow conditions. The close agreement of larger instability of the flow as diagnosed from the Eady-index and the increase of fast growing analysis errors suggests that the flow configuration in 1999 played a vital role for poorer forecast performance.

5. SVD analysis of flow dependent forecast errors

The qualitative differences between the flow in 1998 and 1999 plays an important part in understanding the differences in forecast accuracy. To study the relationship between forecast errors and the flow more qualitatively we use a singular value decomposition analysis of the two multivariate fields.

Ferranti et all (2000) diagnosed the flow dependent errors of wintertime forecasts by applying the Singular Value Decomposition (SVD) analysis to pairs of forecast errors and analyses. The SVD procedure isolates the linear combination of data from forecast errors and the linear combination of analysis anomalies that has a maximum covariance. The output of the SVD routine produces two sets of time coefficients that have a maximum linear correlation, one for anomalies of forecast errors and one for flow anomalies. Applying those time coefficients to the original fields of analyses and errors produces a pair of fields: the analysis anomaly and the corresponding forecast error anomaly. In their diagnosis of forecast errors for seven winters Ferranti et all (2000) found that flow dependent errors are approximately twice as large as seasonal mean errors.

5.1 Flow dependent error differences

For the investigation of summer forecast performance the question arises whether the increase of forecast errors in 1999 is associated with a particular flow anomaly. It seems that the Singular Value Decomposition (SVD) analysis represents an appropriate tool for to answer this question.

It is worth mentioning how the SVD analysis is applied here. Instead of working with the full fields as input, only a limited number of EOFs from a principal component analysis is used. The EOF filtering of the data pairs decreases the number of degrees of freedom in each field and produces results that are more stable with respect to sampling variability. In the first step of the SVD analysis where the principal components are calculated, the analysis and forecast error fields are used from three selected areas: the Northern Hemisphere, the Atlantic-European area and Europe. Forecast errors are presented to the SVD procedure in two different forms. By using forecast errors where the sign has been kept unchanged we obtain the signed flow dependent forecast error. By using the magnitude of forecast errors (so called L1 norm), the SVD technique identifies

flow dependent error magnitudes that can to some extent be compared to standard L2 norm statistics presented before.

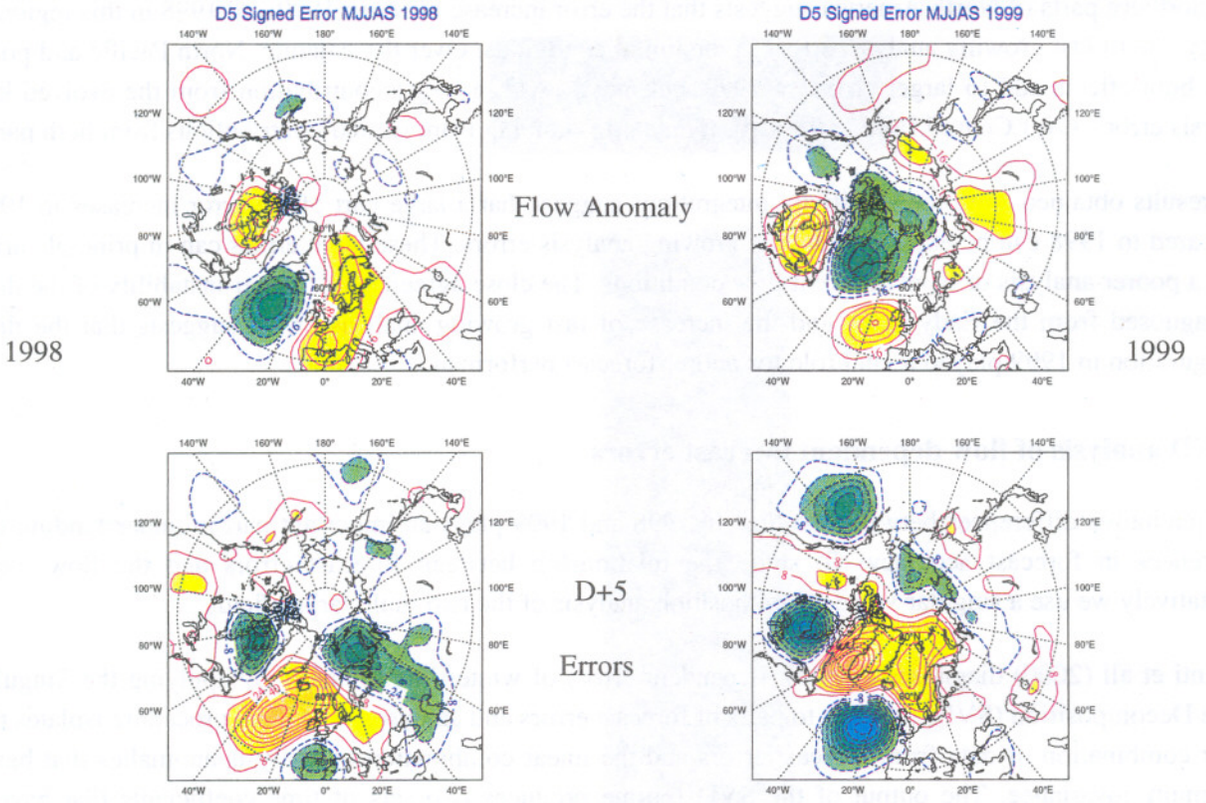

Figure 7: The dominant pair of SVD modes of 500 hPa height forecast error anomalies and of the 500 hPa height anomalies of the verifying analysis. Verifying analysis anomaly: upper panels, forecast error anomalies: lower panels, summer 1998: left panel, summer 1999: right panel. Positive (negative) values are shown in yellow-reddish (greenish) colours mean higher (lower) than normal values. Units: m. Contour intervals: 8 m for errors and 16 m for analysis anomalies.

Consistent results of flow dependent systematic errors for different forecast ranges are obtained for the extended summers of 1999 and 1998 when the error and analysis fields are based on the principal components over the Atlantic-European sector (20N to 80 and 90W to 60E). Applying the SVD analysis to the two summer data sets separately, in both years the largest positive covariances between the time series of analysis anomalies and error anomalies are found for a strong westerly flow anomaly in the central Atlantic (upper panels of Fig. 7) and in both summers the day-5 forecast errors (lower panels of Fig. 7) indicate a weakening of this flow pattern. The westerly flow anomaly is more pronounced in 1999 than in 1998 and the forecast errors show a spatially more consistent picture of weakening of the westerly flow over most parts of the North Atlantic. As the SVD analysis identifies pairs of anomalous patterns, the results support also the interpretation for a negative sign of the two anomaly fields which means that an easterly flow anomaly is weakened in the forecast as well.

The result obtained for performing the SVD analysis independently for the two extended summers may arise from two effects: The high level of errors for the 1999 summer may not have been reached in summer 1998 because the type of extended westerly flow anomaly has a smaller frequency in 1998, or for similar flow conditions the errors may be larger in 1999.

Common errors for similar flow conditions can be investigated more closely by applying the SVD analysis to the combined time series of the two summers instead of running the diagnostics on 1999 and 1998 independently. It is not surprising that the output from the SVD diagnostic from the combined summer seasons of 1998 and 1999 identifies again a flow pattern of increased westerlies in the analysis (Fig. 8) that has been found in the individual seasons as well. The output of the SVD analysis consists of two time series of coefficients, one for the flow anomaly and one for the forecast error. Additional information on occurrences of flow types can be gained by investigating the property of the time coefficient for the flow anomaly. After separating the time coefficients valid for the combined 1998-1999 summer into the 1998 and 1999 parts, the frequency distribution identifies the strength and the frequency of the flow anomalies in the two summers. In Fig. 9, which shows the histogram of the flow anomaly time coefficient, the positive x-axis values indicate a westerly flow anomaly in the Atlantic-European area and negative values correspond to an easterly flow anomaly. The strong westerly flow anomaly, characterized by large positive values on the x-axis, has been more frequently observed during summer 1999 than during summer 1998, whereas a weak easterly anomaly has been more frequently observed in 1998 than in 1999.

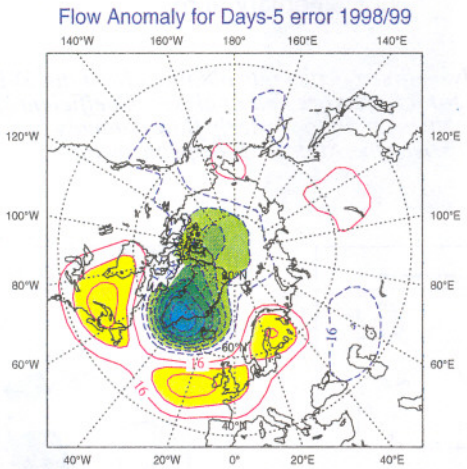

Figure 8: The dominant mode of 500 hPa height analyses obtained from the SVD analysis applied to the combined summer 1998 and 1999 day-5 forecast errors and analyses. Positive (negative) values are shown in yellow-reddish (greenish) colours mean higher (lower) than normal values. Units: m. Contour interval 16 m.

We can use these SVD time coefficients of the anomalous flow valid for the combined summers to calculate the error contribution for the separate summers. When we project the full error fields of height for the summer 1999 and 1998 on this flow anomaly by using the leading SVD time coefficients we obtain the forecast error components of the two summers separately for this flow anomaly (Fig. 10). Similar to the result obtained from

a separate SVD analysis for the two seasons we obtain larger flow dependent errors during the summer 1999 compared to 1998. Larger errors are found in particular over the North Atlantic and over the Norwegian Sea.

