Development of the ERA-40 Data Assimilation System

A.J. Simmons

European Centre for Medium-Range Weather Forecasts
Shinfield Park
Reading, RG2 9AX, U.K.

Abstract

The development of the ERA-40 data assimilation system is reviewed. Changes made to the operational ECMWF assimilation system since production of the ERA-15 analyses are summarized, as are additional changes developed specifically for ERA-40. Examples are presented of the assessment of ERA-40 through the examination of trends and through short-range forecast verification.

1. Introduction

Ideally, reanalysis is the analysis of past observational data using a fixed, tried-and-tested, data assimilation system. In practice, however, in developing and running a reanalysis system one has to address a number of issues and make a number of compromises.

Firstly, for a data assimilation that covers a period of several decades and has to be completed within a production period of about two years, the computational cost of completing a single analysis will typically have to be substantially less than for an advanced current operational data assimilation system. Choosing an earlier, affordable operational version of the data assimilation system may mean foregoing a number of desirable recent developments. Accordingly a new configuration that incorporates as many of the recent operational improvements as possible may have to be developed.

Secondly, the best available data assimilation system is likely to be based on the use of data from current observing systems, and thus require adaptation to use data from older observing systems. Moreover, when the reanalysis system is run in production mode, its performance must be monitored carefully to identify when intervention is required to adapt to changes in any pre-processing applied by data producers.

Thirdly, a system that is tried-and-tested inevitably has a number of known deficiencies. Whilst recently developed changes may be implemented to address some of these deficiencies, acceptable and reasonably tested solutions to others may not be available. Account needs to be taken of this in the generation of products, the validation and the documentation of the reanalysis.

Finally, errors and addressable deficiencies in formulation may well be found during the course of production of a reanalysis. Decisions have to be made as to whether to introduce corrections or refined formulations in the interest of improving later analyses, or to let the problems remain in the system in the interest of continuity of the full set of analyses.

Each of these types of issue has been faced in the development and production runs of the ERA-40 data assimilation system, as discussed in several of the presentations at the workshop. In this paper an overview is given of the development of the data assimilation system for ERA-40. More detailed discussion of a number of aspects of the system, particularly relating to the use of satellite data, has been given in other presentations by ECMWF speakers.

The computational cost of ECMWF's current operational four-dimensional variational data assimilation (4D-Var) and T511 horizontal resolution is too high for it to be used for ERA-40. A modified form of the three-dimensional variational analysis (3D-Var) used operationally in the ECMWF medium-range prediction system between January 1996 and November 1997 (Andersson et al., 1998) has thus been adopted. 3D-Var is today used operationally at ECMWF to produce short cut-off analyses from which forecasts are run to provide boundary conditions for the limited-area short-range forecasting systems of a number of the ECMWF Member States. It is also the method used for operational global analysis at a number of other centres (e.g. Parrish and Derber 1992; Lorenc *et al*. 2000) and was used for the NCEP/NCAR reanalysis (Kalnay et al., 1996).

The spectral representation of atmospheric fields chosen for ERA-40 has triangular truncation of the spherical harmonic expansion at total wavenumber 159. This T159 resolution is finer than the T62 resolution used for the NCEP/NCAR reanalysis, and the T106 resolution used for ECMWF's fifteen-year analysis for 1979-1993 (ERA-15; Gibson et al., 1997). Other fields are represented on a form of reduced Gaussian grid (Hortal and Simmons, 1991; Courtier and Naughton, 1994) with a quasi-uniform grid spacing of about 125km, the same as used for ERA-15. The 60-level vertical resolution (discussed later) is that currently used operationally at ECMWF.

It is beyond the scope of this paper to give a comprehensive documentation of the analysis scheme and assimilating model used for ERA-40. They are elements of the "Integrated Forecasting System" (IFS) developed jointly by ECMWF and Météo-France. Much relevant information can be found in the IFS documentation that can be viewed on ECMWF's public website (http://www.ecmwf.int), and further information (including specification of the products archived from ERA-40 and details of their spatial representation) can be found in the site's ERA-40 pages. Here brief summaries are presented of relevant changes to the ECMWF's standard assimilation system made since the ERA-15 analyses were produced and of changes to the assimilation system made specifically for ERA-40. Examples of the monitoring of ERA-40 production through examination of trends and verification of short-range forecasts are also presented.

