Observations, assimilation and the improvement of global weather prediction - Some results from operational forecasting and ERA-40

Adrian Simmons

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, UK

1. Introduction

ECMWF's Annual Seminar was last devoted to data assimilation in 1996. In the opening lecture at that Seminar, Roger Daley discussed how, over the preceding fifteen years, data assimilation had evolved from being a minor and often neglected sub-discipline of numerical weather prediction to become not only a key component of operational weather forecasting but also an approach that was important for environmental monitoring and estimation of the ocean state. The seven years since then have seen numerical weather prediction reap considerable benefit in terms of improved forecast accuracy from earlier and ongoing investments in the scientific and technical development of data assimilation. The application to environmental monitoring has advanced with the emergence of a broad user-base and new initiatives in reanalysis, and ocean-state estimation has become established operationally as a component of seasonal forecasting systems. Moreover, as can be seen generally in the Proceedings of this 2003 Seminar, considerable advances continue to be made in the techniques and range of application of data assimilation.

The principal aims of this contribution are to illustrate some basic aspects of atmospheric data assimilation and to illustrate some aspects of the improvement of forecasts, particularly for recent years. To do this, results are drawn from the recently completed ERA-40 reanalysis of the atmospheric observations made since mid 1957, and from ECMWF's operational analyses and forecasts. Some results from the operational forecasting systems of the Met Office and NCEP are also used. Attention is concentrated on analyses and deterministic forecasts for the extratropical troposphere, predominantly for the 500hPa height field. This is partly because of the continuing widespread use of such forecasts (in addition to newer products such as from ensemble prediction or related directly to weather elements), partly because it is for the 500hPa height field that the most comprehensive records are available, and partly to enable the simplest of messages to be presented. Significant improvements have also been achieved in analyses and forecasts for the tropics and stratosphere, although both regions pose additional challenges. For example, Hólm *et al.* (2002) discuss issues related to the hydrological cycle in the tropics, and stratospheric issues are discussed in the proceedings of the ECMWF workshop on the topic held earlier this year.

The basics of the atmospheric observing system and data assimilation are discussed in the following section. Section 3 presents some results from ERA-40, discussing the evolution of the observing system since 1957, the corresponding evolution of analysis increments and data fits, and the corresponding evolution of forecast accuracy, which is compared also with the change in accuracy of ECMWF's operational forecasts over the second half of the reanalysis period. Further aspects of the improvement in operational forecasts are presented in the Section 4. Section 5 draws some inferences relating to data assimilation derived from studies of the differences between analyses and forecasts from ECMWF and the Met Office, and from the differences between successive ECMWF forecasts. Some results on the scale-dependence of forecast error are presented in Section 6, which is followed by a brief concluding discussion.

2. Basic aspects

Fig. 1 presents maps of the operational coverage of atmospheric data received at ECMWF for the six-hour period beginning at 09UTC on 5 September 2003. Such maps are presented daily on the ECMWF website (http://www.ecmwf.int). Data of each of the types illustrated are assimilated operationally at ECMWF, although in the case of scatterometer and ozone data, observations from the ERS-2 satellite (denoted by the orbits confined to the vicinity of Europe and the North Atlantic in the bottom centre and right plots of Fig. 1) are at the time of writing being passed only passively through the assimilation system for monitoring purposes rather than being actively assimilated.

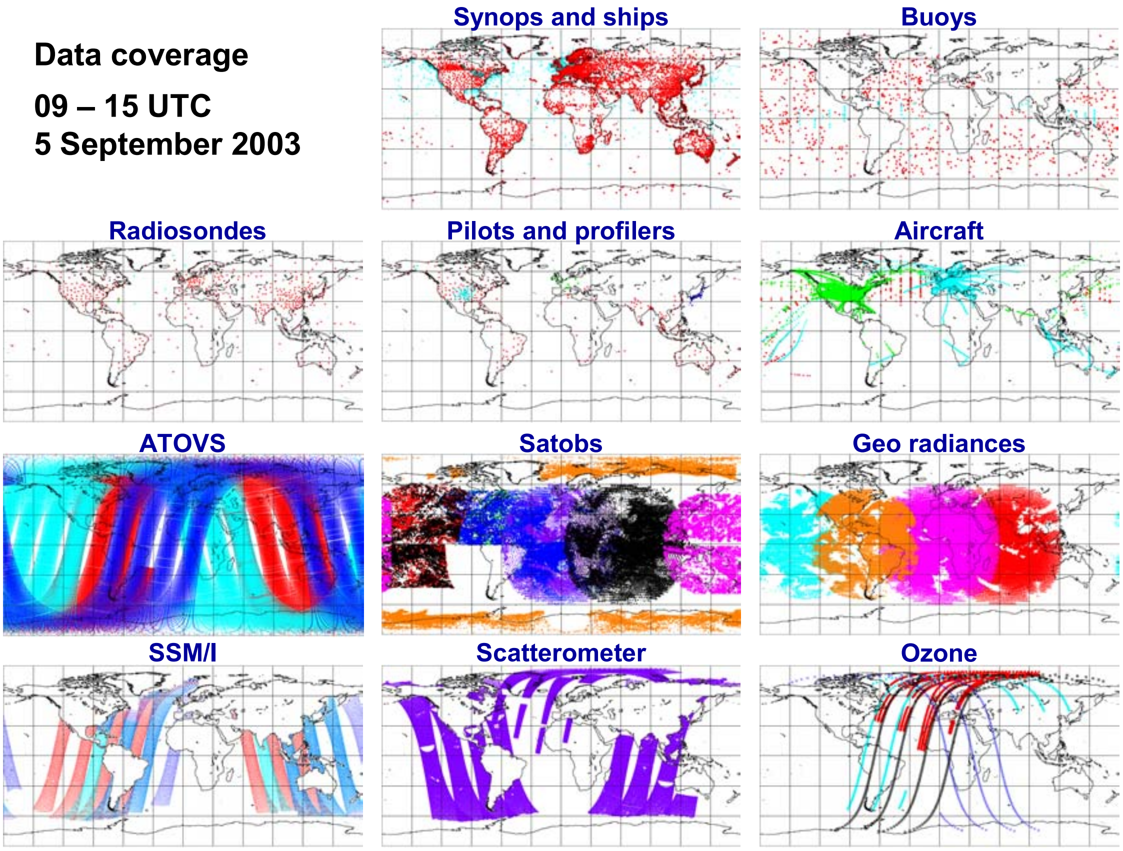

Figure 1: Operational data coverage plots for the six-hour period centred on 12UTC, 5 September 2003.

The atmospheric observing system illustrated in Fig.1 comprises a range of *in-situ* and remotely-sensed measurements. These different types of measurement have different accuracies and different temporal and spatial coverage. Some measurements are of quantities such as wind and temperature that are prognostic model variables, values of which are required to initiate forecasts. Others provide information that is only indirectly related to the model variables, the radiances in different parts of the electromagnetic spectrum measured from satellites in particular. Use of data from the satellite-based component of the observing system is discussed in Thépaut's contribution to these Proceedings.

Data assimilation is a process in which a model integration is adjusted intermittently to bring it in line with the latest available observations related to the model's state variables. It is shown schematically in Fig. 2 for a six-hourly assimilation with a three-dimensional (spatial) analysis of observations such as the 3D variational (3D-Var) system used by ECMWF previously for its main operations (Andersson *et al.*, 1998), recently for ERA-40 and currently for short-cut-off analyses from which forecasts are run to provide boundary conditions for short-range limited-area models. 3D-Var systems are also used, for example, by the Met Office in the UK (Lorenc *et al.*, 2000) and by NCEP in the USA (Parrish and Derber, 1992).

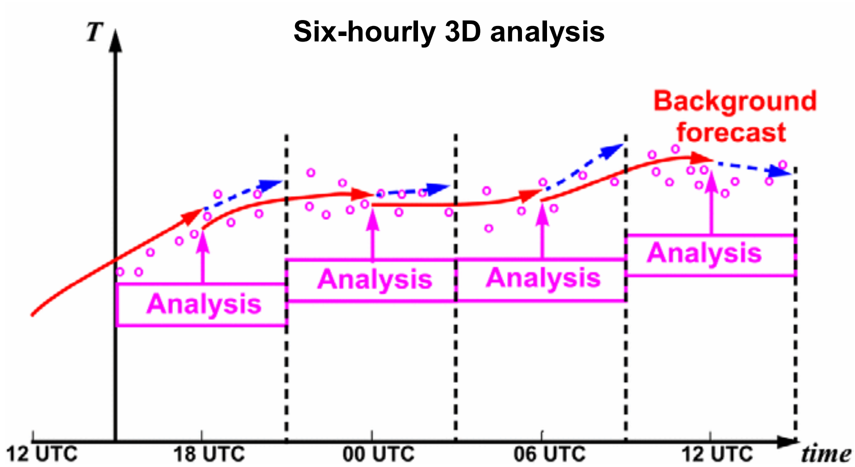

Figure 2: Schematic of the data assimilation process for a six-hourly three-dimensional data assimilation such as used for ERA-40. The ordinate denotes a model variable, temperature for example.

As illustrated in Fig. 2, the six-hour "background" forecast from the preceding analysis is adjusted by adding an "analysis increment". The increment is determined from an analysis of the "innovations", the differences between observations and equivalents derived from the background forecast. In the so-called "FGAT" approach (First Guess at the Appropriate Time) these differences are computed at the actual times of the observations, but are assumed to be valid at the main synoptic hour for which the analysis is carried out. The analysis for a particular time in the configuration shown makes direct use of observations taken within three hours of analysis time, and is influenced by earlier observations through the information carried forward by the background forecast. The weighting given to the innovations in the analysis procedure depends on the expected errors of the observations and background forecast.

The analysis increment is typically small compared with the change made by the background forecast from the preceding analysis. An example is presented in Fig.3, which shows maps of the mean-square changes in 500hPa height produced by the background forecasts from 06UTC to 12UTC (left) and by the 12UTC analyses (right) for the year 2001 from the 3D-Var ERA-40 data assimilation. The main Atlantic and Pacific storm tracks are evident in the northern hemisphere plot, as is a local maximum over the Mediterranean. Larger changes occur in the southern hemisphere storm track encircling Antarctica. Mean-square analysis increments are typically smaller than the background-forecast changes by an order of magnitude or more, as can be seen from the extratropical averages printed on each map. Little detail of the pattern of increments can be discerned when plotted as in Fig. 3 with a colour shading appropriate to showing the pattern of background-forecast changes; these increments will be discussed further in the following section, where they are shown with a more-discriminating shading in Fig. 7.

The quality of the background forecast depends both on the quality of the preceding analysis and on the quality of the forecast model. The predominant role of the background model in evolving the estimated state of the atmosphere is indicative of the need for as accurate as possible a model to enable as accurate as possible an analysis. A good model representation of synoptic- and planetary-scale dynamics and thermodynamics is needed to capture the main evolution of a field such as 500hPa height, but a good model representation of local weather elements is needed as well, not only for direct analysis and prediction of such elements but also to enable correct extraction of the information related to larger scales that is contained in the observations of these elements.

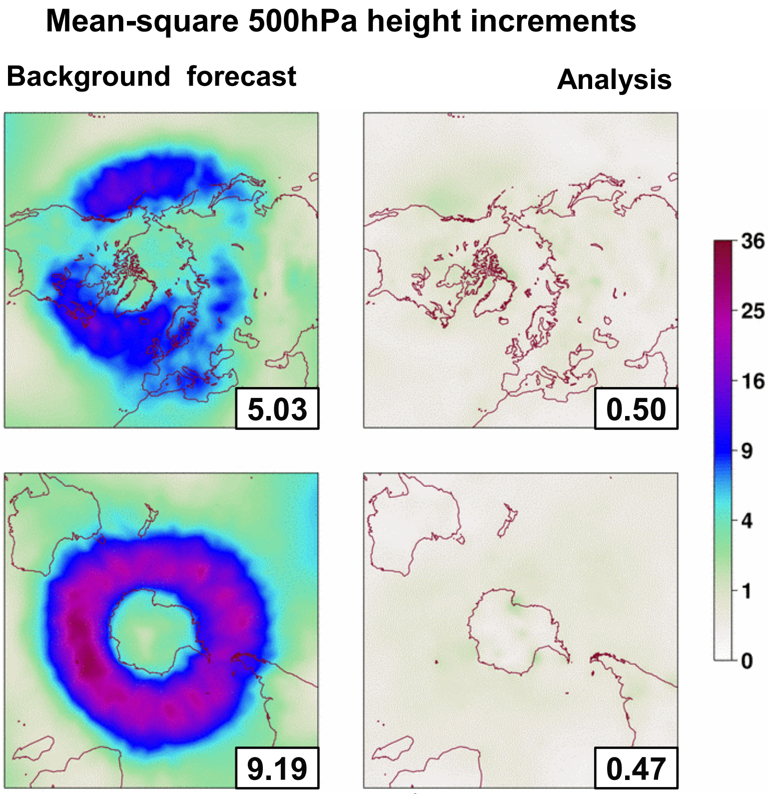

Figure 3: Mean square increments in 500hPa height (dam^2) for the ERA-40 background forecasts from 06 UTC to 12UTC (left) and for the 12UTC ERA-40 analyses (right) for the year 2001, shown for the northern (upper) and southern (lower) hemispheres. Values averaged over each extratropical hemisphere are shown in the bottom-right corner of each map.