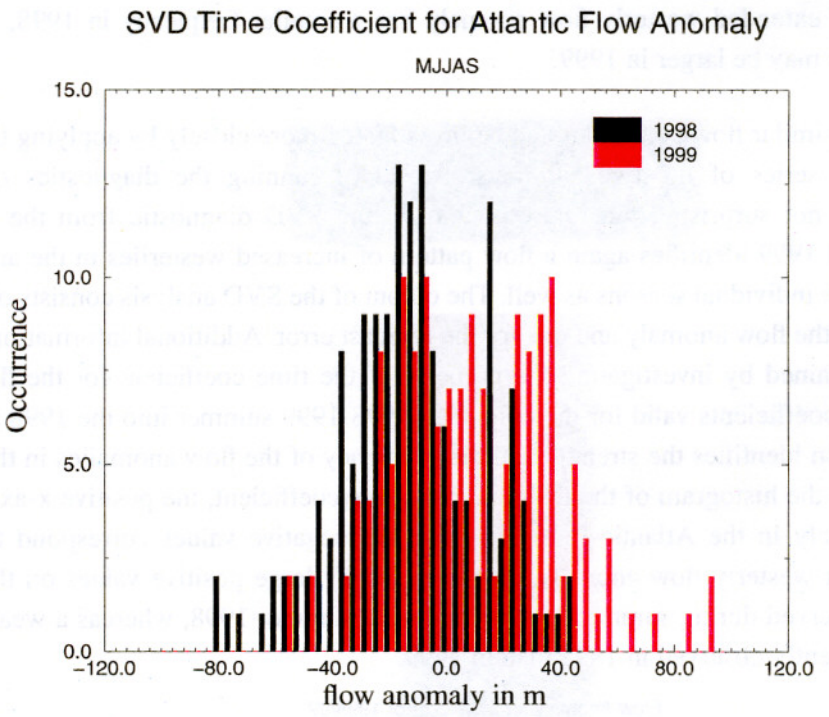

Figure 9: Histogram of the SVD analysis time coefficients obtained from the SVD analysis applied to the combined summer 1998 and 1999 day-5 forecast errors and analyses. Coefficients for summer 1999 in red and summer 1998 in black. Positive x-axis values apply to westerly flow anomaly and negative values to easterly anomaly over the Atlantic-European sector.

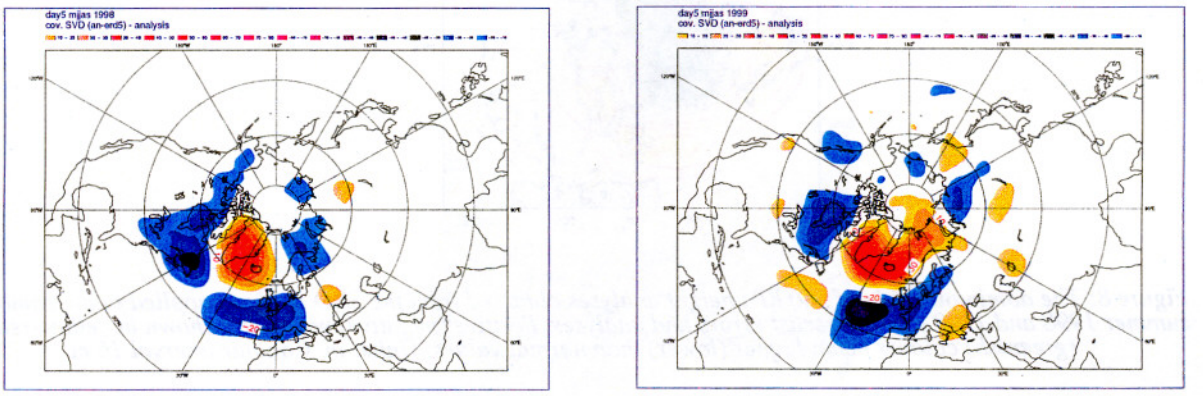

Figure 10: Projection of day-5 forecast error fields of individual summers on to the leading analysis SVD modes of combined summers. Time coefficients were obtained from the SVD analysis applied to the combined summer 1998 and 1999 day-5 forecast errors and analyses. Summer 1998: left panel, summer 1999: right panel. Red colours: positive errors, blue colours: negative errors. Units: m. Contour intervals 10 m.

5.2 Flow dependent error magnitudes

By using the absolute value of forecast errors instead of the signed value as input to the SVD analysis we obtain the flow dependent component of the magnitude of forecast errors. The comparison is based on diagnosing the two winter seasons separately. For the Atlantic the leading component of the flow pattern showing the highest correlation with the anomalies of absolute day-5 errors are shown in the top panels of Fig. 11 (1998 left panel, 1999 right panels). The selected flow patterns are in agreement with those found for the signed forecast error. The larger north-south gradient in 1999 suggests a stronger westerly flow anomaly during that summer.

Comparing the flow dependent errors of the two summer seasons (lower panels of Fig. 11) larger than normal error magnitudes are found in both years for flow configurations that represent an enhanced westerly flow in the Atlantic. The more pronounced flow anomaly in 1999 corresponds to larger error magnitudes. Whereas in summer 1998 slightly larger than normal errors are found over the North Atlantic and South-East Europe, in 1999 significantly larger errors are found mainly over the North Atlantic and Northern Europe.

The evolution of the different error patterns in 1998 and 1999 has been traced back to the forecast-range of day-3 and day-1. Fig. 12 shows the anomalous error magnitudes for both time ranges and seasons. The selected flow patterns of the analysis are very similar to those selected for the day-5 and are therefore omitted. For the forecast range of day-1 error magnitudes during the summer of 1998 larger than normal for this flow anomaly are confined to a small area to the West of Greenland, whereas in 1999 positive anomalous errors were covering a larger area over Greenland and further to the West.The sequence of anomalous error magnitude pattern for day-1, day-3 (Fig. 12) and day-5 (Fig. 11) show a general tendency of errors around Greenland to propagate downstream but also to spread to larger areas. The spreading is particular noticeable from day-1 to day-3 in summer 1999.

The SVD analysis for squared errors, the so called L2 norm produces *horizontal structures* of error differences that are almost identical to those obtained here for the L1 norm (error magnitudes). We can therefore estimate the contribution of flow dependent errors on total errors by comparing the flow dependent error magnitudes (Fig. 12 and 11) with seasonal rms errors (Fig. 3). The results from the SVD analysis are consistent with the rms error statistics in showing that Northern Europe and the Norwegian Sea have been effected by larger errors in 1999 than 1998. The additional information we gain here is, that the dominant signal of the error increase in 1999 occurs during episodes of strong westerly flow. Furthermore the SVD analysis gives a much clearer picture of error propagation. Day-1 errors around Greenland are larger in 1999 and evolve quickly into larger errors over the Atlantic and Northern Europe by day-5.

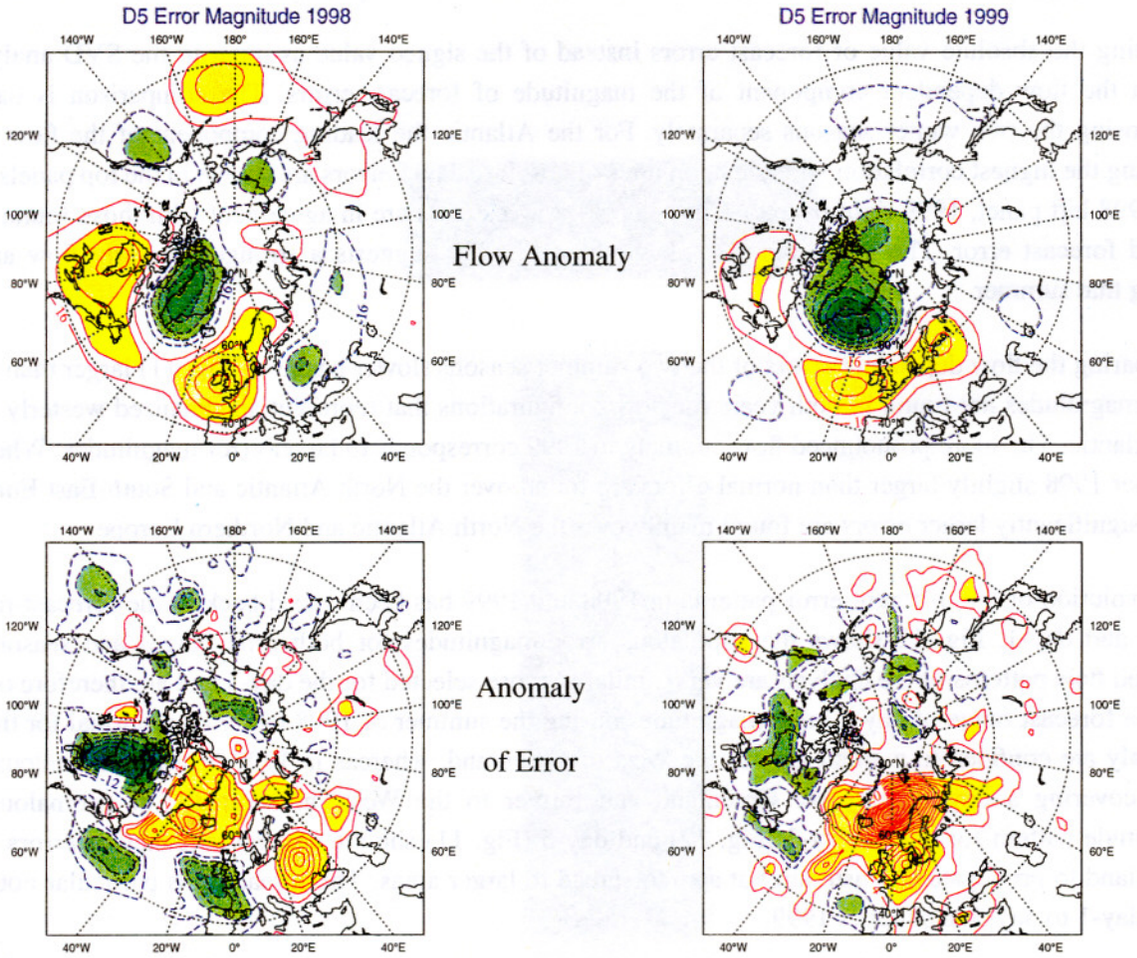

Figure 11: The dominant pair of SVD modes of 500 hPa height forecast error magnitude anomalies and of the 500 hPa height anomalies of the verifying analysis. Verifying analysis anomaly: upper panels, forecast error magnitude anomalies: lower panels, summer 1998: left panel, summer 1999: right panel. Positive (negative) values are shown in yellow-reddish (greenish) colours mean higher (lower) than normal values. Units: m. Contour intervals: 4 m for errors and 16 m for analysis anomalies.