2. Operational changes year-by-year since ERA-15

The data assimilation system used for ERA-15 was cycle 13r4 of the IFS, and was essentially the same as that used operationally by ECMWF from April 1995 to January 1996, apart from its use of T106 rather than T213 horizontal resolution. Since then there have been between one and four changes each year to the operational forecasting system that are of relevance to ERA-40. ERA-40 is based on cycle 23r4 of the IFS, though with some modifications. The changes are summarized below on a year-by-year basis. Simmons and Hollingsworth (2002) discuss the substantial improvement in overall medium-range forecast accuracy that has resulted from them.

2.1 1996

January, IFS cycle 14r3

This change saw the replacement of Optimal Interpolation (used for ERA-15) by 3D-Var (Andersson et al., 1998) for the main analysis. Introduction of 3D-Var enabled assimilation of scatterometer data, initially from the ERS-1 satellite, with data from ERS-2 replacing ERS-1 data in July. The January change also included revisions to the parametrizations of vertical diffusion and convection.

September, IFS cycle 15r5

3D-Var was extended in September 1996 in the areas of observation screening (Järvinen and Undén, 1997), quality control (Andersson and Järvinen, 1999) and calculation of background-error variances (Fisher and Courtier, 1995). In addition, modifications were made to the land-surface parametrization (introducing the freezing of soil moisture) and to the boundary-layer parametrization in stable conditions (Viterbo et al., 1999). Both changes substantially reduced the wintertime cold bias in near-surface temperatures that was a deficiency of the ERA-15 analyses. An error in the treatment of detrainment in the convection parametrization was also corrected (Simmons et al., 1999).

December, IFS cycle 15r7

A two-time-level version (Temperton et al., 2001) of the model's semi-Lagrangian advection scheme was introduced as part of this change, and advection was applied for the first time to the prognostic cloud fields. Prior to the earlier introduction of 3D-Var, specific humidity in the stratosphere had been set at each analysis cycle to the value 2.5×10^{-6}. Problems were experienced when humidity was subsequently allowed to evolve through 3D-Var data assimilation cycles, due to a spreading upwards of analysis increments of tropospheric origin, and increments in specific humidity were thus suppressed for low background values (Simmons et al., 1999).

Large differences in two-metre temperature between ERA-40 and ERA-15

The change in low-level temperature from ERA-15 to ERA-40 is especially pronounced for the Antarctic. In addition to the effect of the model changes introduced into operations in September 1996, there are differences in surface and low-level atmospheric temperatures in this region due to more-accurate specification of the orography, a changed modelling of sea-ice and treatment of the main permanent ice shelves as land rather than sea-ice. Fig. 1 shows differences in two-metre temperature between monthly-mean analyses for July 1989 from ERA-40 and ERA-15. The ERA-40 analysis is substantially warmer over the Antarctic plateau, by up to 22K, generally cooler by a few degrees over sea-ice, and substantially colder, by up to 15K, over the Ronne and Ross ice shelves.

2.2 1997

May, IFS cycle 16r2

The major change in this cycle was the introduction of a new formulation of the background forecast error constraint J_b (Derber and Bouttier, 1999). In addition, the variational quality control was extended to the assimilation of TOVS radiance data.

August, IFS cycle 16r4

Among the changes introduced at this time were a new bias correction of TOVS radiances (Harris and Kelly, 2001), more accurate calculation of saturation vapour pressure for cold temperatures (Simmons et al., 1999), refinement of the quality control, and improvements to the use of scatterometer data and representation of ocean albedo.

November, IFS cycle 18r1

This change comprised the operational introduction of 4D-Var (Mahfouf and Rabier 2000, and refs.) with six-hourly cycling.

December, IFS cycle 18r3

Major modifications to the parametrizations of deep convection, radiation and clouds were implemented this month (Gregory et al., 2000).