ECMWF and Météo-France currently use four-dimensional variational (4D-Var) data assimilation operationally. The basic approach of 4D-Var is illustrated schematically in Fig. 4. An adjustment to the background forecast valid at the beginning of the assimilation window is determined iteratively so as to reduce the discrepancy between the observations within the assimilation window and the equivalent forecast values, with the adjustment limited to be consistent with the estimated errors of the observations and

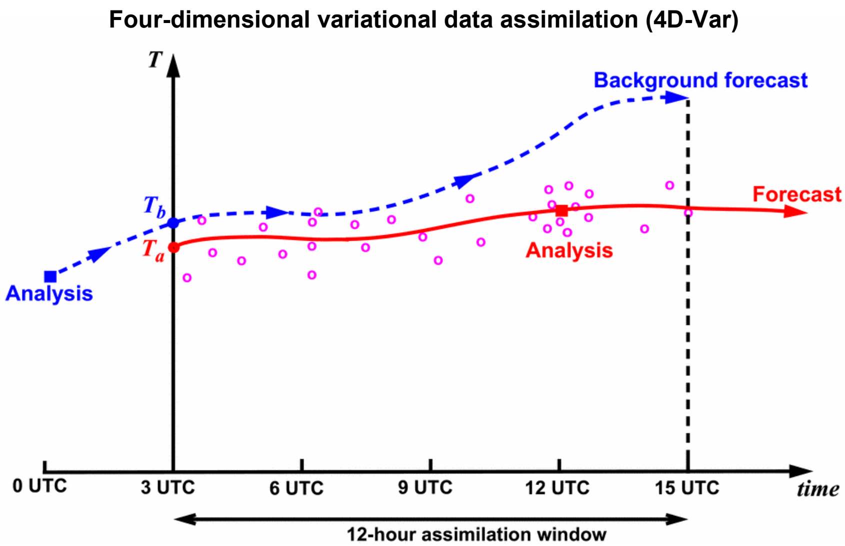

Figure 4: Schematic of the data assimilation process for the twelve-hourly four-dimensional data assimilation used operationally at ECMWF. The ordinate denotes a model variable, temperature for example.

background forecast. Fig. 4 shows the case of a twelve-hourly assimilation window as implemented at ECMWF. As illustrated in the figure, background and medium-range forecasts are actually run from analysis times of 00 and 12UTC, rather than the start of the assimilation window. This is because ocean-wave and land-surface conditions are adjusted at these times through separate analysis procedures and also because of the specific way the "incremental" formulation of 4D-Var (Courtier et al. 1994) is implemented for the atmospheric analysis.

A comprehensive introduction to variational data assimilation is given in Rabier's contribution to these Proceedings.

3. Some results from ERA-40

ERA-40 is a project in which observations made from September 1957 to August 2002 have been reanalysed using a six-hourly three-dimensional variational version (3D-Var) of the ECMWF data-assimilation system. The assimilating model used the same 60-level vertical resolution as is used currently for operational forecasting at ECMWF, but much lower horizontal resolution (T159 rather than T511 spectral truncation; a corresponding grid resolution of about 125km rather than a little under 40km.) The data assimilation was based on a version (cycle 23r4) of the forecasting system operational in the second half of 2001, modified to include a few newer features subsequently introduced operationally in cycles 25r1 and 25r4.

Production of the ERA-40 analyses was possible only because of the considerable external support the project received. Most of the older observations were supplied by NCAR via NCEP, EUMETSAT supplied reprocessed satellite winds, sea-surface-temperature and sea-ice analyses were provided by the Met Office and NCEP, and additional observational datasets were supplied by many other organizations. Formal validation partners (KNMI, MPI Hamburg, Météo-France, the Met Office, NCAR and Reading University) and others provided valuable feedback. Considerable funding was provided by the European Commission through the Fifth Framework Programme, computing support was provided by Fujitsu Ltd, support towards planning meetings and a workshop was provided by WCRP and GCOS, and staff were seconded by IAP, JMA and PCMDI to work on the development phase of the project.

Production of the ERA-40 analyses was completed in April 2003. Products on a regular 2.5^{o} grid are available from ECMWF’s public data server for research and educational use, and the complete set of full-resolution products is available from ECMWF’s Meteorological Archive and Retrieval System (MARS) and from ECMWF Data Services. Sets of products are (or will become) available also from some national data centres. Ten-day forecasts are being produced from each of the 00UTC and 12UTC ERA-40 analyses for each complete calendar year, employing the model version used for the ERA-40 data assimilation.

Considerable further information on the project may be viewed on the ECMWF website (http:// www.ecmwf.int).

The global observing system has changed considerably over the years since it was enhanced in 1957 in preparation for the International Geophysical Year of 1958. The ever-present observation types are the synoptic surface observations from land stations and ships, and the vertical soundings from radiosonde and pilot balloons. The accuracy of the radiosonde measurements improved over the period, but spatial and temporal coverage declined. Fig. 5 compares, for example, the distribution and frequency of the soundings available to ERA-40 for the months of March 1958 and March 1997. 20% fewer soundings were available for the more recent month. In 1958 the North Atlantic Ocean was covered by soundings from a network of fixed weather ships, and coverage at high northern latitudes was better than in 1997. In 1997, a much reduced number of soundings were available over the former Soviet Union and the North American soundings did not match the quite complete twice-a-day record achieved in 1958.

The other observation types illustrated in Fig. 1 have more than compensated for the decline in radiosonde coverage. Fig. 6 shows counts, using a logarithmic scale, of the aircraft, buoy, satellite-radiance and satellite-wind observations used daily in ERA-40 from September 1957 until the end of 2001. Vertical lines mark the enhancements in data availability that occur at the beginning of 1973 and in late 1978. 1973 sees a significant increase in the number of available aircraft observations, the first few buoy observations, and radiances available from the first of the VTPR instruments flown on the early NOAA satellites. 1979 sees the major enhancement of the observing system that took place initially for the Global Weather Experiment (or FGGE, the First Global Experiment of the Global Atmospheric Research Programme.) The VTPR data are replaced by data from the series of three (MSU/HIRS/SSU) TOVS instruments, winds first become available from geostationary satellites, ozone data become available from the TOMS and SBUV instruments, and there is a substantial increase in buoy and aircraft data. Observation counts decline for a while after the end of FGGE, but generally increase thereafter, helped by the additional satellite winds derived specifically for ERA-40 by EUMETSAT's reprocessing of the data from Meteosat-2 (1982-1988). Newer satellite instruments from which data are assimilated in ERA-40 are SSM/I from 1987 onwards, the ERS altimeter from 1991, the ERS scatterometer from 1993 and AMSU-A from 1998.

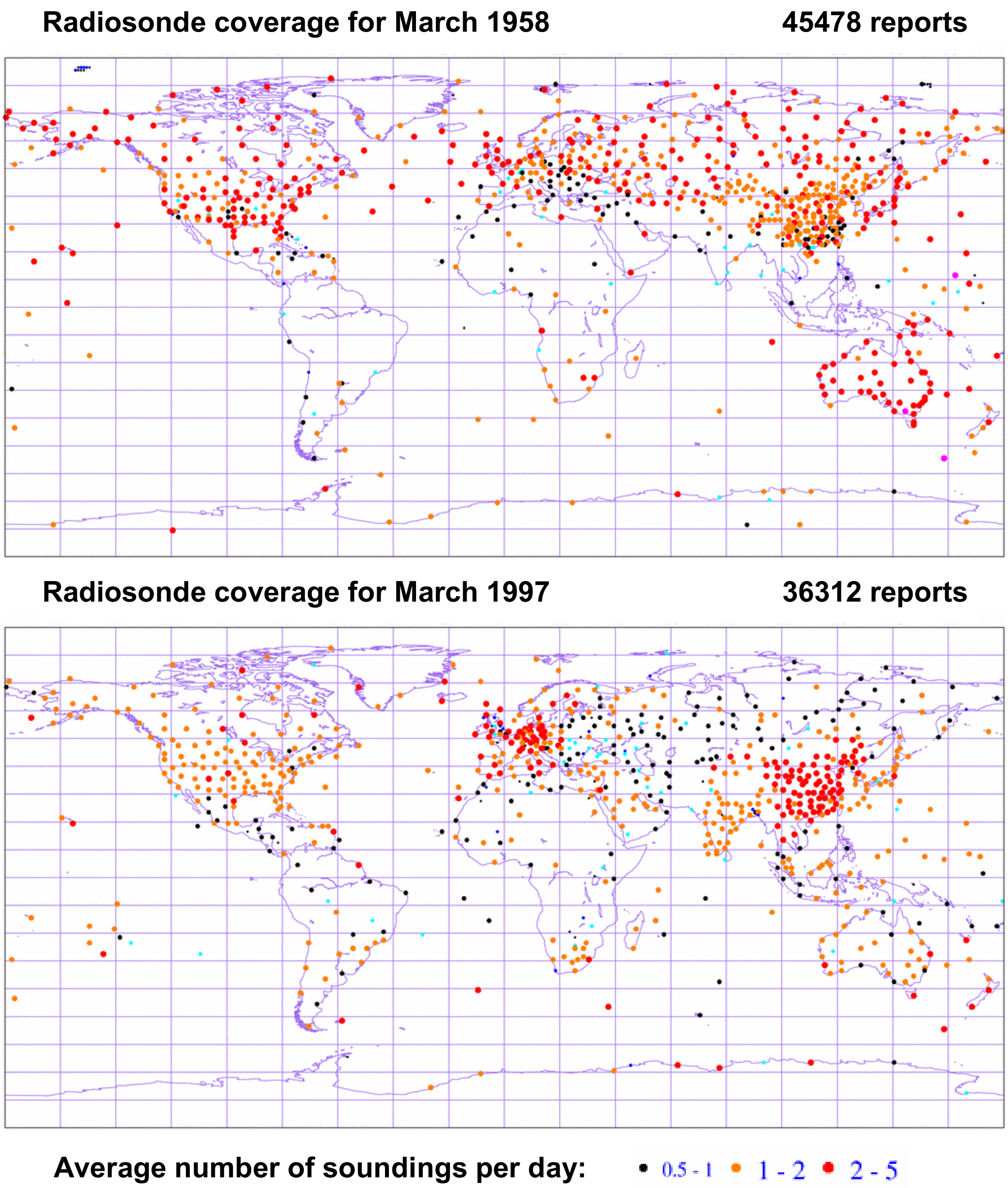

Figure 5: ERA-40 radiosonde data coverage for March 1958 (top) and March 1997 (bottom).

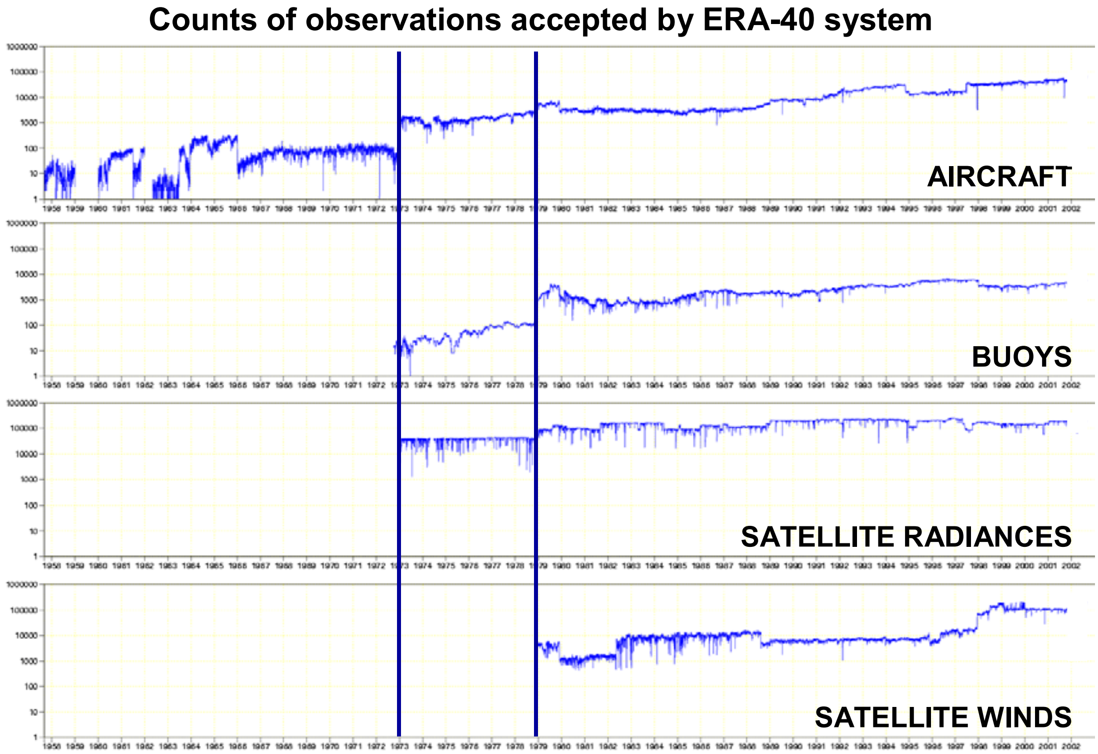

Figure 6: Counts (plotted with logarithmic scale) of observations used daily in ERA-40 from September 1957 until the end of 2001. Vertical lines mark the enhancements in data availability that occur at the beginning of 1973 and in late 1978.

The improvement in the observing system over the past four decades is reflected in a marked reduction in the general magnitude of analysis increments over the period of ERA-40. Fig. 7 presents maps of the root-mean-square increments in 500hPa height for the years 1958 and 2001. Increments are much smaller in 2001 over the land and neighbouring ocean regions where there are marked increments in 1958. Isolated small-scale maxima over the oceans in 1958 indicate where radiosonde data from islands and the fixed weather ships correct the background forecast. Local impact is still evident in 2001 for isolated radiosonde stations over Antarctica and northern Russia. Increments are particularly large along the west coast of North America in 1958, where background forecast error that developed over the poorly observed Pacific Ocean can be corrected on its first encounter with the North American radiosonde network. Larger increments occur in 2001 than in 1958 over oceanic regions that are almost devoid of observations in 1958, due to assimilation of the satellite, buoy and aircraft data that today provide coverage of such regions. Over regions well covered by radiosonde data in both years, smaller increments are made in 2001 because of both more accurate background forecasts and more accurate radiosonde measurements.

Generally smaller analysis increments can also result from improvements in data assimilation systems. The left-hand plots in Fig. 8 show root-mean-square 500hPa height increments over the southern hemisphere for 1989 from ERA-40 and from the earlier ERA-15 reanalysis, which used the optimal interpolation method superseded by 3D-Var for ECMWF operations in early 1996. Increments are generally smaller for ERA-40 than for ERA-15. This is a good result for ERA-40 if it stems from a greater accuracy of the background forecasts (in closer agreement with the observations) rather than from a failure of the analysis to draw sufficiently closely to observations. The former appears to be the case, as the 24-hour ERA-40 forecasts generally match radiosonde data better than the 24-hour ERA-15 forecasts, as illustrated by the profiles of root-mean-square temperature and vector-wind fits evaluated over the extratropical southern hemisphere shown in the right-hand plots of Fig. 8.