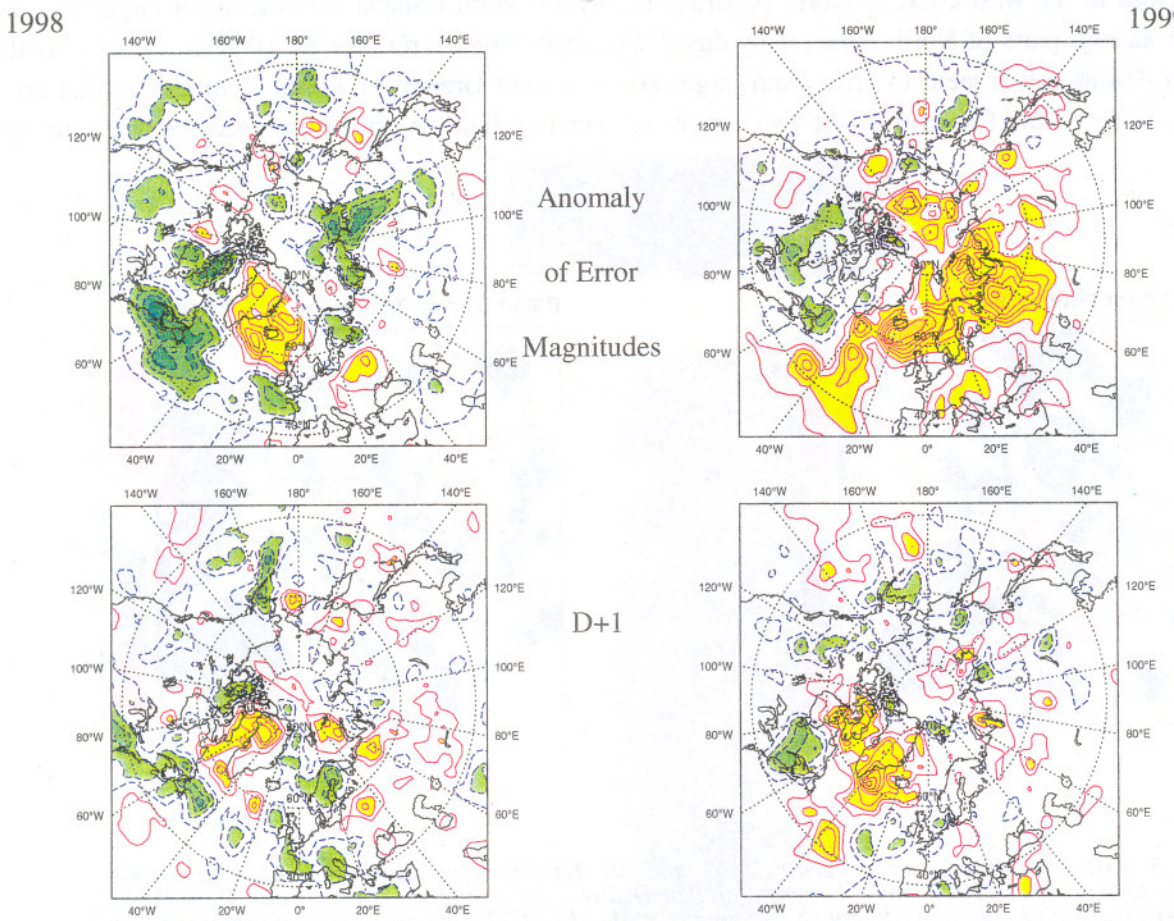

Figure 12: The same as Fig.11 but only for forecast error magnitude anomaly. Summer 1998: left panel, summer 1999: right panel. day-3 forecasts upper panels, day-1 forecasts lower panels. Positive (negative) values are shown in yellow-reddish (greenish) colours mean higher (lower) than normal values. Units: m. Contour intervals: 2 m for day-3 errors and 0.75 m for day-1 errors.

6. Difference between UK Met. Office and ECMWF forecasts

The differences between the forecasts from ECMWF and the UK Met. Office have been used to investigate further the nature of forecast errors during the extended summer period of 1999. The rms error differences between the day-5 forecasts from the two centres, each verified against their own analyses show mainly smaller errors for the ECMWF forecasts (Fig. 13, left panel). However for North America and parts of the Atlantic the ECMWF forecasts have larger errors than forecasts from the UK Met. Office for this extended summer period. The locations with larger errors from the ECMWF forecasts occur mostly where an increase of errors between 1999 and 1998 has been found for ECMWF forecasts (see Fig. 3 bottom right panel). For the shorter range of 2 days (Fig. 13, right panel) ECMWF forecasts have larger errors over widespread parts of high latitudes. For this range we find an even higher similarity between the pattern of error differences (ECWMF - UK Met.Office) and the pattern of error increase from 1998 to 1999 (Fig. 3 top right panel).

Following the error differences through several forecast ranges it seems that the larger errors of ECMWF forecasts close to the west coast of North America and over Central Canada are crucial for larger forecast errors over eastern parts of North America by day-5. Similarly larger errors for ECMWF forecasts over the Norwegian Sea at day-5 seem to arise from larger errors around Greenland at day-2. In general the error differences between the forecasts of the two centres are smaller than the error differences between the two years

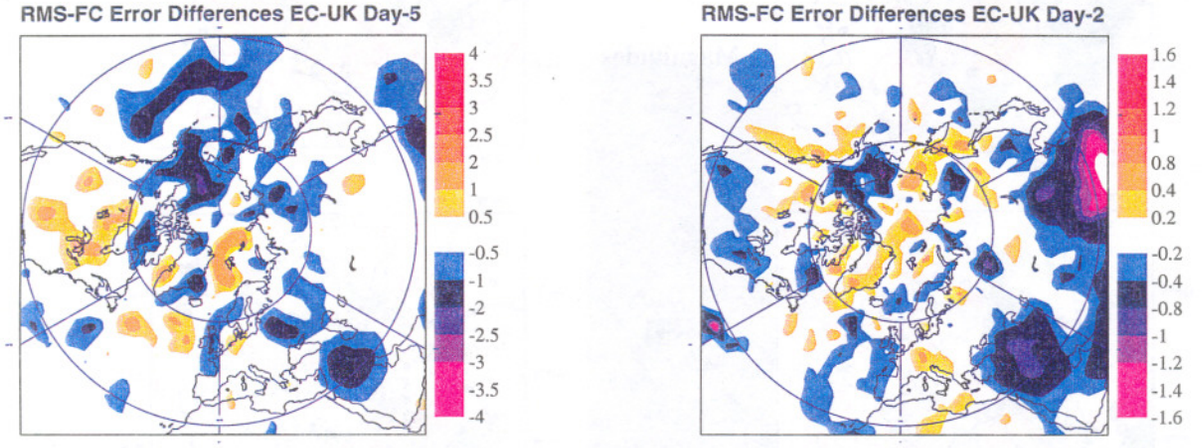

Figure 13: RMS error differences between ECMWF and UK Met. Office forecasts for an extended summer 1999 (May to September 1999). Day-5 forecast (left panel), day-2 forecast (right panel). Units: dam. Positive (negative) values in orange-reddish (bluish) colours mean that ECMWF forecast errors are larger (smaller) than UK Met. Office forecast errors.

As both, the error increase from 1998 to 1999 and the error differences between the two centres are so similar, it seems likely, that both quantities have a similar flow dependence. The extension of the diagnosis to cover the flow dependent part of error differences therefore helps to focus on dominant error differences. To identify the flow dependent part of error differences between forecasts from ECMWF and the UK Met. Office we again make use of the SVD technique. First the forecasts errors are calculated separately for each centre. Then the difference between the magnitude of the forecast errors (ECMWF minus UK) and the analysis fields provides the input pair for the SVD analysis to investigate the flow dependent differences in the performance of the two systems. The focus is still on the Atlantic-European sector for which the EOF's of the two input components to the SVD have been calculated.

The leading pair of SVD components can now be interpreted in the way that for the identified height anomaly positive or negative error differences indicate that the ECMWF forecasts have larger or smaller error magnitudes respectively for this flow type. The SVD analysis applied in this way to day-5 forecast error differences and analyses finds again a westerly (or easterly) flow anomaly as the dominant height anomaly pattern that has the largest correlation to forecast errors differences (Fig. 14, right panels). For the medium forecast range the ECMWF errors are larger to the west of Greenland, over the North Atlantic, over the Norwegian Sea and over parts of Western Europe. Smaller errors are found over eastern parts of North

America. Those regions, for which larger day-5 errors for ECMWF have been diagnosed for this westerly flow, fall into regions where we found an increase in rms-errors for ECMWF forecasts between summer 1999 and 1998.