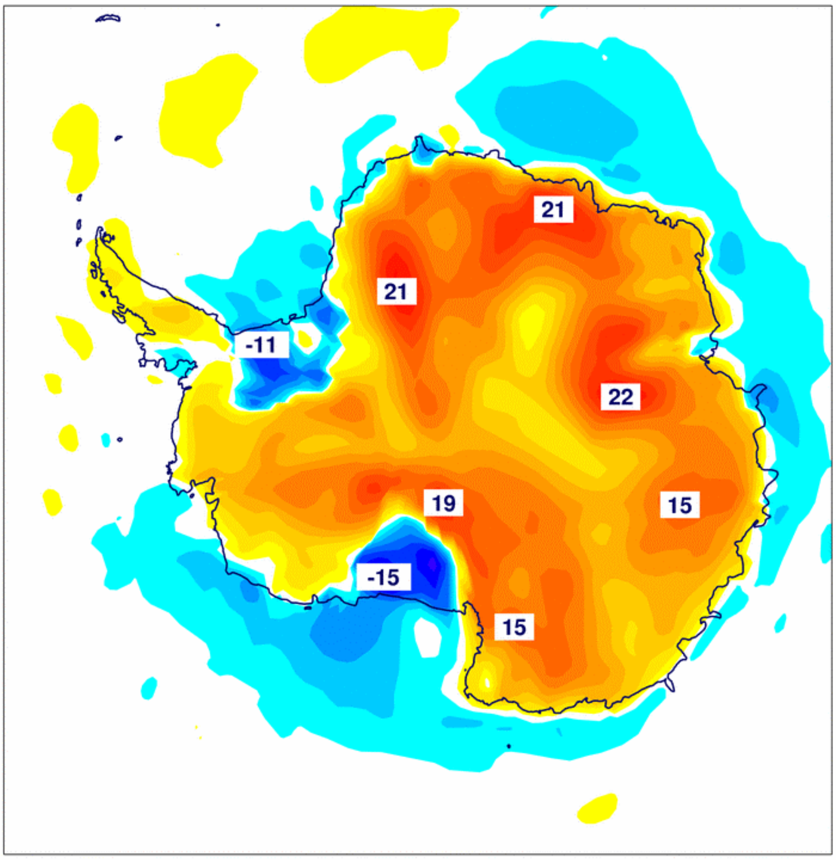

Fig. 1 Differences over the Antarctic in two-metre temperature (K) between ERA-40 and ERA-15 analyses averaged for July 1989. Colour shading begins at values of ±1K, with contour intervals of 2K. Red/yellow shading indicates where ERA-40 analyses are on average warmer than ERA-15 analyses.

2.3 1998

April, IFS cycle 18r5

The spectral resolution of the operational model was increased to T319 when this cycle was introduced. The "linear grid" option made possible by use of semi-Lagrangian advection was adopted, the computational grid remaining that of the previous T213 model version. The higher-resolution model orography was based on a new basic dataset that in particular gave an improved representation of Antarctica (noted earlier). A new version of the two-time-level semi-Lagrangian scheme was also introduced (Hortal, 2002), as was a seasonal variation of background (snow-free) albedo.

June, IFS cycle 18r6

Two-way coupling of the atmospheric model with the ocean-wave model was an important component of this change (Janssen et al. 2002). The use of radiosonde data was revised, assimilating temperature rather than height data and assimilating significant- as well as standard-level data. The assimilation was also changed to use hourly surface data (Järvinen *et al.* 1999) and SSM/I humidity data (Gérard and Saunders 1999).

Horizontal resolution of the ERA-40 assimilation system

The assimilating model for ERA-40 uses the linear-grid option adopted operationally in April 1998. As noted earlier, the spectral resolution is T159, compared with the T106 resolution of ERA-15. The ~125 km resolution reduced Gaussian grid is however the same as that used for ERA-15.

One consequence of ERA-40's use of finer spectral resolution, but the same grid-point resolution, is a very marked reduction in the amplitude of the unrealistic spectral ripples in the model orography that are most obvious over oceans or flat land close to major mountain ranges. Fig.2 compares the ERA-40 and ERA-15 orographies in the vicinity of South America. A logarithmic contour is chosen to emphasize the reduced amplitude of spectral ripples for ERA-40, which can be seen not only over the Pacific Ocean to the west of the Andes but also over the Amazon basin.