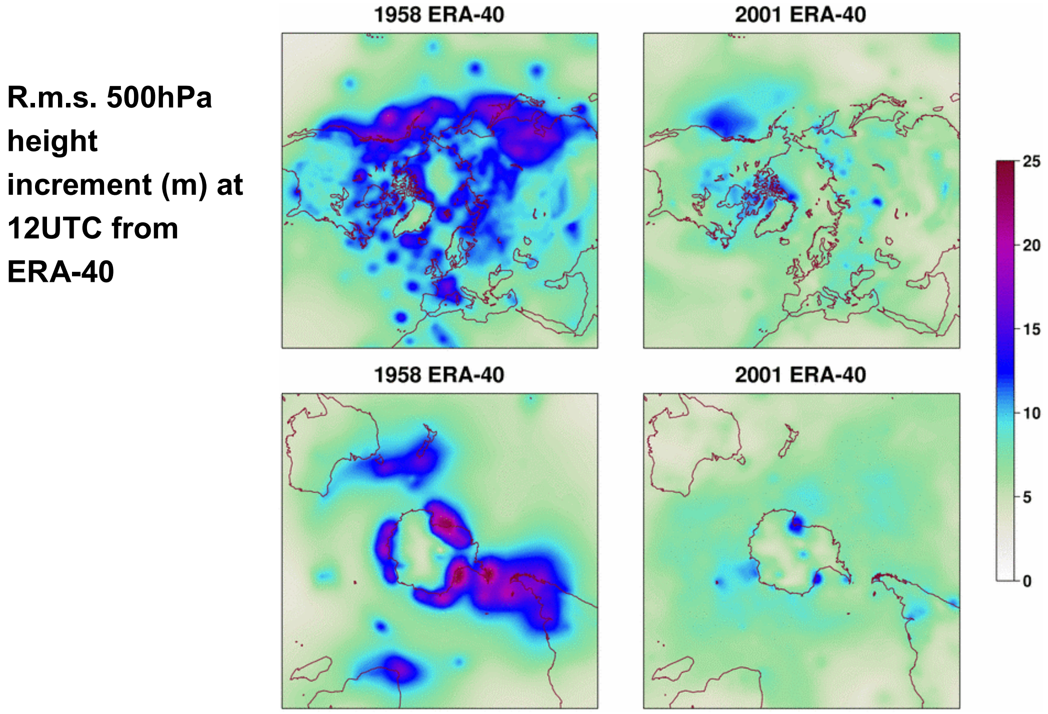

Figure 7: Root-mean-square analysis increments in 500hPa height (m) over the northern (upper) and southern (lower) hemispheres from ERA-40 for 1958 (left) and 2001 (right).

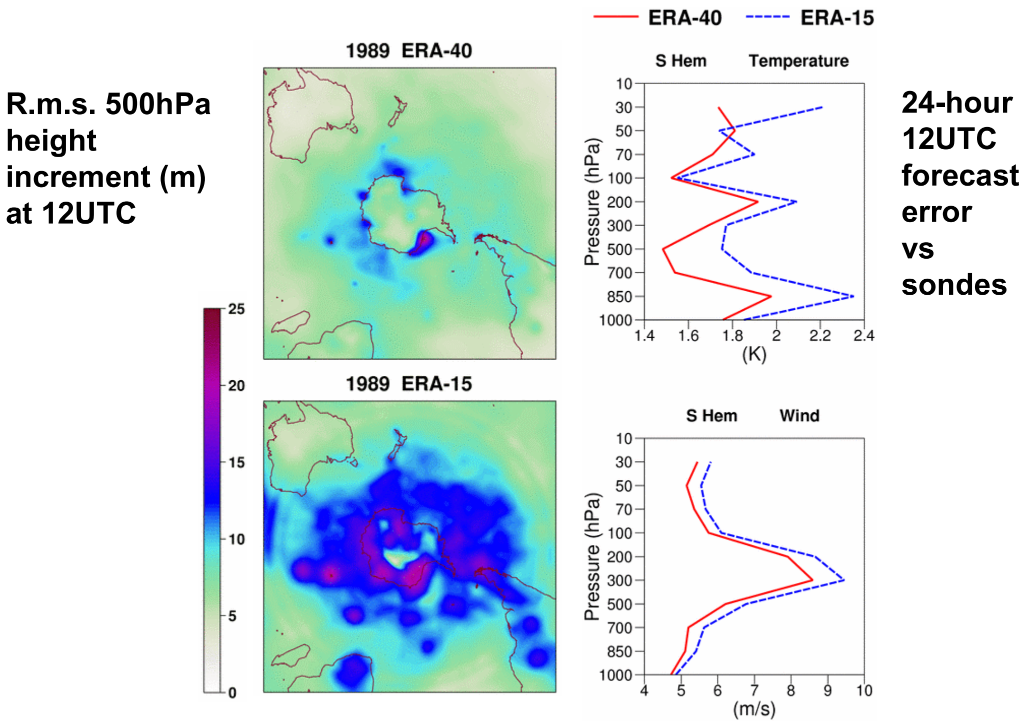

Figure 8: Root-mean-square analysis increments in 500hPa height over the southern hemisphere for 1989 from ERA-40 (upper left) and ERA-15 (lower left) and vertical profiles of root-mean-square errors of 24-hour forecasts verified against radiosondes over the extratropical southern hemisphere for temperature (K, upper right) and vector wind (ms^{-1}, lower right)

Fig. 9 shows time series of background and analysis fits to SYNOP surface pressure measurements for the extratropical southern hemisphere. The number of observations used per day is also shown. Variations in the latter are a consequence not only of variations in the number of observations available but also of variations in the acceptance of the available observations by the data assimilation system, as data that deviate too far from the background are rejected in the data-screening that precedes the variational analysis.

A general improvement in the fit to the surface-pressure measurements occurs over the period of ERA-40. Changes in data coverage might affect this, as increased coverage in the subtropics, where variance is lower, would tend to lower values in plots such as these in which no area-weighting is applied. However, the sudden improvement in fit at the beginning of 1979 almost certainly is a consequence of the major improvement of the whole observing system that took place at the time, as illustrated in Fig. 6. An improvement at the beginning of 1973 can also be seen.

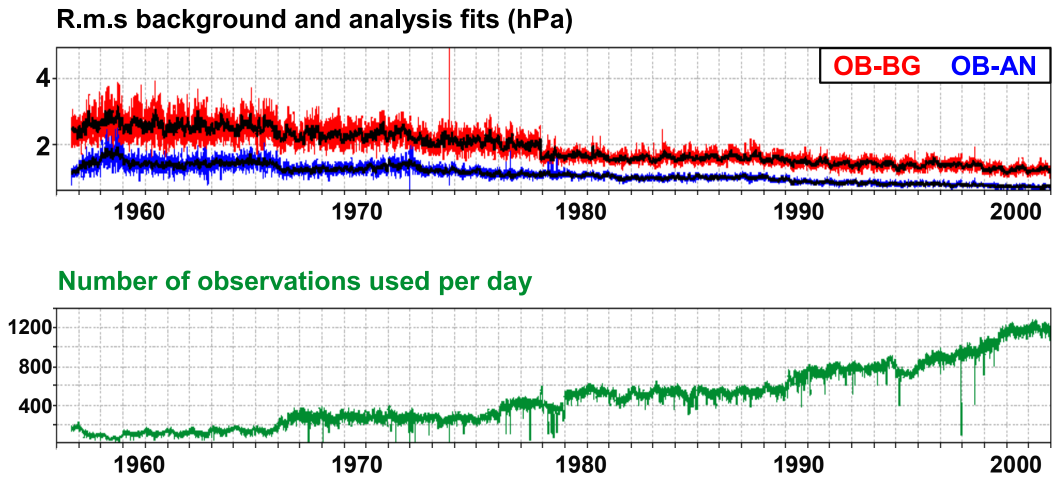

Figure 9: Root-mean-square ERA-40 background (red: daily; black: 15-day moving average) and analysis (blue: daily; black: 15-day moving average) fits to 00UTC SYNOP surface pressure observations over the extratropical southern hemisphere (upper) and the number of these observations used per day (lower).

Further evidence of the improvement of the analyses over time is provided by objective verification of the medium-range forecasts being run from the ERA-40 analyses. These forecasts are being produced using the same T159 version of the model used in the ERA-40 data assimilation. The upper panel of Fig. 10 shows (in colours varying from yellow to magenta) anomaly correlations of 500hPa height over the extratropical northern hemisphere for those years from 1958 to 2001 for which a complete set of ten-day forecasts has been completed. There is a general improvement of forecast quality throughout the period, but even in the early years the skill level is quite high, with the 60% anomaly correlation reached close to six days ahead or beyond in each annual average. Interannual variations in predictability are also evident; of the years completed to date, the highest correlations later in the forecast range occur not for a recent year but for 1981. Relatively low values occur for 1999, as they did in operations at the time, as will be seen in the following section.

Also shown in the upper panel of Fig. 10 are the results for the northern hemisphere from ECMWF operations for 1980, 2001 and for the 12 months ended 31 August 2003. ERA-40 performance for even the earliest years is considerably better than the performance of ECMWF operations in 1980. In contrast, ERA-40 performance for 2001 is considerably poorer than that of ECMWF operations for this year. The ERA-40 assimilation system was based on the version of the forecasting system that was operational in the second half of 2001. It thus differed from operations in 2001 principally because of its use of a 3D-Var rather than a 4D-Var analysis, and its use of T159 rather than T511 horizontal resolution for the background and medium-range forecasts. The results for 2002/03 indicate further recent improvement of the operational system.

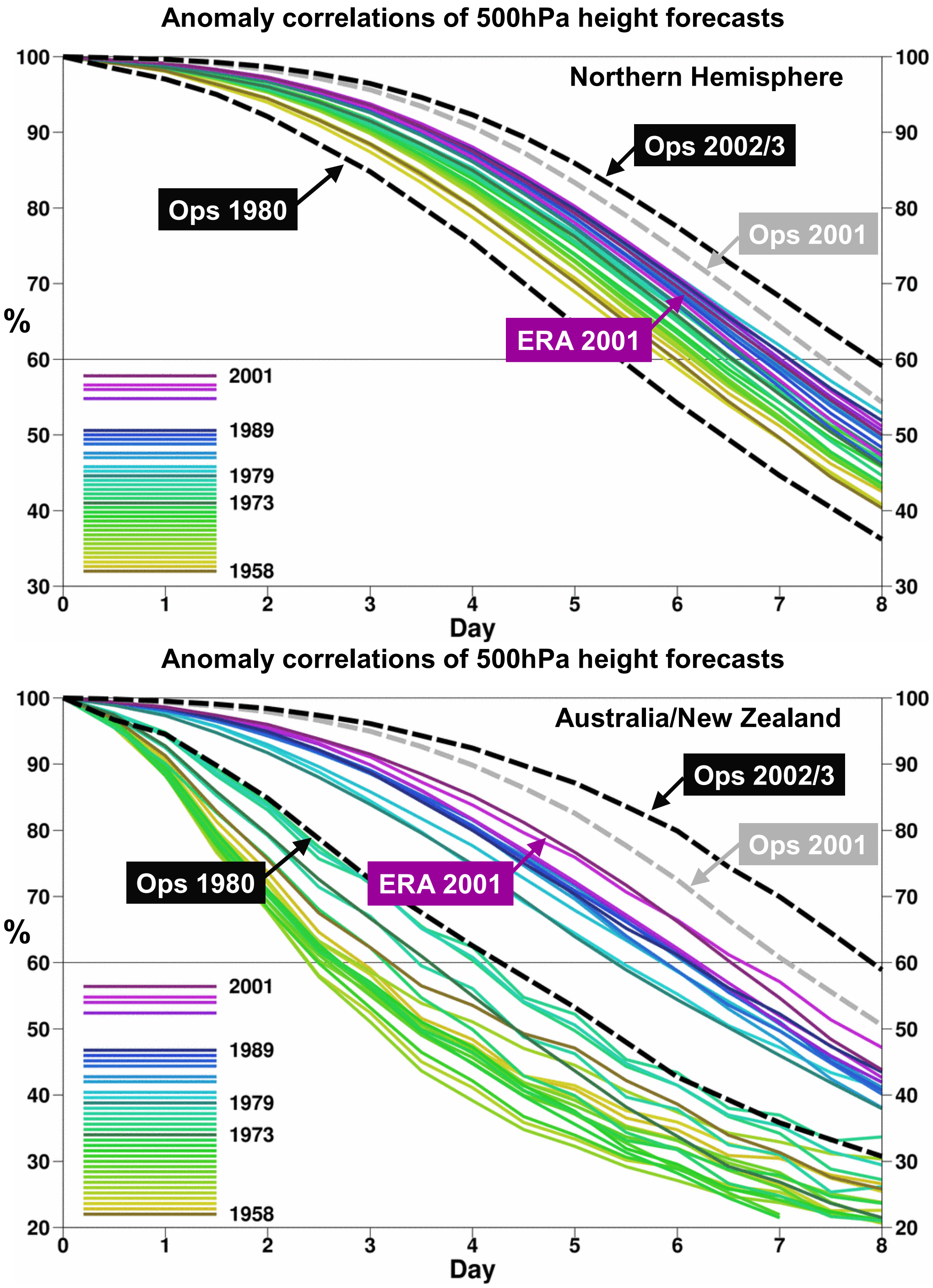

Figure 10: Anomaly correlations of 12UTC 500hPa height forecasts as functions of forecast range for the extratropical northern hemisphere (upper) and for a smaller region encompassing Australia and New Zealand (lower). Results are shown for many of the ERA-40 years (verified against corresponding ERA-40 analyses, denoted by the colour scale shown in legend) and for ECMWF operations (verified against the corresponding operational analyses, labeled Ops) for the calendar years of 1980 and 2001 and for the year ending 31 August 2003.