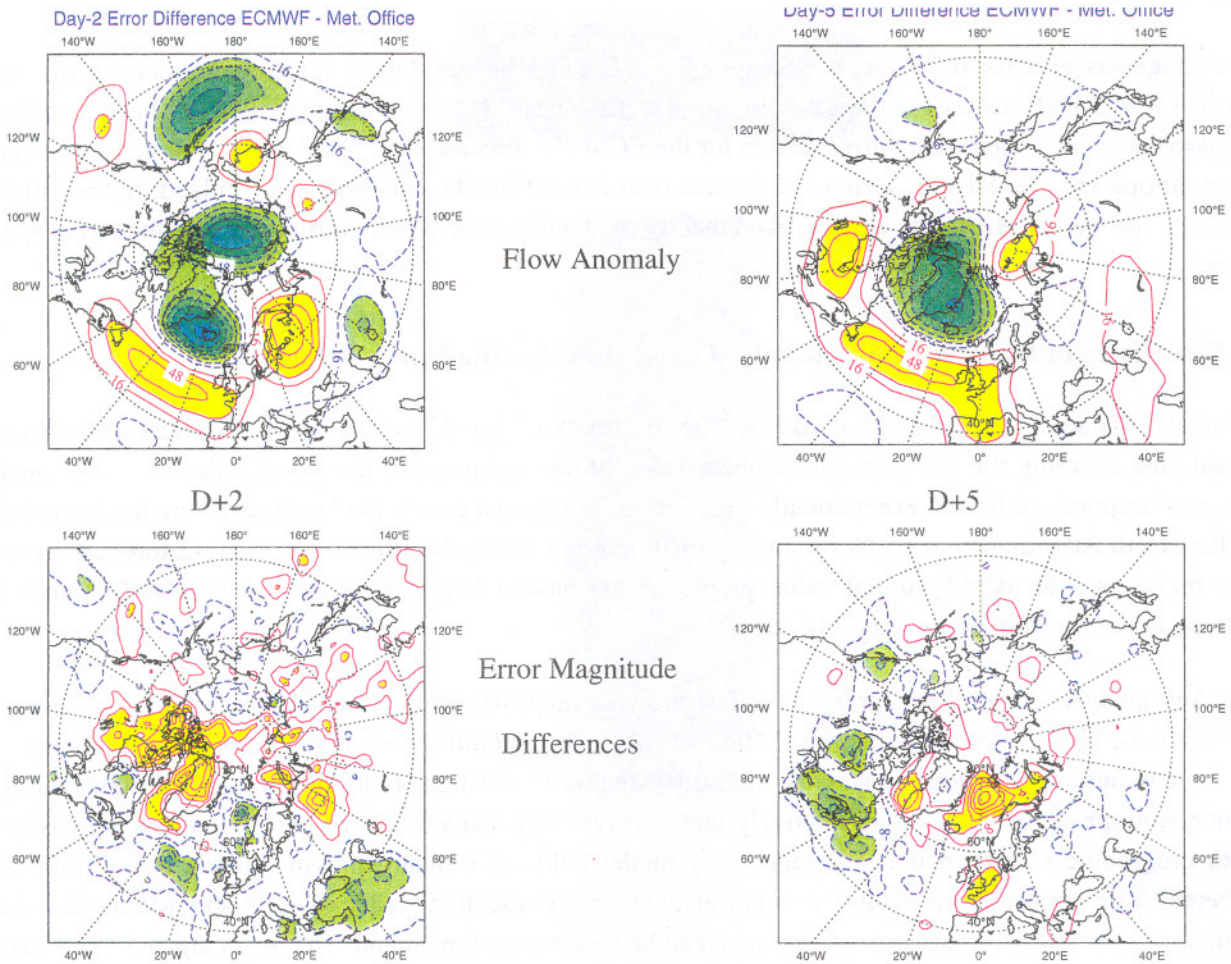

Figure 14: The dominant pair of SVD modes of 500 hPa height forecast error magnitude anomalies and of the 500 hPa height anomalies of the verifying analysis. The forecast errors represent magnitude differences between the ECMWF and UK Met. Office forecast error magnitudes. Verifying analysis anomaly: upper panels, forecast error magnitude difference anomalies: lower panels, day-2 forecasts: left panel, day-5 forecasts: right panel. Positive (negative) values are shown in yellow-reddish (greenish) colours mean larger (smaller) errors for ECMWF forecasts compared to UK Met. Office forecasts. Units: m. Contour intervals: 2 m for day-2 errors, 8 m for day-5 errors and 16 m for analysis anomalies.

Going back to day-2 forecasts (Fig. 14, left panels) a similar flow anomaly has been selected as for the day-5 error differences. ECMWF errors at this shorter range are larger for eastern parts of North Canada and areas around Greenland. The diagnosis of the sequence of similar SVD analyses for day-1, day-2, day-3 and day-5 forecast errors of the two centres indicates, that the larger errors for the ECMWF forecasts are propagating in 5 days from the region of North Canada and Greenland into the North Atlantic and most northern parts of Europe.

The output of the SVD analysis supports the interpretation that for an easterly flow anomaly as well. Whereas ECMWF forecasts have larger errors than UK Met. Office forecasts when the westerlies are stronger than normal, ECMWF forecasts have smaller errors when the westerlies are weaker than normal. However, the result is not exactly symmetric, as a flow configuration with a strong easterly flow anomaly is less frequent than a strong westerly flow anomaly.

The diagnosis of error differences between ECMWF and UK Met. Office forecasts has shown that larger errors for ECMWF forecast in the medium range seem to occur for a flow anomaly that has been identified to be associated with lager than normal errors for the ECMWF forecast itself. The origin of large forecast errors over Europe and the Atlantic seem to be located over North Canada and close to Greenland. In the following sections special investigations are described that try to identify more closely errors that originate from specific areas.

7. Sensitivity of forecast errors to initial conditions for the North Atlantic

Normally the adjoint calculations to identify the sensitivity of day-2 forecast errors to initial conditions use a target area covering the Northern Hemisphere from 30 degrees north to the North Pole. To isolate analysis errors of importance for the exceptionally high European forecast errors, the target area here has been limited to the North Atlantic (25 to 65 North and 70 to 10 degrees West). Minimization of day-2 forecast errors for this region should identify initial value problems that have a large impact on forecast performance over Europe in the medium range (day-3 to day-7).

The key analysis errors that represent the fast growing analysis errors (Control forecasts minus Sensitivity integrations, Fig. 15) for temperature at 850 hPa evolved for 24 hours show large values close to the east coast of Canada indicating a large sensitivity of North Atlantic forecast errors to the initial conditions over those continental areas. To what extent relatively large evolved key analysis errors over mountains are associated with orographic effects in the interpolation of model fields to standard pressure levels is not quite clear. Whereas little impact on the evolution of key analysis errors is seen in the lee of the Rocky Mountains and the Himalayas, signs of downstream development can be seen east of Greenland. The major signal of key analysis errors propagates from the east coast of North America to the West-Atlantic by day-3 and to the Central and East Atlantic by day-5. A secondary track of error propagation is seen east of Greenland. By day-7 the evolved key analysis errors have further propagated to the Norwegian Sea and to Northern Europe, having merged with errors originating from the Greenland area. The apparent propagation of evolved key analysis errors refines the picture of error propagation we have seen so far, that increased errors over Northern Europe are strongly linked to analysis errors over North Canada. Error sources close to Greenland seem to have an enhancing effect on errors arising from areas further to the west.

8. Investigation of special events using transplant sensitivity

As mentioned before quite often forecast errors arise from errors in the initial conditions at different locations. It is difficult to locate the exact area which is most important for the forecast quality from the full Northern Hemisphere sensitivity. Hollingsworth et al (1985) have shown that transplanting a different analysis locally into a control analysis can explain the impact of localized analysis differences on forecast performance. This technique is used here to estimate the impact of analysis errors in certain areas. Experiments have been based

on applying sensitivity corrections to the initial conditions in selected areas only, like the polar area, North America and the North Pacific. This procedure is different to the targeting of the sensitivity calculations for a specific area like the Atlantic as shown in the previous section when for each target area the adjoint calculations have to be repeated.