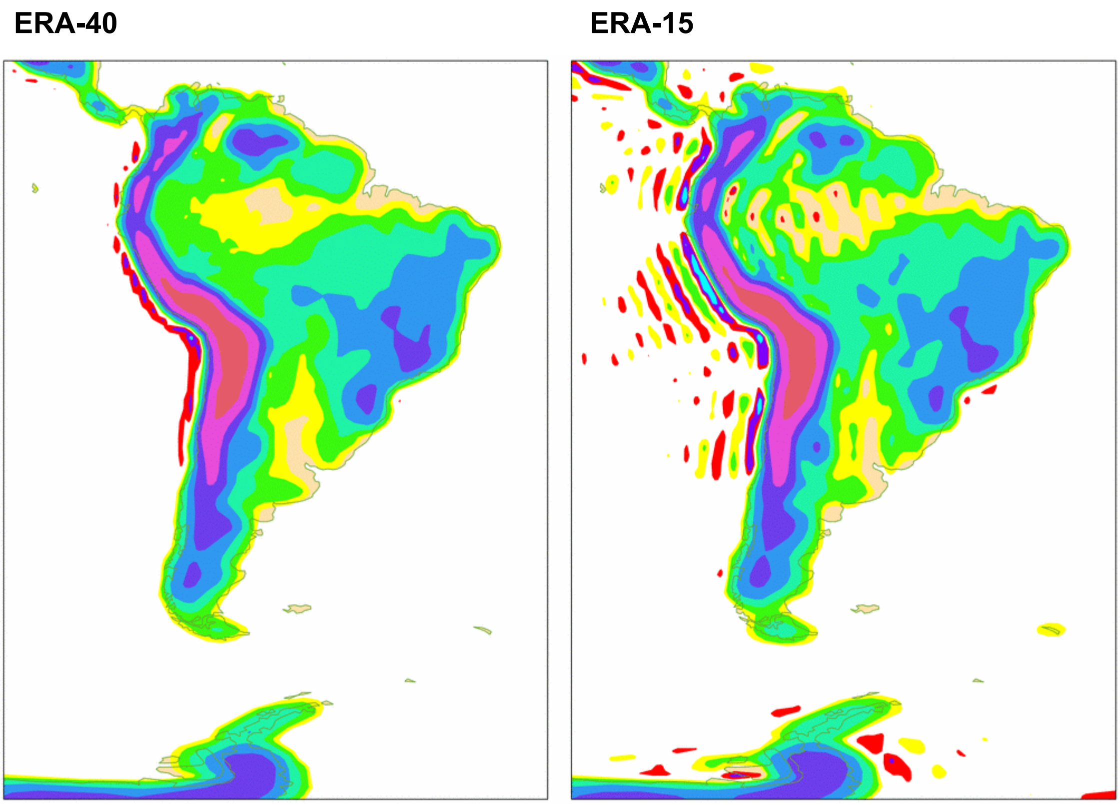

Fig. 2 ERA-40 (left) and ERA-15 (right) model orographies, plotted with contour values ±50, ±100, ±200, 400, 800, 1600 and 3200m.

2.4 1999

March, IFS cycle 19r2

This change (Untch and Simmons, 1999) saw an increase in the number of levels from 31 to 50 in the atmospheric model. Resolution was added in the stratosphere and lower mesosphere, with the top model level raised from 10 to 0.1hPa. A parametrization of the high-level moisture source due to methane oxidation was added.

May, IFS cycle 21r1

In this month assimilation of raw microwave (MSU, AMSU-A) radiances from the TOVS and new ATOVS satellite-borne instruments was introduced to replace the assimilation of pre-processed TOVS radiances (McNally et al., 1999).

July, IFS cycle 21r2

A new coupling of the model physics and dynamics (Wedi, 1999) formed part of this change. A new analysis of soil variables and of two-metre temperature and relative humidity was also activated. Quality control of the raw TOVS/ATOVS radiances was relaxed to allow assimilation of more of these data and there was an increase also in the use of SATOB winds from Meteosat (Rohn et al., 2001). Ship winds were applied at the anemometer height where known and otherwise at a more representative height than the 10m assumed previously. The assimilation was extended to include data from US wind profilers, and bias correction of radiosonde temperatures was introduced (the previous height-bias correction having ceased to be applicable in June 1998).