Overall, the ERA-40 system appears to be about as skilful as the operational ECMWF system in 1998 or 1999, and about as skilful as the operational (3D-Var) systems of the Met Office and NCEP in 2001 (not shown), at least as regards 500hPa height forecasts. The northern-hemisphere results in Fig.10 show that whilst the observing system for medium-range prediction has improved over the years, a greater improvement in forecasts has been derived from the improvements in data assimilation and forecast models achieved since 1980.

A different picture is seen for the southern hemisphere. The lower panel of Fig. 10 shows anomaly correlations for a region including Australia and New Zealand where observational coverage is sufficient for some reliance to be placed on the quality of the verifying analyses throughout the period. Forecast quality is poor in the 1950s and 1960s, picking up a little in the mid-seventies when radiances from the VTPR instruments on the early NOAA satellites were assimilated. A dramatic jump in forecast quality occurs at the end of 1978, when as noted earlier the observing system was improved considerably with the introduction of radiances from the TOVS instruments and the addition of winds from geostationary satellites and many more data from drifting buoys and commercial aircraft. Observing-system improvements beyond 1979 have had larger impact on southern- than northern-hemisphere forecast accuracy, bringing forecast-skill levels closer. The same is true of forecasting-system improvements. This is seen not only by comparing the extent to which the ERA-40 and operational medium-range forecasts improve between 1980 and 2001, but also by examining the results of the pre-operational trials of recent forecasting system changes discussed below.

4. Recent improvements in operational forecasts

Fig. 11 presents root-mean-square errors of forecasts of 500hPa height and mean-sea-level pressure for the extratropical northern and southern hemispheres. Time series from 1990 onwards are shown for three-day and five-day forecasts from three global prediction systems, that of ECMWF and those of the Met Office and NCEP, the two national centres that come closest to matching ECMWF's performance according to these measures of forecast accuracy. ECMWF results for the four-day range are also presented. The plots show annual running means derived from the verification statistics that forecasting centres exchange monthly under the auspices of the World Meteorological Organization. Each centre's forecasts are verified by

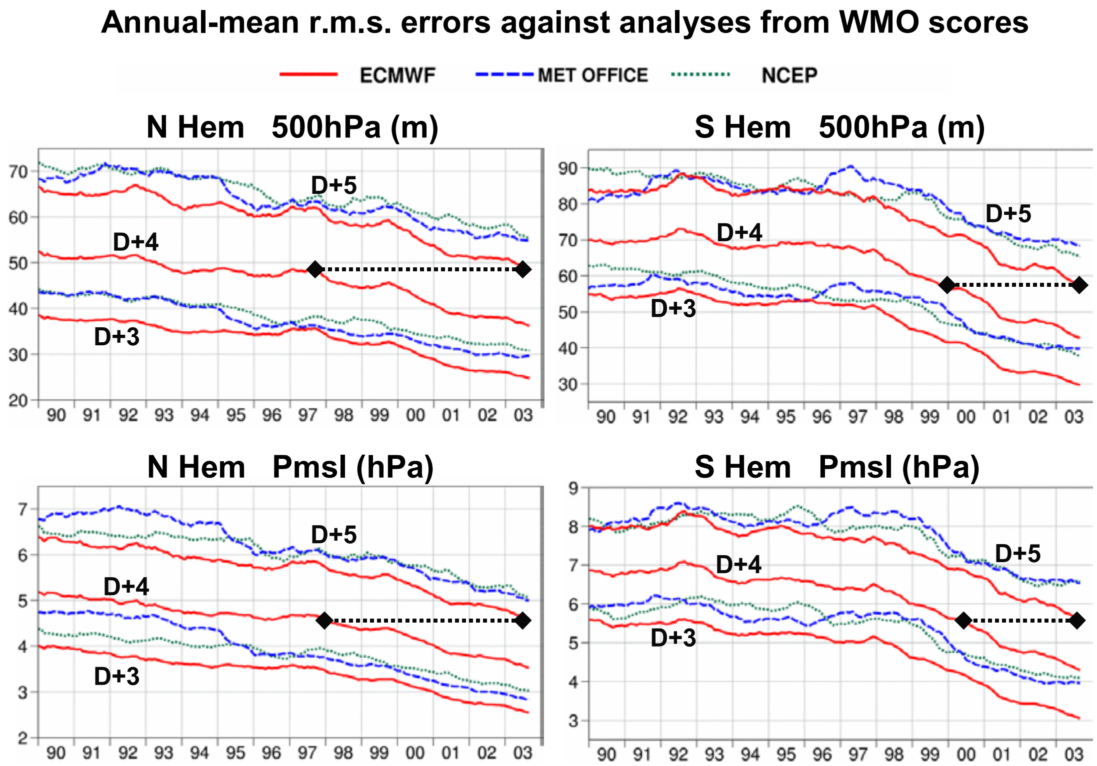

Figure 11: Root-mean-square errors of 3- and 5-day forecasts of 500hPa height (upper; m) and mean-sea-level pressure (lower; hPa) for the extratropical northern (left) and southern (right) hemispheres. Results from ECMWF, the Met Office and NCEP are plotted in the form of annual running means of all monthly data exchanged by the centres from January 1989 to August 2003. ECMWF 4-day forecast errors are also shown. Values plotted for a particular month are averages over that month and the 11 preceding months, so that the effect of a forecasting-system change introduced in that month is seen from then onwards.

comparison with its own analyses. Results are presented for initial forecast times of 12UTC for ECMWF and the Met Office, and 00UTC for NCEP. Additional Met Office results from 00UTC have been available since 1995, but are not plotted as they are very similar to those presented for 12UTC. The ECMWF forecasts are routinely produced with a cut-off time for data reception prior to the final analysis cycle that is several hours later than used by the other centres, but evidence from ECMWF forecasts produced with earlier cut-off times indicates that differences in the forecasting systems rather than in data reception are the primary cause of the differences in forecast accuracy illustrated here.

Fig. 11 shows a general trend towards lower forecast errors in both hemispheres, for both 500hPa height and mean-sea-level pressure. The improvement since 1996 in ECMWF forecasts for the northern hemisphere amounts to around a one-day extension of the forecast range at which a given level of error is reached, today's four- and five-day forecasts being respectively about as accurate on average as the three- and four-day forecasts of six or seven years ago. The rate of improvement over this period has been especially rapid in forecasts for the southern hemisphere, amounting to a one-day gain in predictability in just three to four years.

The starting point for the rapid recent improvement in ECMWF forecasts shown in Fig. 11 was the operational introduction of four-dimensional variational (4D-Var) data assimilation (Mahfouf and Rabier 2000, and refs.) in late November 1997. The improvements in short-range ECMWF forecasts since then can be linked very directly to a series of subsequent forecasting-system changes. Table 1 shows the principal

IFS cycle	Date of change	Reduction in rms error (m) NHem	SHem	Nature of change
18r1	1997/11/25	0.47	1.77	4D-Var, with six-hourly cycling
18r6	1998/06/29	0.42	0.05	Use of sig and std level T,q from sondes, improved QC and serial error correlation for SYNOP/DRIBUs. Coupled wave model
19r2	1999/03/09	0.37	0.44	Increased stratospheric resolution
21r1	1999/05/05	0.22	0.90	Direct assimilation of MSU and new AMSU-A radiances
21r4	1999/10/12	1.49	2.31	Error statistics for Jb from EnsDA, SSMI wind data, corrected use of humidity observations, increased PBL resolution, improved cloud, convection and orography
22r3	2000/06/27	0.57	1.94	New error variances (for Jb and Jo), additional ATOVS data, improved land-surface, sea-ice and LW radiation
23r1	2000/09/12	0.42	0.65	Twelve-hourly 4D-Var and related refinements of data assimilation system
23r3	2000/11/21	0.80	0.66	Increased (T511/T159) model/analysis resolution, increased angular resolution in wave model
24r3	2002/01/22	0.34	0.42	Assimilation of Quikscat and additional ATOVS data, finite elements in vertical, improved land-surface, sea-ice and LW radiation
25r4	2003/01/14	0.19	0.74	Assimilation of GOES clear-sky radiances, MODIS winds, more HIRS, SAR, third ATOVS satellite, direct use of SSM/I, improved Jb and incremental 4D-Var algorithm, improved cloud and convection schemes

Table 1: The operational changes to the ECMWF forecasting system over the past six years that had the largest impact on root-mean-square errors of one-day 500hPa height forecasts for the extratropical northern and southern hemispheres as measured in pre-operational trials.

changes, when they were implemented and the impacts on one-day forecast errors measured during pre-operational trials. In most cases impact is larger in the southern than in the northern hemisphere. Simmons and Hollingsworth (2002) showed that the actual annual-mean root-mean-square errors of one-day 500hPa height forecasts over the five years up to 2001 matched well the errors that would have occurred had the changes introduced between November 1997 and November 2000 given exactly the same average forecast improvements in operational use as were measured in the pre-operational trials. The same result holds when the calculation is extended to include the latest changes to the forecasting system and data use. It indicates that the overall recent improvement in short-range forecasts is indeed due overwhelmingly to changes to the forecasting system and availability of data rather than to circulation regimes that were unusually easy to predict in the last few years.

Verification of forecasts by comparison with radiosonde observations confirms the improvement in forecasts shown in Fig.11. Fig. 12 presents root-mean-square errors of three-, four- and five-day ECMWF 500hPa height and 850hPa wind forecasts verified against radiosondes from 1995 onwards. Values for the northern hemisphere 500hPa height are quite similar to those from verification against analyses at these forecast ranges. The only difference of note for the southern hemisphere is that the rapid reduction in height forecast errors is seen to begin some six months earlier in 1997 in Fig. 12 than in Fig. 11. This earlier fall in error measured by comparison with radiosonde data appears to be associated with the introduction in May 1997 of a new formulation for the background error constraint in the then-operational 3D-Var system (Derber and Bouttier 1999). This change brought a marked improvement in tropical forecasts, and may thus have given a greater improvement in verification against those radiosondes located in the southern subtropics than in verification against analyses over the whole southern extratropics. A major element of recent changes to the forecasting system involved wider utilization and new methods of assimilation of satellite data, so it is particularly reassuring to see a large reduction in forecast errors as measured against the southern hemisphere radiosonde network.

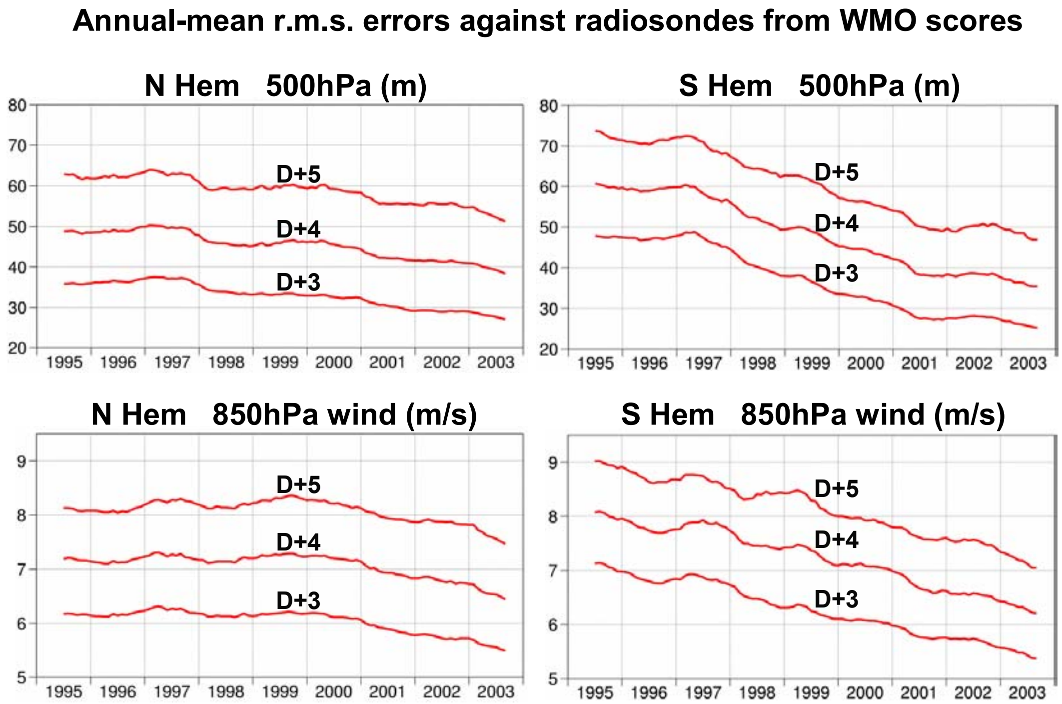

Figure 12: Root-mean-square errors of 3-, 4- and 5-day ECMWF 500hPa height (upper; m) and 850hPa vector wind (lower; ms^{-1}) forecasts for the extratropical northern (left) and southern (right) hemispheres, plotted in the form of annual running means of monthly data for verification against radiosondes from July 1994 to August 2003. Values plotted for a particular month are averages over that month and the 11 preceding months.

The levels of skill of northern and southern hemisphere forecasts cannot be compared simply in terms of root-mean-square errors because of inter-hemispheric differences in the levels of variance of the fields. Comparison can, however, be made directly in terms of anomaly correlation coefficients, which are closely related to mean-square errors normalized by corresponding variances (see e.g. Simmons *et al*. 1995). Fig. 13 presents anomaly correlations of 500hPa height based on ECMWF's operational three-, five- and seven-day forecasts from January 1980 to August 2001. Running annual means of the monthly-mean skill scores archived routinely over the years are plotted for the two hemispheres.

Fig. 13 shows a higher overall rate of improvement in the forecasts for the southern hemisphere. In the early 1980s, the skill levels of the three- and five-day forecasts for this hemisphere were only a little better than those of the five- and seven-day northern hemisphere forecasts. At the time this was not surprising in view of the sparsity of conventional ground-based and aircraft observations in the southern hemisphere (Bengtsson and Simmons 1983). Today, however, the skill at a particular forecast range in the southern hemisphere is only a little lower than that at the same range in the northern hemisphere. There is little doubt that improvements in the availability, accuracy and assimilation of satellite data have been major factors contributing to the relative improvement in forecast skill. Evidence for this comes both from the pre-operational testing of changes to the forecasting system and from observing system experiments such as discussed by Dumelow in these Proceedings.