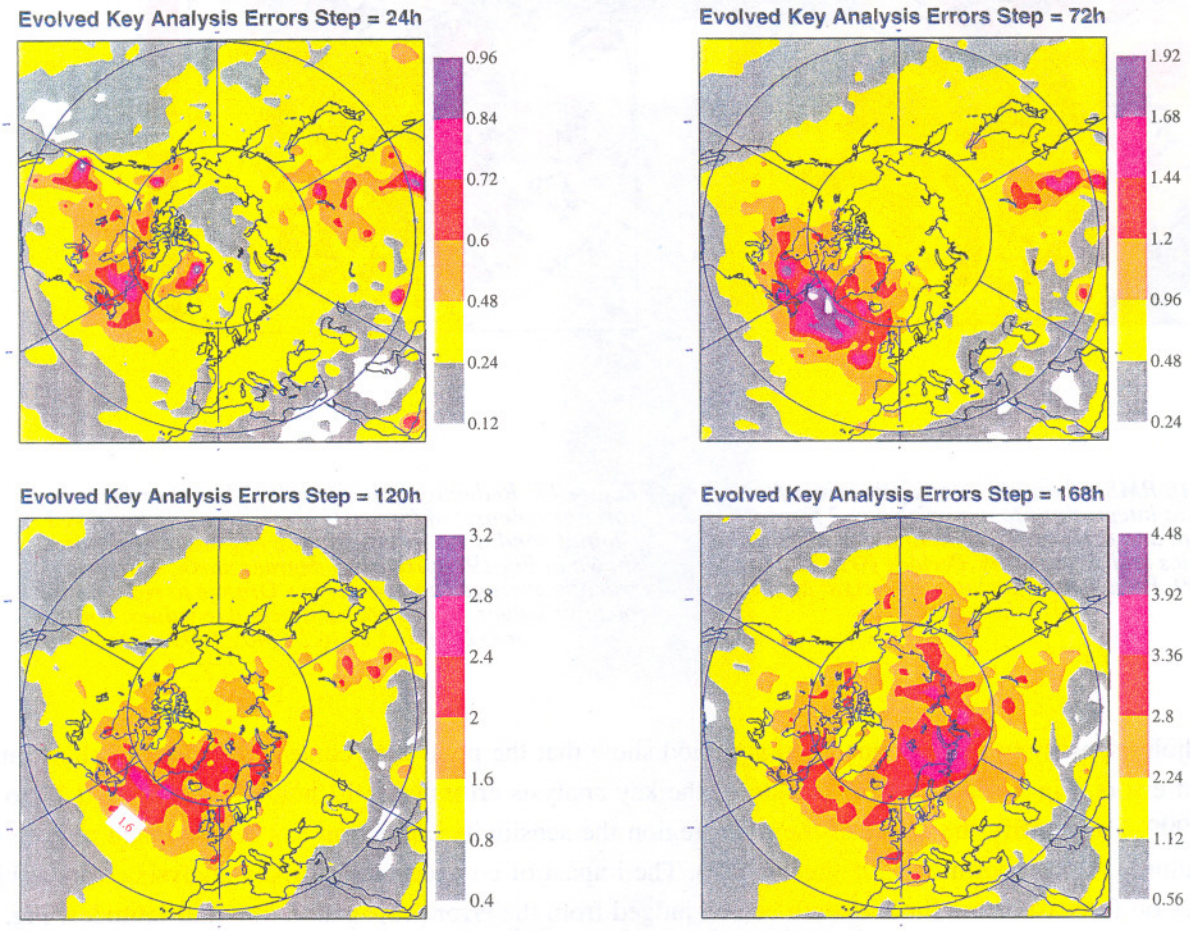

Figure 15: Key analysis errors from adjoint sensitivity integrations targeted to the North Atlantic and evolved for 24, 72, 120 and 168 hours. Parameter: Temperature at 850 hPa. Units: K.

8.1 Polar problems in May

The first period investigated was 10-20 May 1999 (10 to 20). During this period large negative differences between UK Met. Office and ECMWF low level temperatures were observed. The fact that low level temperatures in the ECMWF analysis and short range forecast were too high could also be seen from a large bias of the first guess temperature against an observation site on the polar ice.

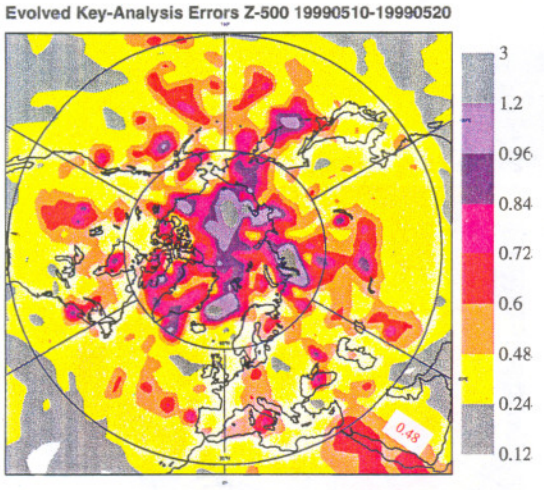

Fig. 16 RMS values of key analysis errors from adjoint integrations to minimize day-2 forecast errors and evolved for 12 hours. Use of linear physics and 9 iterations. Period: 10 to 20 May 1999. Parameter: geopotential height at 500 hPa.Units: dam.

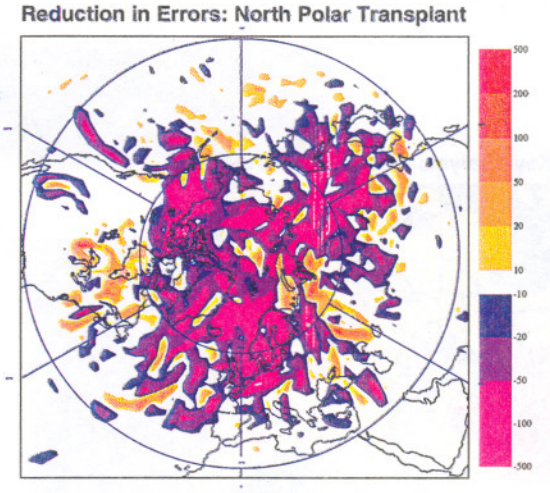

Figure 17: Reduction (blue-purple colours) or increase (orange colours) of forecasts errors by using improved initial conditions (compensating key-analysis errors shown in Fig 19) north of 65 degrees north. Parameter: vertically averaged kinetic energy. Orange to red colours: positive values, bluish colours: negative values. Units: m2s2. Period 10 to 20 May 1999.

The adjoint sensitivity calculations for that period show that the polar area contributed much to short range forecast errors. Fig 16 shows the evolution of the key analysis errors after 12 hours for the period 10 to 20 May 1999. To quantify the effect of the polar region the sensitivity correction has been applied only to the high latitude region (north of 65 degrees north). The impact of compensating for key analysis errors at high latitudes on forecast errors further south can be judged from the error propagation into mid-latitudes. Fig. 17 shows the reduction of day-3 forecast errors in terms of kinetic energy integrated over 850, 500 and 300 hPa (blue colours: polar transplant forecasts better than operational forecasts). The reduced polar errors yield a marked improvement in the Atlantic-European area and over the North-West Pacific as well. This transplant experiment has shown that analysis errors in the polar area can have a detrimental impact on forecast skill over Europe. However, the result is likely to be dependent on the flow conditions. For the May period investigated here, a strong baroclinic flow extended from Greenland over North Atlantic into Europe.

8.2 July baroclinic flow conditions

The beginning of July 1999 was quite unusual in that a strong east-west baroclinic zone extended over the Central US as can be seen from the Eady Index for that period (Fig. 18). The evolved key-analysis errors for the first 10 days in July show enhanced sensitivity over the Atlantic and North Pacific as well as over polar areas (Fig. 19). For this period transplant sensitivity experiments covering a number of sensitive areas have been run. A similar experiment as for the May period to compensate analysis errors in the polar region showed

only a fairly small amount of error propagation to mid-latitudes, as in early July the major frontal zone was located far to the south, away from the polar region studied for the May period.

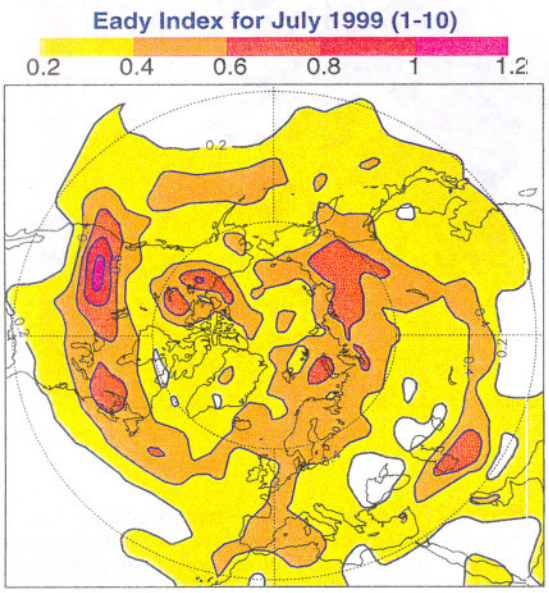

Figure 18: Low level Eady-index for the first 10 days of July 1999 as an estimate of maximum growth rate of baroclinic waves. Units: day^{-1}.

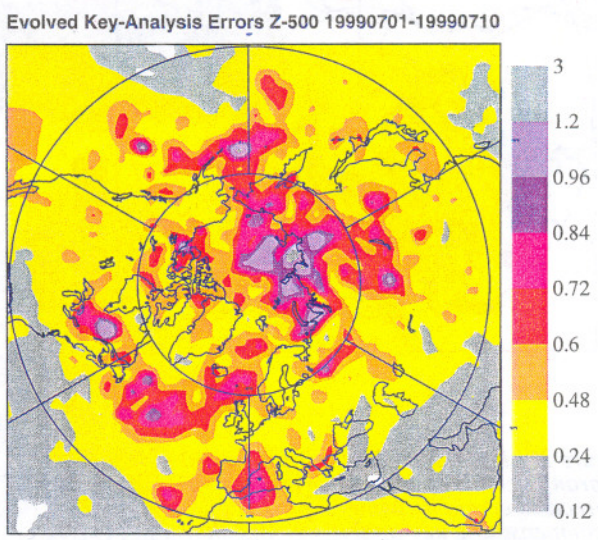

Figure 19:Key analysis errors from adjoint integrations to minimize day-2 forecast errors and evolved for 12 hours. Use of linear physics and 9 iterations. Period: 1 to 10 July 1999. Parameter: geopotential height at 500 hPa. Units: dam

a) North America transplant

Correcting the analysis only over North America (60-30N) was intended to demonstrate the effect of the exceptional baroclinic zone on the error evolution further downstream. Fig. 20 shows to what extent the day-3 errors are reduced by compensating the key analysis errors over North America. Within 72 hours errors originating from North America have propagated into Western and Northern Europe.

b) *North Pacific transplant*

A further area of large sensitivity was found over the North Pacific. Transplant experiments based on corrected initial conditions in the area 145E/-135W/60N/25N have been carried out for the first 10 days in July again. Analysis errors evolving from this region propagate quickly into the into North-America and further into the Atlantic by day-3 (Fig. 21). It seems that the baroclinic frontal zone over North America acts as an efficient channel for error propagation from the Pacific to the Atlantic.