October, IFS cycle 21r4

The forecast-error estimates used in the specification of the background constraint J_b of the variational analysis were derived from an ensemble of 3D-Var analyses with perturbed observations and model physics from this cycle onwards. The assimilation was also improved by use of SSM/I marine-wind data, by correcting an error in the use of use of radiosonde humidity data and by improved bias correction of the TOVS/ATOVS radiances. Vertical resolution was increased to 60 levels, with the change on this occasion concentrated in the planetary boundary layer, where the height of the lowest level was lowered from 33m to 10m (Teixeira 1999). The representations of clouds (Jakob and Klein 2000), convection and orography were revised. A report on this major set of modifications is given by Jakob et al.(2000). In addition, ozone was introduced as a new prognostic model variable, although with neither assimilation of observations nor interaction with the radiation parametrization.

Vertical resolution of the ERA-40 assimilation system

The 60-level vertical resolution used for ERA-40 is that operational since October 1999. The distribution of these 60 levels is compared with that of the 31 levels used in the ERA-15 model in Fig. 3. The greater number of levels in the planetary boundary layer can be seen, though not their precise location (for that see Jakob et al., 2000). Upper tropospheric resolution is similar for the two configurations. The improved stratospheric representation for ERA-40 is evident; the vertical layer spacing is close to 1.5km over much of the stratosphere in the 60-level model.

Differences in oceanic 10-metre wind speed between ERA-40 and ERA-15

Kållberg(1997) discussed the fits to observation and spin-up of oceanic 10m winds in ERA-15. He inferred that the analysed winds were too low in places due to the assimilation of 10-metre wind observations from island stations that were unrepresentative of winds over neighbouring open seas, and that winds were likely to be too high elsewhere because of the assumption that ship wind observations were at 10m rather than at the generally higher levels at which the anemometers were located. Both analysis deficiencies have been corrected for ERA-40, by blacklisting unrepresentative island winds and adopting the changed use of ship winds introduced operationally in July 1999[1]. Fig.4 presents an example of differences in wind speed between ERA-40 and ERA-15.

[1] In the absence of information on the actual anemometer height, ship winds are assigned to a fixed height of 25m derived from averaging a recent WMO list of reporting ships. No allowance is made for a tendency for anemometer heights to increase over the period of ERA-40. This could cause a slight spurious drift in the ocean winds over the course of the reanalysis