Interannual variations in skill are also evident in Fig. 13, especially for the northern hemisphere at the five- and seven-day time ranges. In particular, there is a pronounced minimum in the northern hemisphere scores arising from relatively poor performance over the year to August 1999. A corresponding maximum can be seen in the time series of root-mean-square errors shown in the left panels of Fig. 11. This is evident for the Met Office forecasts as well as for those of ECMWF.

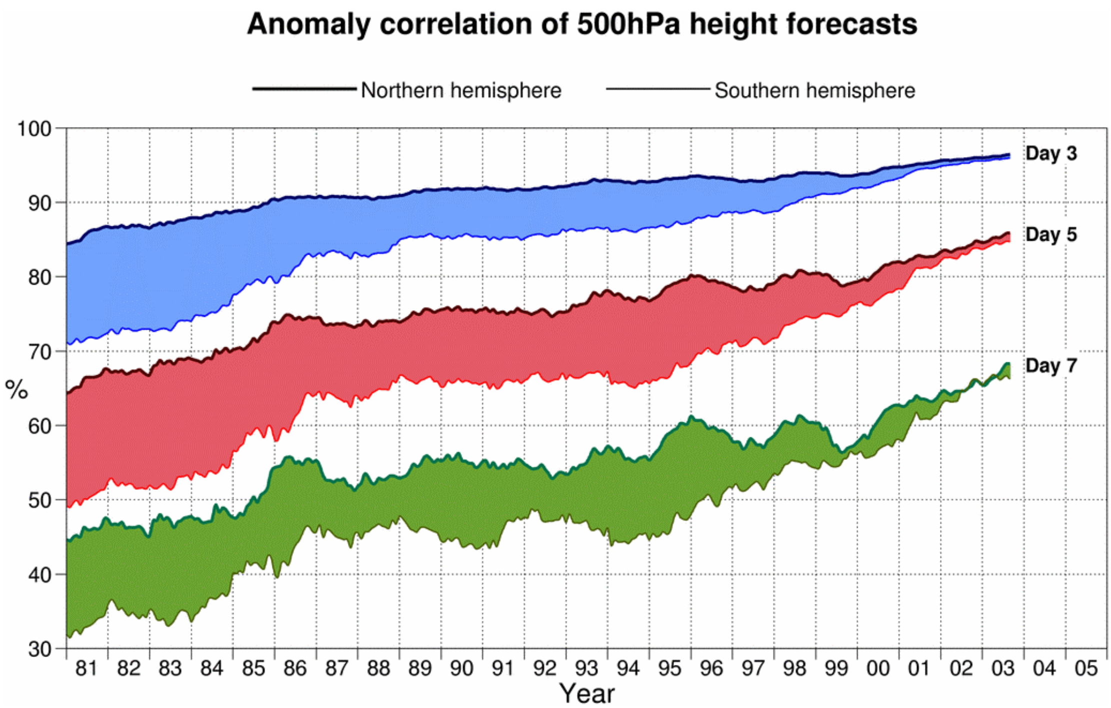

Figure 13: Anomaly correlation coefficients of 3-, 5- and 7-day ECMWF 500hPa height forecasts for the extratropical northern and southern hemispheres, plotted in the form of annual running means of archived monthly-mean scores for the period from January 1980 to August 2003. Values plotted for a particular month are averages over that month and the 11 preceding months. The shading shows the differences in scores between the two hemispheres.

5. Analysis and forecast differences

Several aspects of the performance of data assimilation and forecast systems can be illustrated by comparing results from the systems of different centres, or by comparing successive forecasts from one particular centre.

Table 2 presents root-mean-square differences between ECMWF and Met Office analyses of 500hPa height over the extratropical northern and southern hemispheres. Results are shown for the periods from December to February (DJF) and from June to August (JJA) for each of the past six years. Differences between the analyses have been reduced very substantially over these years, particularly for the southern hemisphere. The fact that the analyses from the two centres have become much closer to each other does not necessarily imply that both sets have become much more accurate, but given the substantial improvements in forecast accuracy achieved by both centres it may be inferred that both centres' analyses have indeed become significantly closer to the truth.

	N Hem DJF	S Hem DJF	N Hem JJA	S Hem JJA
1998	**14.1**	**21.2**	**10.6**	**29.7**
1999	**15.6**	**21.4**	**10.3**	**27.4**
2000	**12.2**	**16.3**	**8.9**	**17.0**
2001	**9.8**	**11.6**	**8.2**	**14.2**
2002	**10.1**	**11.5**	**8.3**	**12.3**
2003	**7.8**	**10.2**	**5.5**	**10.7**

Table 2: Root-mean-square difference between ECMWF and Met Office 500hPa height analyses (m) over the extratropical northern and southern hemispheres for December to February (DJF) and June to August (JJA) for the past six years.

Maps of the mean-square differences between the ECMWF and Met Office analyses for DJF for the years 1997/98 (labelled 1998 in Table 2), 2000/01 and 2002/03 are shown in Fig. 14. The reductions in mean-square differences indicated in Table 2 can be seen to have occurred at virtually all locations. Difference patterns, however, remain largely the same. They are relatively small over substantial areas of the northern continental land masses, but are larger in a band stretching from the Arctic southwards over central Asia, where radiosonde coverage is poorer than over other northern land areas, and also over western and northern North America, where the background fields of the data assimilation suffer from the lower accuracy of upstream analyses over the Pacific and Arctic oceans. Differences are larger over the mid- and high-latitude oceans, particularly in the southern hemisphere, and the largest values for each hemisphere are generally found at polar latitudes. Local minima can also be seen, for example at the South Pole and around parts of the coastline of Antarctica, reflecting the availability of radiosonde observations which both analyses fit closely.

Differences for the latest DJF are compared with those for the latest JJA in Fig. 15 (with different contouring than used in Fig. 14). Differences in the middle-latitude storm tracks are, not surprisingly, larger in winter than summer. However, the largest northern hemisphere differences (and by inference relatively large analysis errors) occur over the Arctic in summer, where *in-situ* measurements are sparse and satellite data are difficult to use in the lower troposphere due to the nature of the underlying surface. The geographically- and seasonally-fixed background-error correlations used in the current ECMWF analysis system may also be less representative for this region than elsewhere. Propagation and amplification of forecast error that originates from analysis error over the Arctic is a known source of medium-range error over Europe in summer. It was, for example, particularly prevalent in the summer of 1999 (Simmons *et al.,* 2001).

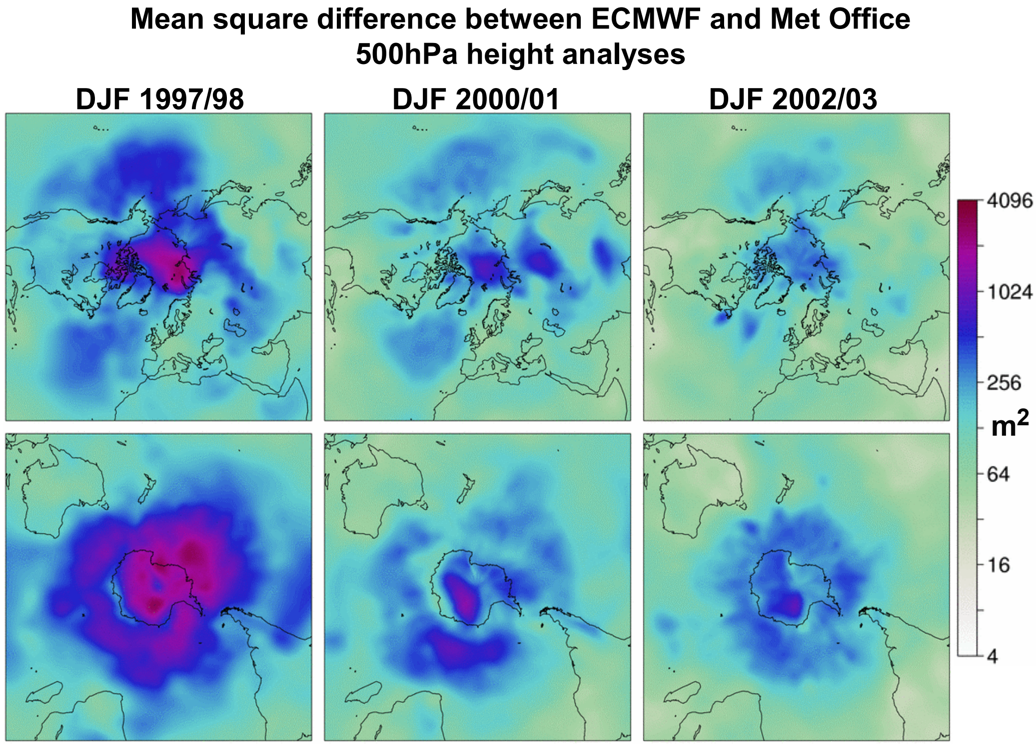

Figure 14: Mean square differences (m^2) between the 500hpa height analyses of ECMWF and the Met Office for the northern (upper) and southern (lower) hemispheres computed over the months of December, January and February for the years 1997/98 (left), 2000/01 (centre) and 2002/03 (right).

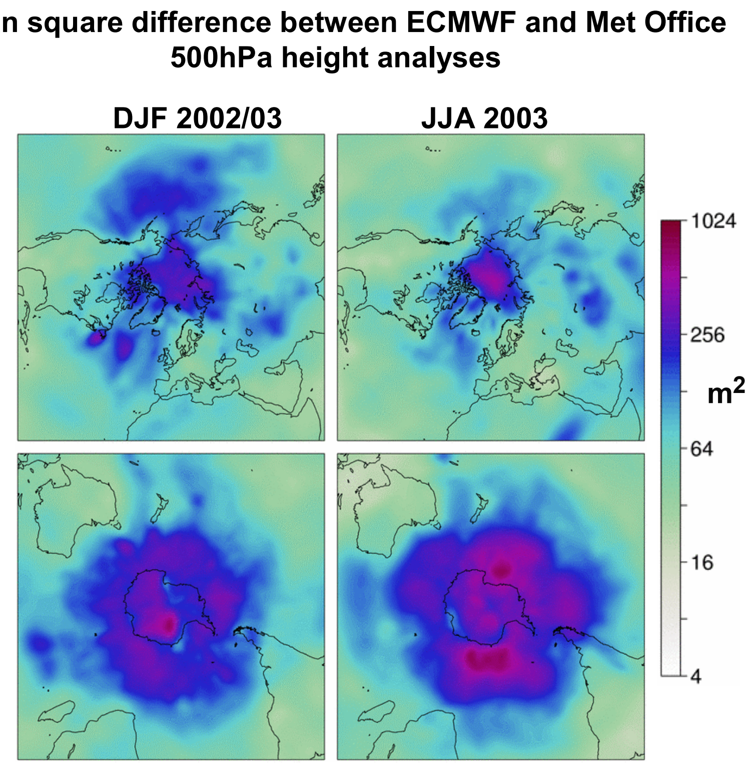

Figure 15: Mean square differences (m^2) between the 500hpa height analyses of ECMWF and the Met Office for the northern (upper) and southern (lower) hemispheres computed over the months of December, January and February 2002/03 (left) and June, July and August 2003 (right).

The fact that differences between the analyses of ECMWF and the Met Office (and between ECMWF and NCEP for that matter) have become substantially smaller in recent years does not imply for certain that analysis errors have declined, as the data assimilation systems of the different centres could have converged to such an extent that each share common pronounced errors. In this case, however, one would expect that the short-range forecast errors of the centres would be highly correlated, which is not the case.

The upper panel of Fig. 16 shows correlations between ECMWF and Met Office forecast errors (with verification against corresponding analyses) computed for DJF 2002/03. They are plotted as functions of forecast range, for the extratropical northern and southern hemispheres. Under perfect-model assumptions and neglecting error in the verifying analyses, such correlation coefficients would asymptote to the value 0.5 for large forecast ranges (when the two sets of forecasts and the verifying analyses can be regarded as three sets of random deviations from climatology). The panel shows that correlation coefficients close to 0.5 are in fact reached quite early in the forecast range, after about day 2. Correlations are weaker at shorter time ranges. Extrapolating correlations and the measured forecast errors themselves (shown in Fig. 17) enables estimates of analysis errors to be made, as discussed by Simmons and Hollingsworth (2002). These indicate values of the order of 4 to 5m for 500hPa height for the ECMWF analyses over the most recent winter and summer, about half the observation error of radiosonde measurements.