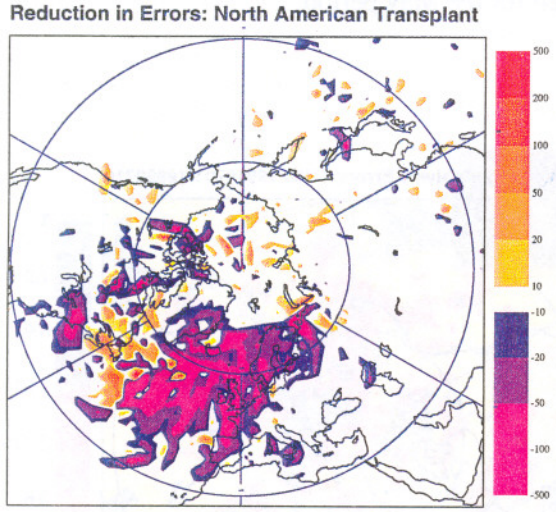

Figure 20: Reduction (blue-purple colours) or increase (orange colours) of day-North America transplant3 forecasts errors by using improved initial conditions (compensating key-analysis errors shown in Fig 19) over North America. Parameter: vertically averaged kinetic energy. Units: m2s2. Period: 1 to 10 July 1999

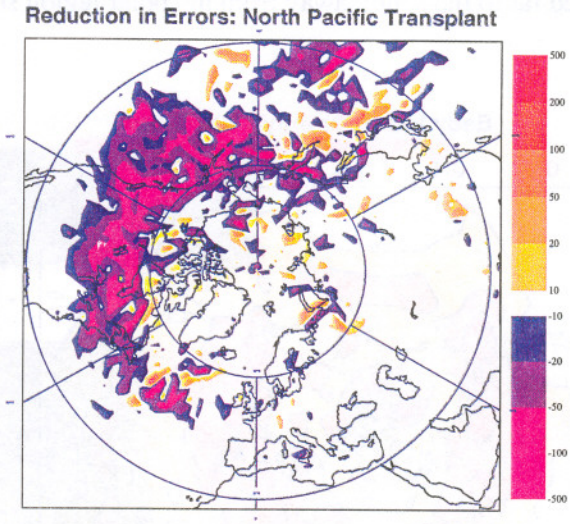

Figure 21: Reduction (blue-purple colours) or increase (orange colours) of day-3 forecasts errors by using improved initial conditions (compensating key-analysis errors shown in Fig 19) over the North Pacific. Parameter: vertically averaged kinetic energy. Units: m2s2. Period: 1 to 10 July 1999

9. Improvements from the new model cycle (21r4)

Two versions of the ECMWF forecasting system have been used operationally for the summer 1999 season, the so called cycles 21r1 and 21r2. Details of model and analysis system differences between the two cycles are listed in a companion paper by Simmons (2000). For almost the entire extended summer season a revised forecasting system (cycle 21r4) has been tested before it became operational on 12 October 1999. The improvements of forecast quality of cycle 21r4 in terms of anomaly correlation have been discussed in Simmons (2000). Here we use the differences between cy21r4 and the operational forecasts in our diagnostic study to characterize the specific problems of the operational system for the Summer 1999.

9.1 Comparison of cycle 21r4 forecasts with operational forecasts and UK Met. Office forecasts

Fig. 22 shows the geographical distribution of day-1 forecast error differences between the new cycle 21r4 and the operational system. A reduction of forecast errors can be found over most areas, in particular over the Arctic and some low latitude regions. Whereas day-1 forecasts for Northern Canada have been improved, a small degradation is found over central parts of North America. A fair comparison of cycle 21r4 summer errors for 1999 with operational errors of 1998 would require re-running at least a subset of the 1998 period with the cy21r4 formulation of the forecasting system. However, assuming for a moment that cycle 21r4 would already have been operational during the summer 1999 the scale of improvements in forecast performance becomes apparent. Comparing day-1 errors for cycle 21r4 forecasts for summer 1999 with operational forecast errors in 1998 (Fig.23) in most areas a decrease of errors is the dominant signal, only a small error increase occurs over some high latitude region. The remaining error increase at high latitudes is most likely related to a more baroclinic flow in summer 1999 as discussed in section 3.

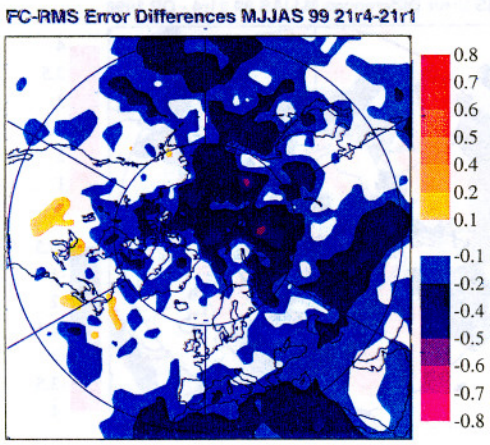

Figure 22: Day-1 rms error differences between cycle 21r4 and operational forecasts for May to September 1999. Units: dam. Orange to red (bluish) colours correspond to positive (negative) values, meaning that cy21r4 errors are larger (smaller) than operational errors.

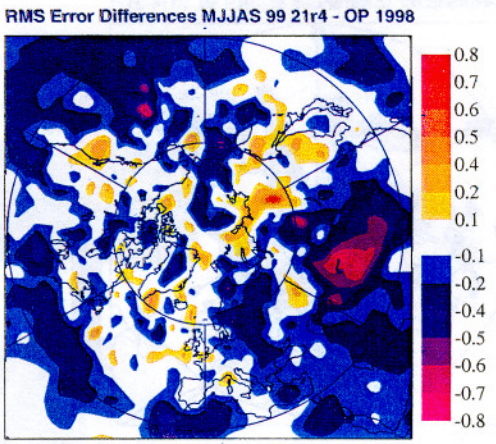

Figure 23: Day-1 rms error differences between cycle 21r4 forecasts of 1999 and operational forecasts of 1998. Period: May to September. Units: dam. Orange to red (bluish) colours correspond to positive (negative) values, meaning that cy21r4 errors are larger (smaller) than 1998 operational errors.

From the sequence of forecast errors maps for day-1, day-2, day-3 and day-5 (only the latter one shown in Fig. 24) it appears that an important part of the improved forecast performance of cycle 21r4 at day-5 over the North Atlantic, Northern Europe and the Norwegian Sea has arisen from an error reduction over North Canada. Comparing cycle 21r4 day-5 forecast errors with 1998 (Fig. 25), the reduction of errors has reversed the sign of forecast error differences over Northern Europe alleviating the forecast problems seen there during the summer of 1999.

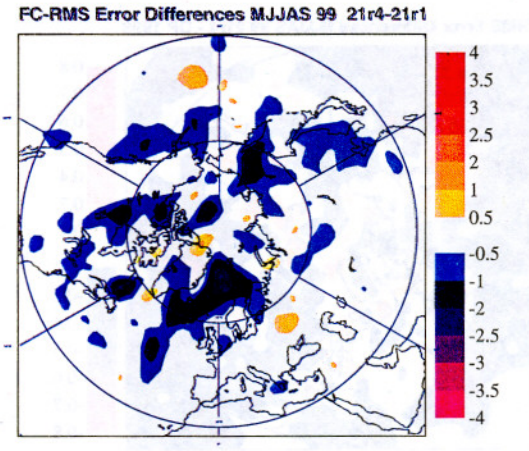

Figure 24: Day-5 rms error differences between cycle 21r4 and operational forecasts for May to September 1999. Units: dam. Orange to red (bluish) colours correspond to positive (negative) values, meaning that cy21r4 errors are larger (smaller) than operational errors.

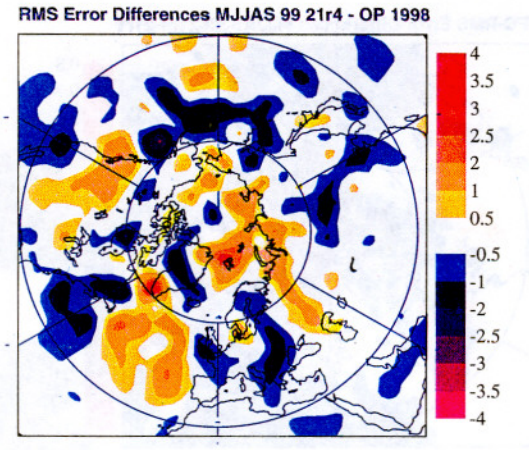

Figure 25: Day-5 rms error differences between cycle 21r4 forecasts of 1999 and operational forecasts of 1998. Period: May to September. Units: dam. Orange to red (bluish) colours correspond to positive (negative) values, meaning that cy21r4 errors are larger (smaller) than 1998 operational errors.