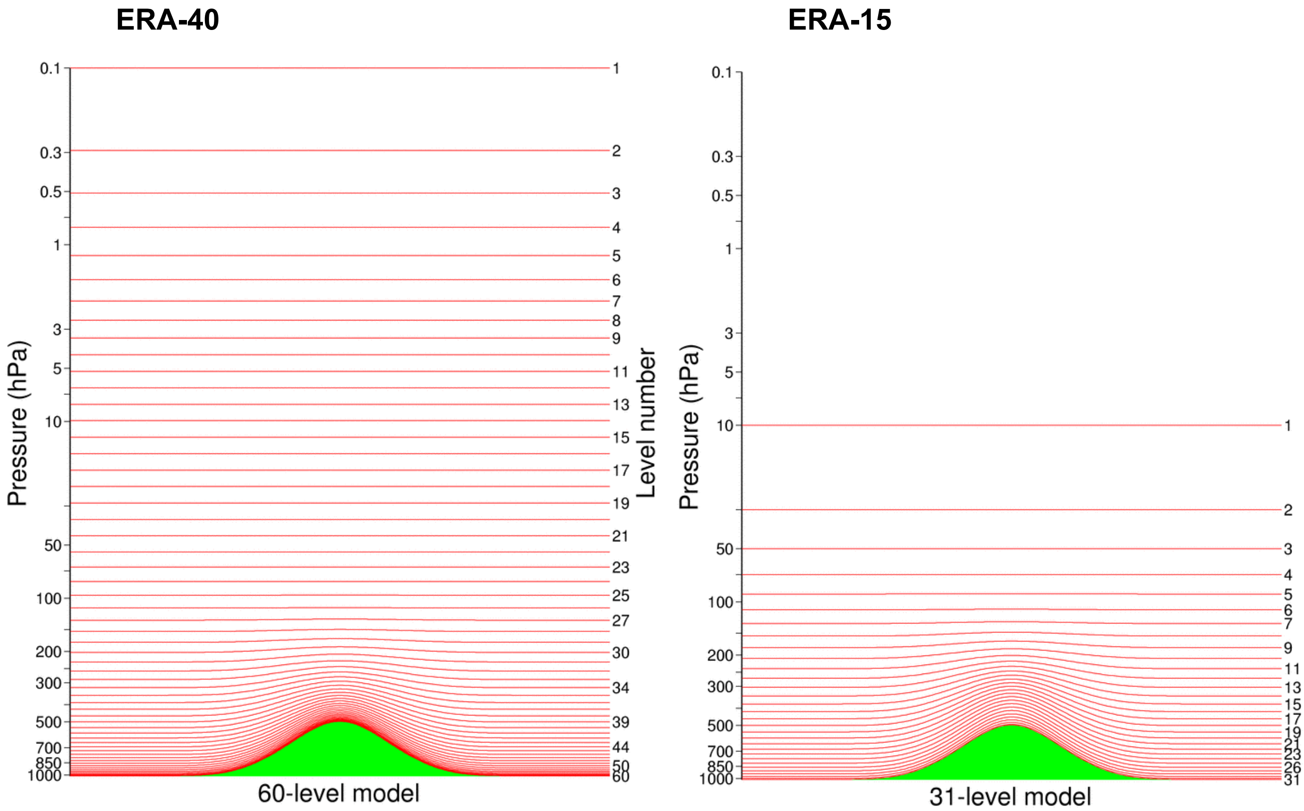

Fig. 3 The distribution of full model levels at which the basic prognostic variables are represented, for ERA-40 (left) and ERA-15 (right).

2.5 2000

April, IFS cycle 22r1

Use of SSM/I data was improved by introduction of a rain check and a revised (background-dependent) bias correction (Marécal et al., 2001). This enabled data from two satellites rather than one to be assimilated, although for much of ERA-40 data are available from only satellite. Suppression of humidity analysis increments was imposed above a diagnosed model tropopause, rather than simply where background humidity was low. Improvements were made to the ocean wave model and the assimilation of altimeter data used to provide it with initial wave heights.

June, IFS cycle 22r3

The observation- and background-error variances used in 4D-Var were revised for this cycle. The quality control of TOVS/ATOVS data was further relaxed, and data from additional channels were assimilated. Major model changes were made to the parametrizations of the land-surface and sea-ice (van den Hurk et al., 2000) and radiation (Morcrette et al. 2001). In addition, the post-processing of model fields was expanded to produce output directly on surfaces of constant potential temperature and potential vorticity, and cloud-affected TOVS/ATOVS radiances were processed passively through the assimilation system to enable production of a range of new diagnostic information on clouds (Chevallier et al., 2001).