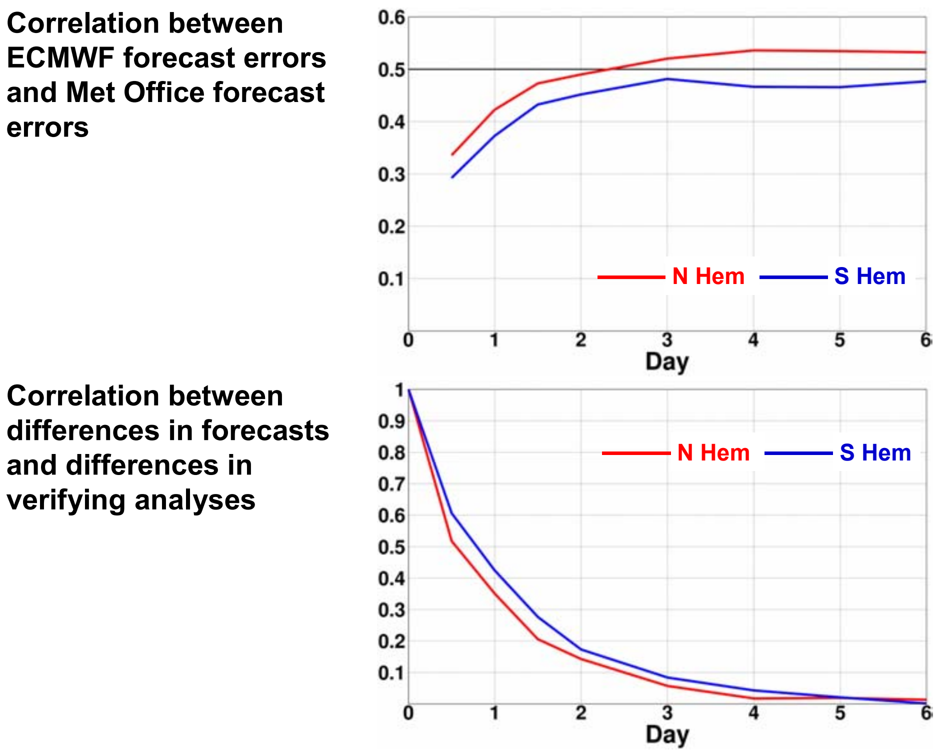

Figure 16: Correlations as a function of forecast range computed for the 500hPa height field over the extratropical northern (red) and southern (blue) hemispheres for 12UTC verifying dates from 1 December 2002 to 28 February 2003. The upper panel shows the correlations between the errors of ECMWF and Met Office operational forecasts (with the forecast from a particular system verified against subsequent analyses from the same system). The lower panels show the correlations between the differences in the forecasts and the differences in the verifying analyses of ECMWF and the Met Office.

The lower panel of Fig. 16 shows correlations between differences in the forecasts and differences in the verifying analyses of ECMWF and the Met Office. The correlation is unity at the start of the forecast range, when the forecast difference is equal to the analysis difference. The correlation drops towards zero with increasing forecast range, as the assimilation of observations drives successive background forecasts and (verifying) analyses further away from the free-running forecast. The correlation falls slightly quicker for the northern hemisphere, due presumably to its better coverage by *in-situ* data. It takes of the order of four to five days for sufficient observations to have been assimilated for 500hPa height analysis differences to lose virtually all memory of earlier differences in background forecasts.

If the correlation between differences in forecasts and differences in verifying analyses illustrated in Fig. 16 is representative of the correlation between true forecast error and error in the verifying analyses, then estimates may be made of the effect of error in the verifying analysis on the measured forecast errors. Simmons and Hollingsworth (2002) discuss this, and show that the effect appears to be rather small. Measured hemispheric root-mean-square errors of recent 500hPa height forecasts are estimated to be within a metre or so of the true errors at the one- and two-day ranges. In contrast, root-mean-square errors of short-range forecasts verified against radiosonde data tend to be dominated by the observation error of the radiosonde measurements, as shown for example by Simmons and Hollingsworth (*loc. cit.*) for the 500hPa height field and by Simmons et al. (2003) in a study of stratospheric temperature and wind forecasts.

Fig. 17 compares root-mean-square errors of ECMWF and Met Office 500hPa height forecasts as functions of forecast range, evaluated over the extratropical northern and southern hemispheres. Results are shown for all forecasts from 12UTC start times verifying in DJF 2002/03. ECMWF forecasts verifying at 12UTC but starting from 00UTC are also shown, with the forecast range shifted by 12 hours in the plots so that, for example, a 36-hour forecast error is plotted as if the range was 24 hours. Results for the two subsequent seasons of 2003 are generally similar.

Fig. 17 shows that the forecasts from ECMWF are considerably more accurate on average than those from the Met Office made with the same starting time, by around 12 hours in the northern hemisphere and by more in the southern hemisphere, as judged by the forecast range at which a particular level of error occurs. The advantage of around 12 hours or more exists throughout the forecast range. The implication is that the initial analysis error is considerably lower in the ECMWF system, leading to lower subsequent forecast errors at all ranges. The similarity of the error graphs for the Met Office forecasts and the 12-hour-old ECMWF forecasts shows that overall error growth rates in the ECMWF and Met Office 500hPa height forecasts are in fact very similar, both forecast models simulating quite realistic levels of variance and amplifying errors of similar magnitude at a similar rate.

Fig. 18 provides a further indication of the link between reducing forecast error in the very short range (and by implication reducing analysis error) and reducing error in the medium range. It shows smoothed time series of the ranges at which ECMWF's operational 500hPa height forecasts have reached certain levels of anomaly correlation over the past two decades. In recent years the improvement of forecasts as measured by the increases in these forecast ranges has been by an amount that varies little beyond a day or so ahead. Improvements in medium-range 500hPa height forecasts thus appear to have stemmed directly from model, analysis and observing-system improvements that have reduced analysis and short-range forecast error.

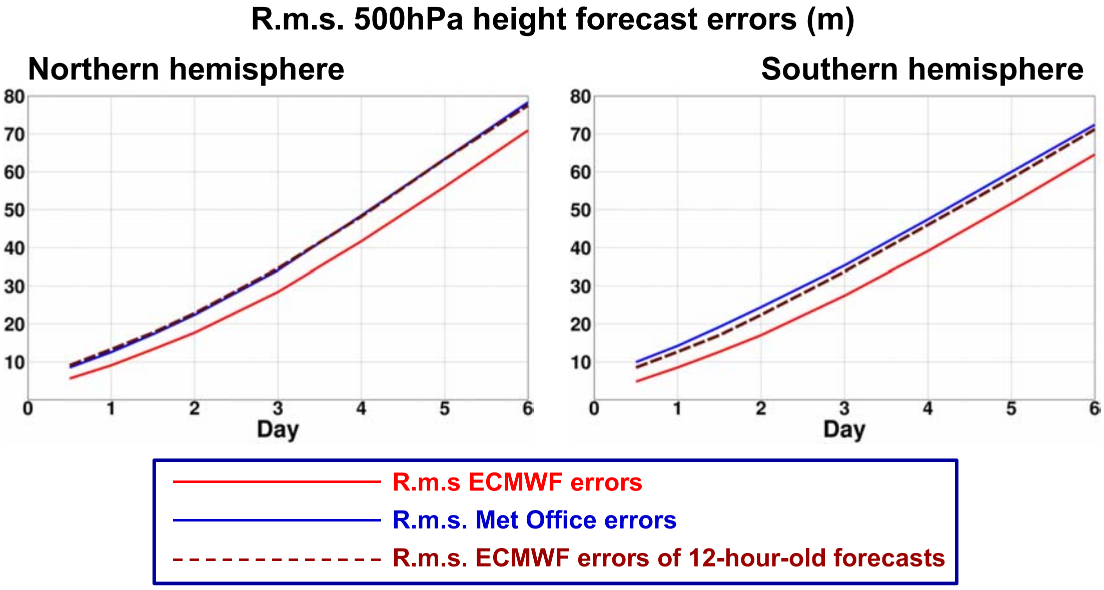

Figure 17: Root-mean-square errors of 500hPa height forecasts (m) for the extratropical northern (left) and southern (right) hemispheres verifying at 12UTC in the period from 1 December 2002 to 28 February 2003. Results are shown for the operational forecasts from ECMWF (red, solid) and the Met Office (blue, solid) at the forecast range indicated on the ordinate. Also shown (brown, dashed) are the errors of operational ECMWF forecasts verifying at the same times but initiated 12 hours earlier (plotted with 12-hour shift in the forecast range).

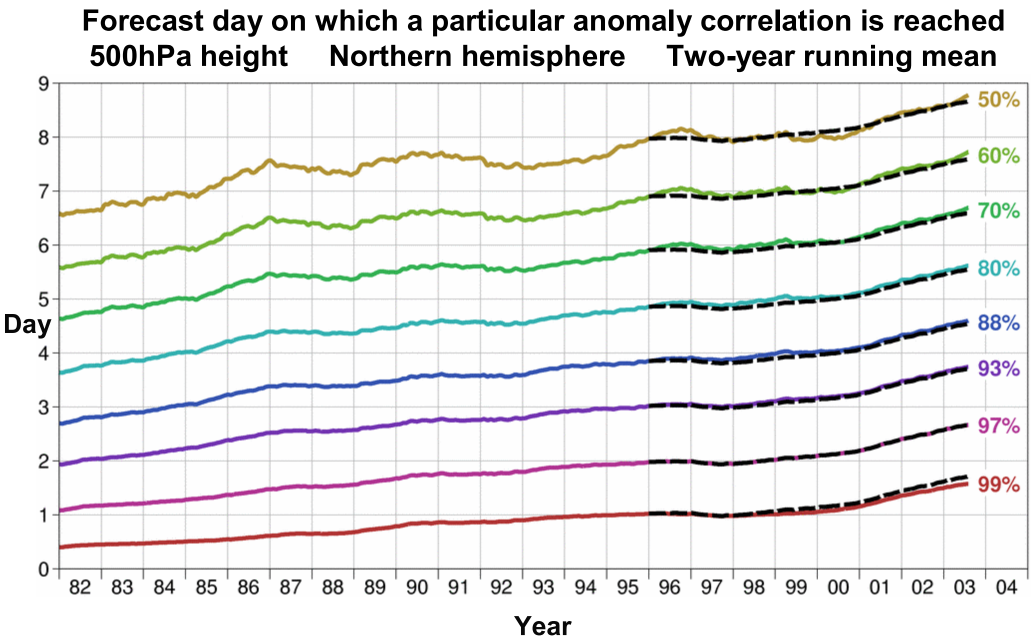

Figure 18: The forecast ranges at which the monthly-mean anomaly correlations of 500hPa height forecasts for the extratropical northern hemisphere reach a set of values from 99% to 50%, based on ECMWF operational forecasts from January 1980 to July 2003. The coloured curves denote actual values plotted as two-year running means, each point on the graph representing the average over the preceding two years. The black dashed curves show displacements of the curve for the 97% level from January 1996 onwards. The curves are displaced so as to match actual values for each percentage level for January 1996.

Fig. 17 can also be viewed as showing that the value of the difference in performance between the ECMWF and Met Office assimilation systems is roughly equivalent to the value of the latest 12-hours of observations, at least as far as the hemispheric accuracy of analyses and forecasts of 500hPa height is concerned. Forecasts from an analysis that uses the most recent observations are not necessarily more accurate than forecasts from an analysis for an earlier synoptic hour that has been produced by a better data assimilation system. In designing an operational forecasting system, compromises have to be made regarding the data cut-off (the wall-clock time at which the data assimilation is started), the times by which forecasts products have to be delivered, and the complexity (and hence computational cost) of the data assimilation and forecast model integration. Traditionally, the data cut-off for short-range prediction has been determined first and foremost by the need to receive the bulk of the latest radiosonde ascents, and a quite stringent limit is placed on the time that can be taken by the forecast-production process. The increasing relative importance of asynoptic observations, particular from satellites, and the development of more accurate but more computationally demanding data assimilation systems and forecast models are making appropriate choices less clear, especially for medium-range prediction. Whatever choices are made, analysis and forecast accuracy remains dependent on the amount of data available for assimilation, and effective telecommunication systems to ensure rapid transmission of information from the point of measurement (satellite or otherwise) to forecasting centres remain a key requirement for optimal operational forecasting.

Fig. 19 casts a different light on the performance of the 12UTC ECMWF and Met Office forecasts. It repeats the plots of root-mean-square errors of these 500hPa height forecasts, and adds plots of the root-mean-square differences between the 12UTC ECMWF and Met Office forecasts and between the 12UTC ECMWF forecasts and the 12-hour older ECMWF forecasts made from 00UTC. The main point of note is that the 12UTC ECMWF forecasts differ from the corresponding Met Office forecasts by a little more than they differ from the verifying analyses (consistent with the forecast error correlations of about 50% noted earlier). The ECMWF and Met Office forecasts do not predominantly share common errors that have resulted from each analysis system having drawn to the same inaccurate observations. Instead, differences in data analysis (quality control decisions, basic analysis methods and specified error statistics) and the growth of differences through the background forecasts of the data assimilation result in two analyses which, were it not for the overall superiority of the ECMWF system, would essentially be two randomly drawn and equally likely estimates of the true state of the atmosphere. As such the two analyses would provide two sets of initial conditions for an ideal ensemble prediction system. This provides the basis for the small "multi-analysis" ensemble of forecasts run daily by ECMWF from the analyses of other centres. Reproducing such variability among analyses is a challenge for ensemble data assimilation systems aimed at providing either optimal perturbations for ensemble prediction or flow-dependent background-error estimates for improved data assimilation, as discussed elsewhere in these Proceedings.

The 12UTC ECMWF forecasts can be seen in Fig. 19 to be much closer to the earlier 00UTC ECMWF forecasts than they are either to the verifying analyses or to the 12UTC Met Office forecasts. The observations that are assimilated over one 12-hour assimilation window are clearly insufficient to remove dependence of the 12UTC analysis on the background forecast initiated 12 hours earlier. Thus, whilst the 12 UTC Met Office 500hPa height forecast is on average similar in accuracy to the earlier 00UTC ECMWF forecast, it provides a more distinct possible alternative to the (probably more correct) 12UTC ECMWF forecast.

Lorenz (1982) discussed the comparison of root-mean-square forecast errors with root-mean-square differences between ECMWF forecasts started a day apart and verifying at the same time. He argued that if the forecast model in operational use at the time was realistic enough for small differences in initial conditions to cause forecasts to diverge at a rate close to that at which separate but similar atmospheric states diverge, then the rate of growth of the forecast differences would provide a limit to the potential accuracy of the forecast that could not be surpassed without analysis or model changes which reduced the one-day forecast error. The evolution of the forecast differences (or "perfect-model" errors) would in particular provide a basis for estimating the intrinsic rate of growth of initially small forecast errors.