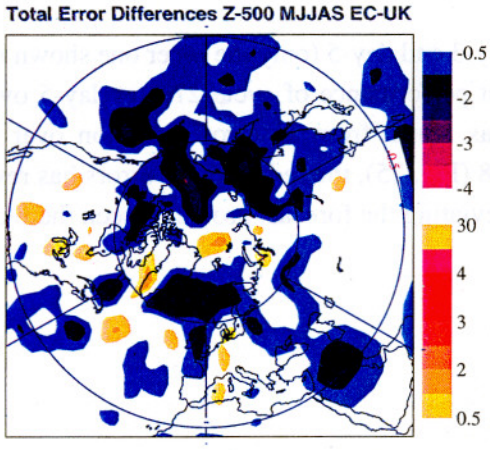

Figure 26: Day-5 rms error differences between cycle 21r4 forecasts and UK Met. Office forecasts. Period: May to September 1999. Units: dam. Orange to red (bluish) colours correspond to positive (negative) values, meaning that cy21r4 errors are larger (smaller) than UK Met. Office operational errors.

The improvement of forecast skill with cycle 21r4 is also apparent by comparing the cycle 21r4 summer forecast errors with forecast errors from the UK Met. Office forecast system operational in 1999. This comparison for day-5 (Fig. 26) shows a marked improvement in particular over Northern Europe (compared to Fig. 13, left panel) where for the cy21r4 the day-5 forecast errors are up to 15 meters lower compared to the UK Met. Office forecast errors.

9.2 Flow dependent errors for cycle 21r4

Applying the SVD technique to the cycle 21r4 analysis and forecast sets shows that the largest reduction of flow dependent errors has been achieved for the highly baroclinic flow configuration over the Atlantic. Fig. 27 shows the leading pair of flow anomaly and day-5 anomalous error magnitudes for cycle 21r4 which should

be compared to the SVD output for the operational forecasts for summer 1999 (Fig. 11, right panels). Though still larger than normal errors are found over the North Atlantic and North-East Europe for a westerly flow anomaly over the North-Atlantic, the errors for cycle 21r4 are much smaller than for the operational forecasts.

The diagnostics presented earlier suggest that unusually large forecast errors may have arisen from analysis problems over the polar area and North Canada. It also seems from Fig. 22 that the new forecasting system of cycle 21r4 improved largely in this area. As we have shown in section 3 that the forecast performance is closely related to the stability of the flow it is important to note that the data assimilation from the new model cycle produces an analysis that has a vertical structure of the atmosphere that is more stable than the analyzed flow in the operational system (Fig. 28, left panel). This is particularly the case over the polar region with the consequence of reduced potential of error evolution from that region. The change in the stability of the flow has the tendency to spread to lower latitudes in the forecast range up to day-2 (Fig. 28, right panel). In the average day-5 forecasts the stability decrease is even larger. A further sign that the polar analysis in the new system 21r4 has been improved is noted in Simmons (2000) by the fact that the cycle 21r4 analysis is closer to the UK Met. Office analysis than the operational analysis.

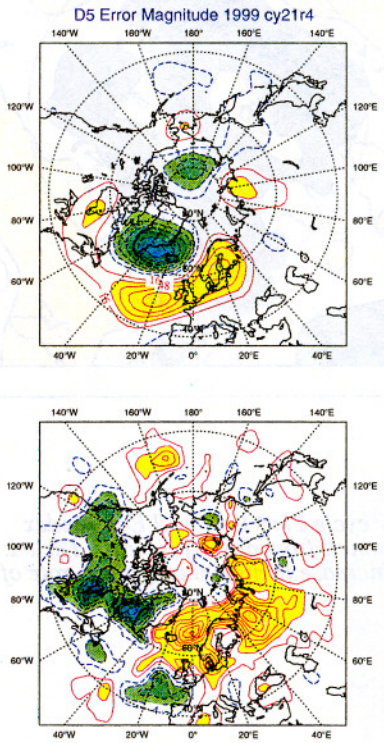

Figure 27: The dominant pair of SVD modes of 500 hPa height day-5 forecast error magnitude and 500 hPa height anomalies of the verifying analysis for cycle 21r4 (summer 1999 to be compared with the right panels of Fig. 11). Verifying analysis anomaly: upper panel, forecast error magnitude anomaly: lower panel. Positive (negative) values are shown in yellow-reddish (greenish) colours mean higher (lower) than normal values. Units: m. Contour intervals: 4 m for day-5 errors and 16 m for analysis anomalies.

9.3 Comparison of analysis performance between cycle 21r4 and the operational system

As some of the diagnostics presented here point to a possible problem in the analysis over Canada the performance of the data assimilation system in that area is of particular interest. One way of evaluating the data assimilation is to calculate statistics of the background (9 hour forecast) and analysis departures from radiosonde observations. This has been done for both systems, for the operational data assimilation and for the new system cycle 21r4, for a period of almost 4 months (7 May to 30 September 1999). Fig. 29 shows the standard deviations for the u and v components of the wind and for temperature for a region from 50 to 90

degrees north and 120 to 70 degrees west. The statistics are shown for the operational data assimilation system (blue lines) and for the experimental system cy21r4 (black lines). Though the fit of the background to observations is only slightly better, a marked reduction in the standard deviation of the analysis-observation departure is found for the two wind parameters, and a modest reduction in the standard deviation of temperature. It is interesting that the background departure of temperature has a bias that implies of a slightly larger static stability at lower levels in the background forecast of cycle 21r4 compared to operational background fields. Similarly the change in the background bias of the zonal wind reduces the vertical wind shear. The increase of static stability and the reduction of vertical wind shear contribute both to a reduction of baroclinicity of the flow in the background confirming the change in the Eady index shown in Fig. 28.

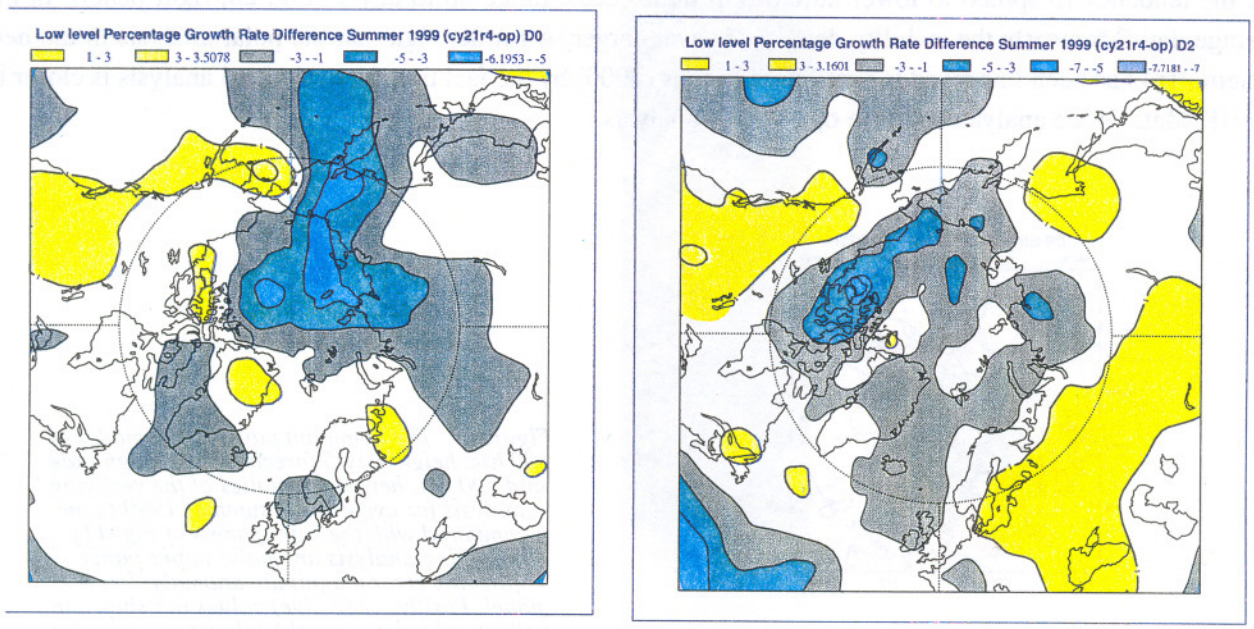

Figure 28: Percentage increase/decrease of the summer cycle 21r4 low level Eady-index compared to operations, left panel: from analyses, right panel: from day-2 forecasts. Period: Summer 1999. Positive values (yellowish colours) mean increase in maximum growth rate of baroclinic waves.