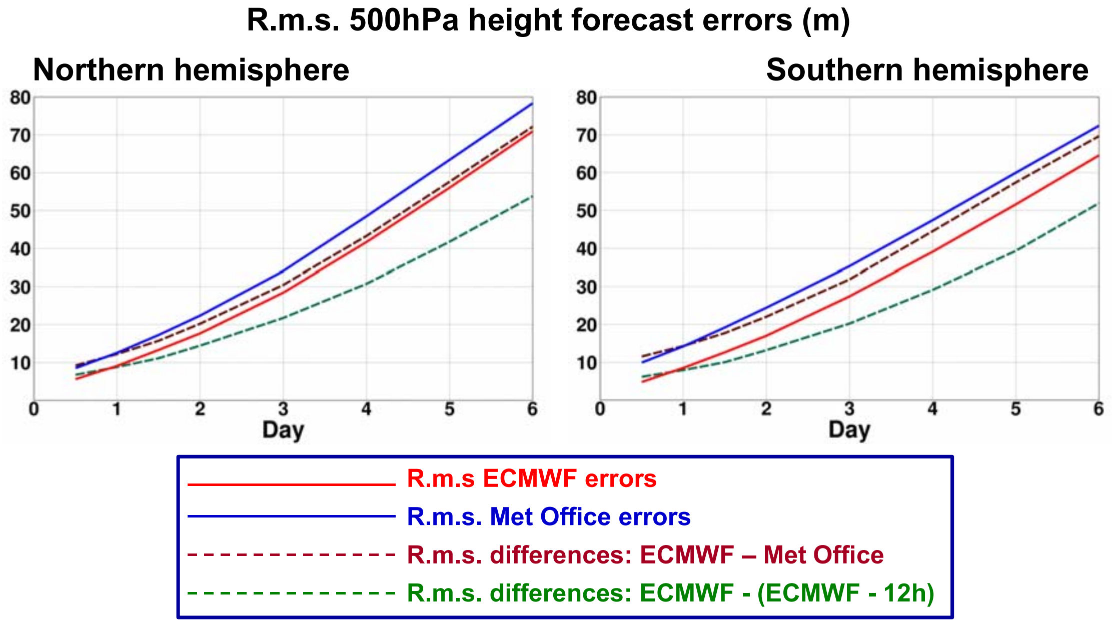

Figure 19: Root-mean-square errors of and differences between 500hPa height forecasts (m) for the extratropical northern (left) and southern (right) hemispheres verifying at 12UTC in the period from 1 December 2002 to 28 February 2003. Errors are shown for the operational 12UTC forecasts from ECMWF (red, solid) and the Met Office (blue, solid) at the forecast range indicated on the ordinate. Also shown are the differences between these ECMWF and Met Office forecasts (brown, dashed), and the differences between these ECMWF forecasts and ECMWF forecasts verifying at the same times but initiated 12 hours earlier(green, dashed).

Measures of the differences between successive numerical forecasts valid for the same time are indicators of forecasting-system performance that are also of direct relevance to bench forecasters. These measures provide indications of the consistency of the forecasting system, the extent to which the latest forecast is consistent with the forecast provided 12[1] or 24 hours earlier. High consistency (low values for measures of the differences between successive forecasts) is clearly a desirable feature if it stems from a basic high accuracy of the forecasts. For forecast of limited accuracy the situation is less clear. Inconsistent forecasts in such cases provide the forecaster with indications of a less predictable situation and of possible alternative developments, but if issued to end-users may unduly undermine confidence in the forecasts.

Fig. 20 compares root-mean-square errors and root-mean-square differences between successive daily operational forecasts of 500hPa height for the northern hemisphere for the winters of 1980/81 (the period studied by Lorenz, 1982) and 2002/03. The errors are very much lower in 2002/03, across the whole forecast range. The reduction since 1981 is about 70% at day one, 45% at day five and 25% at day ten. The gap between the forecast-error and forecast-difference curves is much smaller in 2002/03 than in 1980/81, indicative of model improvement since 1981. This is known to include a significant reduction in the systematic component of forecast error (Simmons *et al*. 1995; Ferranti *et al*. 2002).

1 Forecasts at 12-hourly frequency using comparable data cut-off times have been produced regularly by ECMWF only since 28 March 2001

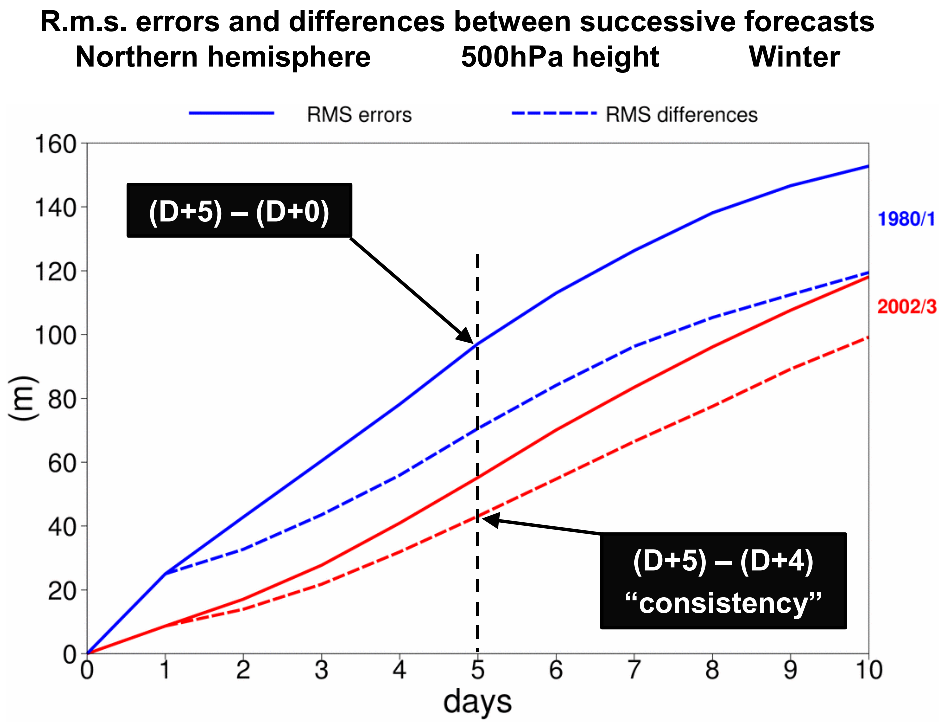

Figure 20: Root-mean-square forecast errors (solid) and root-mean-square differences between successive forecasts (dashed), based on 12UTC ECMWF 500hPa height(m) forecasts for the extratropical northern hemisphere verifying in the periods from 1 December 1980 to 28 February 1981 and from 1 December 2002 to 28 February 2003.

The model improvements over the past two decades have brought with them, as an unavoidable by-product, a faster rate of amplification of forecast errors, and a faster rate of growth of inconsistency between successive forecasts (Simmons and Hollingsworth, 2002). Lower absolute values for forecast errors and the differences between successive forecasts have occurred nevertheless, due to the very much lower starting point for the growth of errors and inconsistency that has been provided by the substantial reduction in analysis and short-range forecast error. Fig. 21 shows, for both extratropical hemispheres, the difference (or consistency) curves plotted for each DJF from 1980/81 to the present. Evident in this figure are the general reduction in one-day forecasts errors, with a distinct fall in particular between 1999/00 and 2000/01, and the much more rapid growth-rate of differences that occurs due to model changes in the late 1980s and early 1990s. Both are even more pronounced for the southern than the northern hemisphere. The increased consistency of forecasts in recent years stems mainly from the reduction in short-range forecast error, as doubling times for small forecast differences have not increased (Simmons and Hollingsworth, loc. cit.). Jung and Tompkins (2003) do however show that the forecasts over the past three years exhibit a slight loss of kinetic energy of transient eddies in the northern hemisphere compared with the early 1990s, although in the southern hemisphere they exhibit less of a spurious gain in energy.

Relatively slow growth of differences for 1999/2001 is due to a deficiency in the operational humidity analysis between October 1999 and April 2000. The lower stratosphere was much too moist over this period, leading through the radiation parameterization in the medium-range forecasts to a changed structure of potential vorticity near the tropopause and a consequent damping of tropospheric transient-eddy activity. What looked at first sight to be a systematic model error in fact was due to a bias in the analysis of a field that adjusted only slowly towards reality as the forecast proceeded.

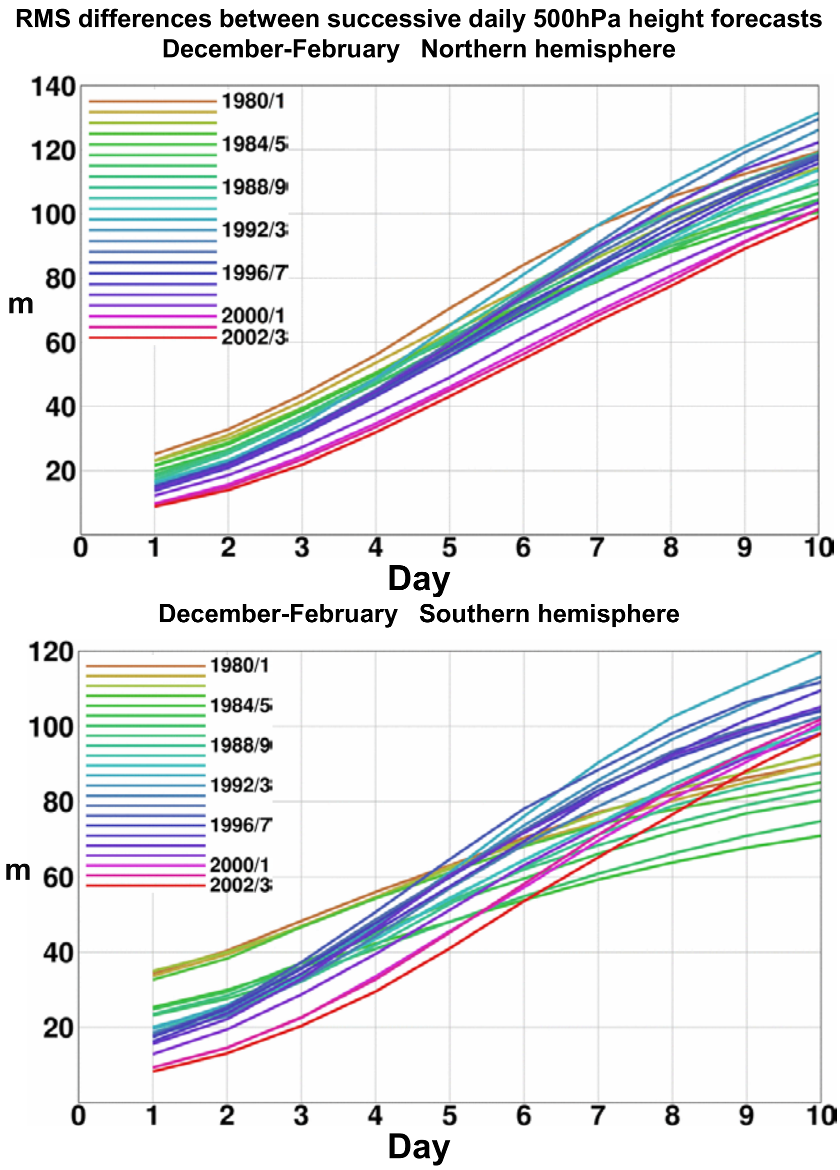

Figure 21: Root-mean-square differences between successive 12UTC forecasts of 500hPa height (m) as functions of forecast range. Results are shown for each December-February period from 1980/81 to 2002/03, for the extratropical northern (upper) and southern (lower) hemispheres.

6. Dependence of forecast error on horizontal scale

The dependence of forecast error on horizontal scale is of interest as regards predictability and forecasting-system performance in general, and more specifically as regards the potential for improved analysis of smaller scales.

Fig. 22 presents log-linear plots of spectra of the global error (measured by verification against analyses) of operational one-day ECMWF 500hPa height forecasts for each DJF since 1980/81. Results are shown for wavenumbers up to 40, the highest wavenumber for which operational forecast results have been saved for the early years. Spectra of the temporal variance of error rather than the mean-square error are presented to avoid the plots being complicated by the relatively large time-mean component of error that occurs at low wavenumbers in the earlier forecasts.

Fig. 22 shows that one-day 500hPa height forecast errors have been substantially reduced over the years at the large synoptic scales represented by a band of wavenumbers centred around about wavenumber nine. In the early years there was a quite pronounced spectral peak of error at these wavenumbers. Error is more uniformly distributed across the spectrum for more recent years, with the flattest spectrum for 2002/03. The sharp decrease in small-scale error with increasing wavenumber seen for the first few years presumably reflects the characteristics of the forecast (and assimilating) model of the time rather than a low level of actual small-scale error. Subsequent forecasting-system changes resulted in an increase in small-scale error, and only following the system changes made in 1999 and 2000 (increases in model and analysis resolution in November 2000 in particular) does the one-day error as measured against analyses drop across the whole spectral range for which results are presented.

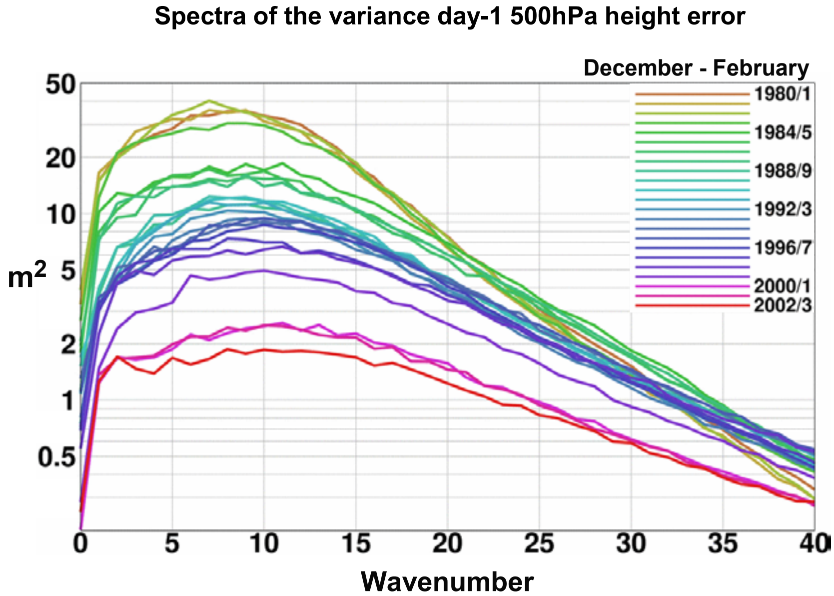

Figure 22: Log-linear plot of the spectra of the variance of the global 500hPa height forecast error (m^2) at one-day range for forecasts verifying in each December-February period from 1980/81 to 2002/03.

Spectra of mean-square errors for DJF 2002/03 are presented in Fig.23 for forecast ranges from twelve hours to ten days. Log-log plots are shown for wavenumbers up to the limit of the T511 model's truncation. Results are presented for 850hPa temperature as there is relatively more amplitude in small scales for this field than for 500hPa height.

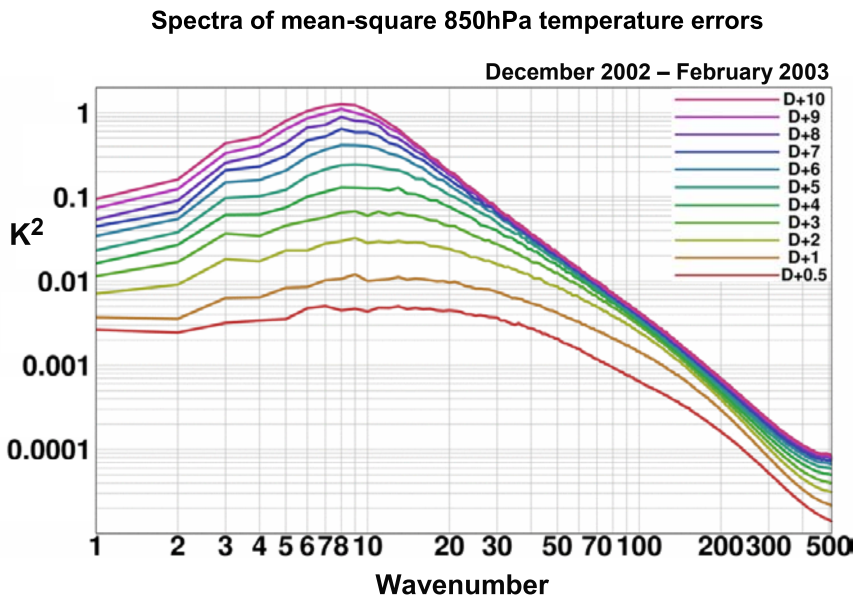

Figure 23: Log-log plot of the spectra of the mean-square global 850hPa temperature forecast error (K^2) at forecast ranges from twelve hours to ten days, for forecasts verifying at 12UTC in the period from 1 December 2002 to 28 February 2003.

The spectrum of 850hPa temperature error for 12-hour forecasts is rather flat over a quite wide range of wavenumbers. A more peaked error spectrum evolves with increasing forecast range, with error continuing to grow throughout the ten days near the spectral peak and at longer wavelengths. At smaller scales error largely saturates within the forecast range, but a slow component of error growth is evident near the truncation limit. Boer (1994) reported similar small-scale behaviour in an earlier version of the ECMWF forecasting system.

A clearer picture of the small-scale behaviour emerges when graphs are redrawn using linear-linear axes, as shown in Fig. 24 for wavenumbers 150 and higher. The figure shows more clearly that there are essentially two time scales for the growth of error in small spatial scales. There is rapid growth and apparent saturation early in the forecast range of a component of error that is likely to represent intrinsic dynamics with small temporal as well as spatial scales. This component decreases markedly in amplitude with decreasing scale, as would be expected given the model's high damping rates for the smallest scales. There is also a second component of error (with a less marked decrease in amplitude with decreasing scale) that evolves much more slowly throughout the forecast range. It is likely that this arises from small-scale components of the temperature field that are directly forced by large-scale dynamical and thermodynamical processes that have much slower intrinsic error growth rates.

Fig. 25 shows log/log plots equivalent to those of Fig.23, but for 850hPa relative vorticity rather than temperature. The vorticity field is inherently smaller scale. Short-range forecast error peaks much further down the spectral range than for temperature, and there is a much larger shift in the shape of the error spectrum as the forecast range increases. As for temperature, error can be seen to grow throughout the forecast range for wavenumbers ten and less, and there is also a component of small-scale error that grows throughout the range.

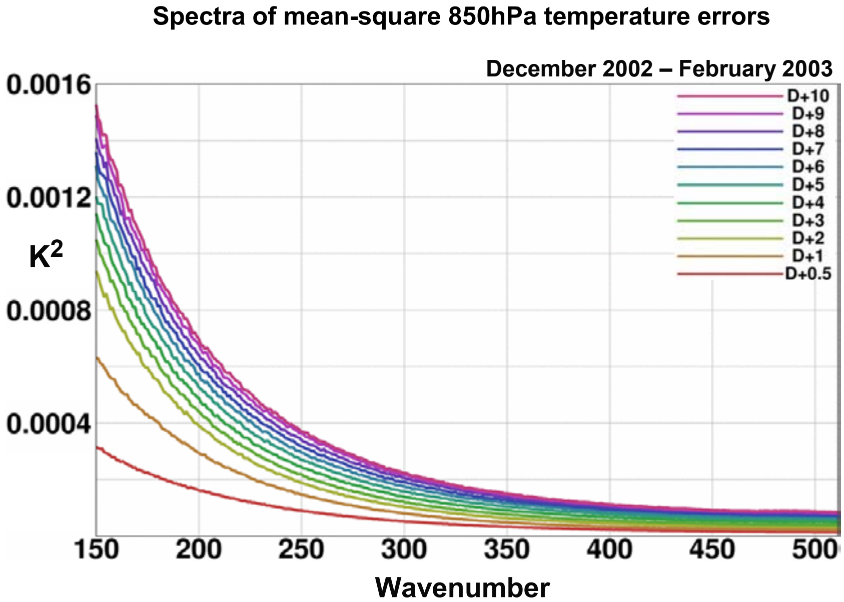

Figure 24: Spectra of 850hPa temperature errors as in Fig. 22, but shown in a linear-linear plot for wavenumbers 150 and above.

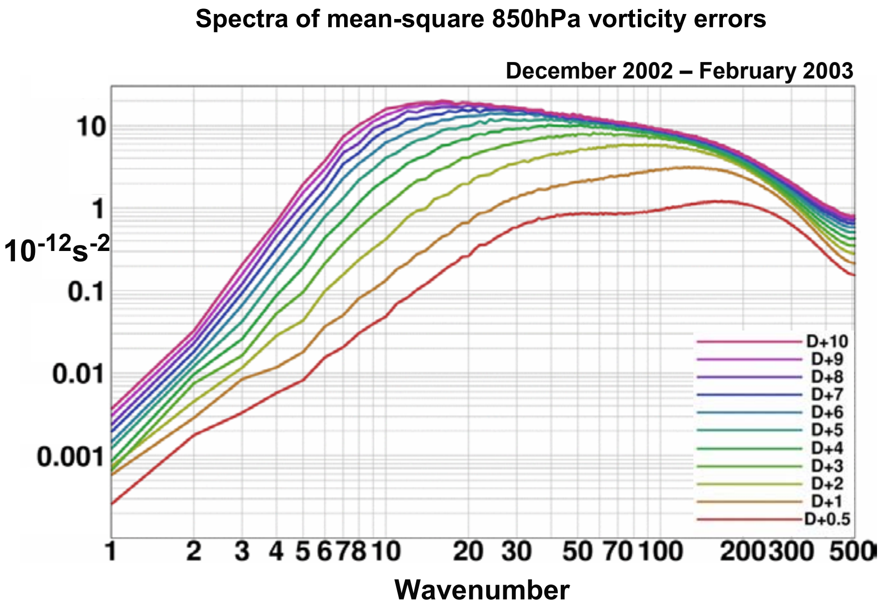

Figure 25: Log-log plot of the spectra of the mean-square global 850hPa relative-vorticity forecast error (s^{-2}) at forecast ranges from twelve hours to ten days, for forecasts verifying at 12UTC in the period from 1 December 2002 to 28 February 2003.

7. Concluding remarks

The first aim of this presentation has been to give a general introduction to the atmospheric observing system and to the way that data from this system are assimilated into numerical models to provide the starting point for numerical weather forecasts. The second aim has been to illustrate the improvement in synoptic-scale weather prediction that has resulted from improvements made to the observing system and data assimilation. To do this use has been made of some of the most recent results from the ERA-40 reanalysis project and from operational forecasting.

The results from ERA-40 show that there has been a clear long-term improvement in the observing system. This is especially so for the southern hemisphere, for which ERA-40 provides striking evidence of the substantial benefit of the changes made to the observing system twenty-five years ago in the build-up to FGGE. ERA-40 also shows nevertheless that a modern data assimilation system (despite being run in a degraded configuration for reasons of computational cost) is capable of providing analyses of the state of the northern hemisphere more than forty years ago that are of sufficient accuracy to yield forecasts that remain skilful well into the medium range. Lack of today's techniques and computer power appear to have been more of a hindrance to general northern hemisphere forecast quality forty years ago than lack of today's observing system, though it must be kept in mind that this remark applies to routine synoptic-scale prediction beyond a day or two ahead, and not necessarily to the short-range prediction of severe weather events where satellite imagery (and observing systems such as ground-based radar) can play such an important role. It should also be kept in mind that the newer observing systems have had to compensate for a decline in radiosonde coverage over key oceanic and high-latitude regions.

There has been a very substantial improvement in operational forecasts over the seven years since the ECMWF Seminar was last devoted to data assimilation. Evidence discussed here and elsewhere indicates that improvements have stemmed in particular from improved data assimilation (improved assimilating models as well as improved analysis techniques) and from new or improved types of observation. It should be noted that attention has been concentrated here on the accuracy of forecasts of 500hPa height, and to a lesser extent wind and temperature in the free troposphere, rather than on forecasts of weather elements such as near-surface temperature, cloud and precipitation. The latter benefit directly from model improvements as well as from the improved definition of the synoptic environment that is documented by the type of verification presented here. General forecast improvements have been achieved by many centres, and it has been demonstrated how the initial analyses of two centres, ECMWF and the Met Office, have converged substantially. Nevertheless, significant (and informative) differences in analyses and forecasts remain.

The spectral breakdown of error discussed in the preceding section shows that there has been a distinct recent improvement in the handling of smaller scales of motion. In the operational ECMWF incremental 4D-Var data assimilation, the spectral truncation of the highest-resolution minimization is at wavenumber 159 and the assimilation period is twelve hours. The results presented show that forecast error is still some way from saturation after twelve or even twenty-four hours for a range of wavenumbers higher than 159, suggesting that there may well be scope for extending the resolution of the minimisation. This might be a rather optimistic view given that the small-scale short-range "forecast errors" considered here are determined by verifying against an analysis that at small scales is very heavily influenced by the model background forecast. A clear benefit accrued however from the resolution changes made in November 2000. The potential for further benefit from additional increases in resolution thus seems well worth exploring.

Acknowledgements

Many have contributed to the achievements reported here. Particular thanks are given to Sakari Uppala and his colleagues in the ERA-40 project team for their successful conclusion of the reanalysis earlier this year and for supplying several of the figures presented here. Tony Hollingsworth is thanked for earlier collaboration on some aspects of the work presented and specifically for drawing attention to the result illustrated in Fig. 18.

References

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljevic, D., Brancovic, C., Cardinali, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Thépaut, J.-N. and Viterbo, P. 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part III: Experimental results. *Quart. J. Roy. Meteor. Soc.,* **124**, 1831-1860.

Bengtsson, L. and Simmons, A.J. 1983: Medium-range weather prediction - Operational experience at ECMWF. *Large-scale Dynamical Processes in the Atmosphere*, Eds. B.J. Hoskins and R.P. Pearce, Academic Press, 337-363.

Boer, G.J. 1994: Predictability regimes in atmospheric flow. *Mon. Wea. Rev.*, **122**, 2285-2295.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. R. Meteorol. Soc.*, **120**, 1367-1388.

Derber, J.C. and Bouttier, F. 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195-222.

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

Hólm, E., Andersson, E., Beljaars, A., Lopez, P., Mahfouf, J.-F., Simmons, A.J. and Thépaut, J.-N. 2002: Assimilation and modelling of the hydrological cycle: ECMWF's status and plans. *ECMWF Tech. Memo.*, **383**, 55pp.

Jung, T. and Tompkins, A. 2003: Systematic errors in the ECMWF forecasting system. *ECMWF Tech. Memo.*, **422**, 72pp.

Lorenc, A.C., Ballard, S.P., Bell, R.S., Ingleby, N.B., Andrews, P.L.F., Barker, D.M., Bray, J.R., Clayton, A.M., Dalby, T., Li, D., Payne, T.J. and Saunders, F.W. 2000: The Met. Office global three-dimensional variational data assimilation scheme. *Quart. J. R. Meteorol. Soc.*, **126**, 2991- 3012.

Lorenz, E.N. 1982 Atmospheric predictability experiments with a large numerical model. *Tellus*, **34**, 505-513.

Mahfouf, J.-F. and Rabier, F. 2000: The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics. *Quart. J. R. Meteorol. Soc.*, **126**, 1171-1190.

Parrish, D.F. and Derber, J.C. 1992: The National Meteorological Center's Spectral Statistical- Interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Simmons, A.J., Mureau, R. and Petroliagis, T. 1995: Error growth and predictability estimates for the ECMWF forecasting system. *Quart. J. R. Meteorol. Soc.*, **121**, 1739-1771.

Simmons, A.J., Andersson, E., Fisher, M., Jakob, C., Kelly, G.A., Lalaurette, F., McNally, A.P., Untch, A. and Viterbo, P. 2001 Forecasting system performance in summer 1999: Part 2 - Impact of changes to the ECMWF forecasting system. *ECMWF Tech. Memo.*, **322**, 31pp.

Simmons, A.J. and A. Hollingsworth 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **128**, 647-677.

Simmons, A., Hortal, M., Kelly, G., McNally, A., Untch, A., and Uppala, S. 2003: Analyses and forecasts of stratospheric winter polar vortex break-up: September 2002 in the southern hemisphere and related events from ECMWF operations and ERA-40. *ERA-40 Project Report,* **5**, 28pp.