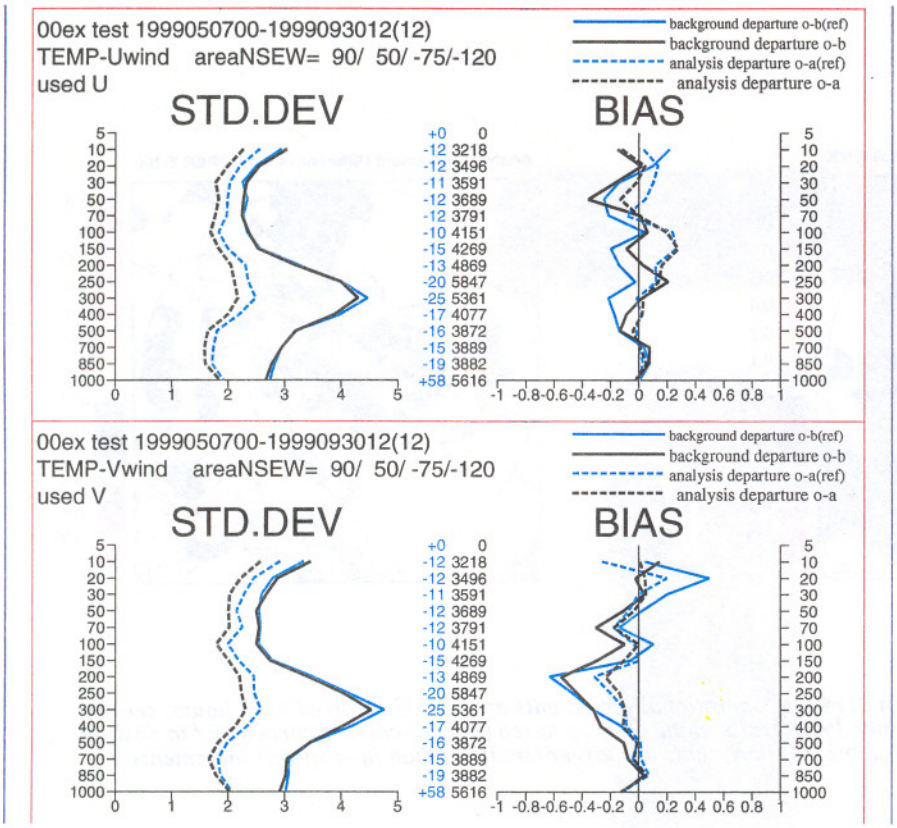

Figure 29a: Standard deviation and bias of the zonal and meridional wind component with respect to radiosonde measurements for an area 50N to 90N and 120W to 75W averaged over a period from 7 May to 30 September 1999. The solid lines are fits of backgrounds to radiosonde observations, the dashed lines are fits of analyses to these observations. Blue lines for the operational system, black lines for the for the new cycle 21r4 system.

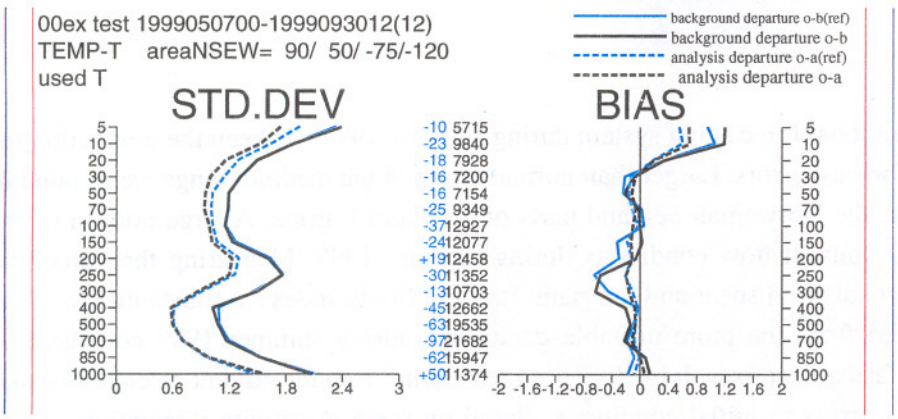

Figure 29b: The same as Fig.29a except for temperature.

A further quality check on the analysis can be made by diagnosing the analysis increments. Operationally the analysis fields and the verifying background fields are archived. As the analysis represents a 3 hour forecast and the background a 9 hour forecast, the background-analysis differences effectively describe 3 hourly evolved increments.

The dominant signal of the evolved increments in the new assimilation (cycle 21r4) is that they are generally smaller than those in the operational system as shown for 12 UTC 500 hPa geopotential height increments in Fig. 30, left panel. The reduction is similar to the reduction of 24 hour forecast errors, with largest values over the polar regions, North and North-West Canada. A secondary maximum of increment reduction is found over Greenland. The change of orography in cycle 21r4 was most significant over Greenland and the different

diagnostics reported earlier have shown that one of the areas to contribute to downstream forecast errors was that around Greenland.

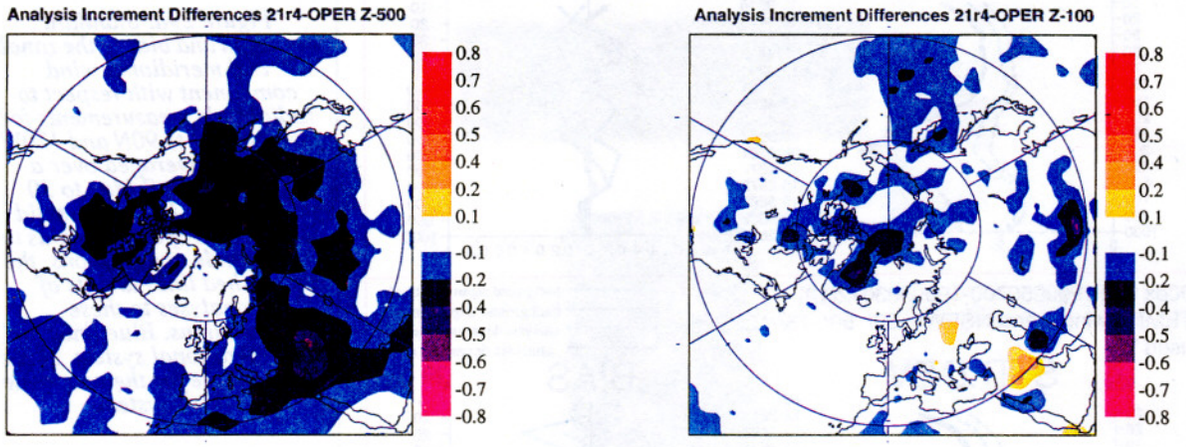

Figure 30: rms differences between 21r4 and operational increments at 12 UTC (evolved for 3 hours, see text for explanation) for May to September 1999. Units: dam. Orange to red (bluish) colours correspond to positive (negative) values, meaning that cy21r4 increments are larger (smaller) than operational increments.

10. Summary and discussion

The poor performance of the operational forecasting system during summer 1999 has been the motivation for an extensive diagnostic study of forecast errors. Larger than normal errors in the medium range were found in particular over the North Atlantic, the Norwegian Sea and parts of Northern Europe. A large portion of the error increase could be linked to unusual flow conditions during summer 1999. Measuring the baroclinic instability of the flow with the vertical wind shear and the static stability (Eady Index), a substantial increase of error growth could be expected from the more unstable conditions during summer 1999 compared to summer 1998. A confirmation of faster error growth in 1999 has been found from the adjoint calculations that identify the sensitivity of forecast errors to initial conditions. Based on these sensitivity integrations it has been shown that a large part of the error increase in summer 1999 has been due to fast growing analysis errors.

By applying the Singular Value Decomposition (SVD) technique to pairs of forecast error data and analyses the flow dependent part of increased forecast errors in 1999 could be identified. Larger errors during the summer of 1999 occur mostly during periods of strong westerly flow over the Atlantic, of which there were more episodes during summer 1999 than during summer 1998. Flow dependent error magnitudes show a clear picture of errors propagating from North Canada and the area around Greenland in the short range forecasts into the North Atlantic and Northern Europe by day-5.

North Canada and Greenland are also areas from where flow dependent error differences between ECMWF forecasts and UK Met. Office forecasts evolve into larger errors over Northern Europe for the ECMWF system in the medium range. Further diagnostics based on regional sensitivity again emphasize the importance of North and North-West Canada for the error evolution over Northern Europe in the medium range.

Some diagnostics have been repeated on parallel runs for the new cycle (21r4) of the forecasting system that became operational in October 1999. A noticeable improvement was found in particular for situations of a strong baroclinic westerly flow over the Atlantic. As a consequence the North Atlantic and Northern Europe are those areas where medium range forecast errors have been improved most with the new cycle.

From several diagnostics we have been able to identify the region where large forecast errors originated. The analysis of flow-dependent errors, error differences between forecasts from ECMWF and the UK Met. Office, and error differences between forecast from the new cycle 21r4 and the operational system all identify the area of North Canada and Greenland as the prime region for initial errors that contribute to Northern European forecast errors in the medium range.

The reduction of medium range forecast errors obtained during the test period of summer 1999 with the new cycle 21r4 provide particularly valuable evidence on the nature of forecast errors. Improvements of the fit of analysis and background fields to observations and a reduction of analysis increments and short range forecast errors seem to have had an important contribution to reduced medium range forecast errors over Northern Europe. Error growth for cycle 21r4 was also positively effected by the fact that the new forecasting system tended to produce an analysis that had a vertical structure with a reduced potential for unstable baroclinic growth.

The diagnosis of the influence of flow changes on error propagation answers the question why winter time forecasts are less likely effected by errors in the source regions identified here. Analysis and model errors from high latitudes are more likely to effect medium range forecast errors in summer than in winter due to the location of the major zone of baroclinic instability which is located further to the north in summer than in winter.

References

Buizza, R.: Forecasting system performance in summer 1999. Part 3: EPS. ECMWF Technical Memorandum No.323

Ferranti, L., Klinker, E., Hollingsworth, A. and Hoskins, B.J., 2000: Diagnosis of systematic errors dependent on flow pattern. To be published

Hollingsworth, A, Lorenc, A.C., Tracton, M.S., Arpe, K., Cats, G., Uppala, A. and Kållberg, P., 1985: The response of numerical weather prediction systems to FGGE level IIb data. Patr I: Analyses. Q. J. R. Meteorol. Soc., **111,** 1-66.

Klinker, E., Rabier, F. and Gelaro, R.: Estimation of key analysis errors using the adjoint technique. *Q. J. R. Meteorol. Soc.*, **124**, 1909-1933