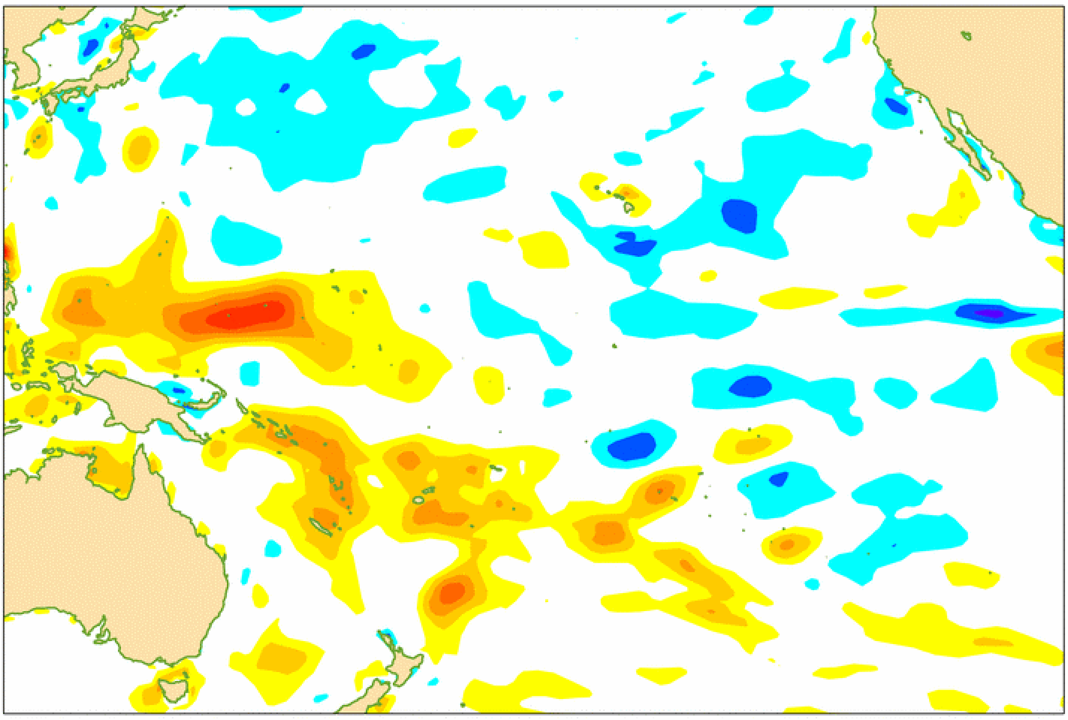

Fig. 4 Difference in 10-metre wind speed over the Pacific Ocean between ERA-40 and ERA-15 analyses, averaged over January 1989. Yellow/red colouring indicates faster winds in ERA-40 and blue colouring indicates slower winds. The threshold for shading is a difference of 0.5ms^{-1} and the interval between shading bands is also 0.5ms^{-1}.

September, IFS cycle 23r1

The main component of this change was an increase from six to twelve hours in the length of the time window used in 4D-Var (Bouttier, 2001). The change also included several related refinements to the incremental analysis system.

November, IFS cycle 23r3

This was a major change to the operational system comprising an increase of resolution of the assimilating (and deterministic forecasting) model from T319 to T511, and an increase from T63 to T159 in the inner-loop resolution of 4D-Var (Miller, 1999; Miller and Hortal, 2001). An increase in spectral resolution and other improvements to the ocean-wave model were also made. Benefit was also derived from the assimilation of data from a second ATOVS satellite.

2.6 2001

June, IFS cycle 23r4

A parametric model of surface emissivity over land and sea-ice was introduced to enable more lower-sounding AMSU (and in the ERA-40 context MSU) data to be used over these surfaces. Corrections were made to the representation and use of the model's land-skin temperature. The parametrization of ozone depletion by heterogeneous chemistry developed by Météo-France (as presented by Simon at the workshop) was adopted.

2.7 2002

Although based on cycle 23r4, several changes that became operational in later cycles were included in the ERA-40 data assimilation system. These included three changes that were introduced operationally in January 2002 in cycle 24r3: a finite-element scheme for the model's vertical discretization, removal of an error in the SSM/I bias correction (which affected stream-1 ERA-40 production up to January 1993) and a pre-conditioning of the 3D/4D-Var minimization (that has negligible impact on results but which significantly increases production efficiency). The assimilation of ozone observations introduced into operations in April 2002 as part of cycle 25r1 was based on that used in ERA-40, with assimilation of near-real-time total ozone retrievals from GOME replacing ERA-40's use of TOMS total ozone as the latter retrievals are not available in time for operational use. SBUV data are used in both ERA-15 and operations.

3. Additional developments and features for ERA-40

Several developments or parameter adjustments within the assimilation system were made specifically for ERA-40.

A FGAT (first-guess at the appropriate time) capability for the 3D-Var analysis was developed for use in ERA-40 in response to feedback received on the original plans for the reanalysis. In this approach, background (first-guess) values are compared with observed values at the observation time rather than the analysis time, and the differences are applied at analysis time. The 3D-Var configuration was also adapted to use the linear grid employed in the model, allowing analysis increments to be derived at the full T159 resolution of the assimilating model (the same resolution at which they are currently derived operationally in 4D-Var). These developments have enabled the reanalysis to fit data more accurately both in space and time.

It was decided to apply a simple uniform rescaling to increase the background error variances used in the variational analysis for ERA-40. This was based both on synoptic examination of trial analyses which indicated that the system was not drawing as closely to observations as was desirable and on the argument that the error variances by default were those appropriate to the background fields of the operational higher resolution 4D-Var assimilation. In view of the lower accuracy expected for ERA-40 background fields, a rescaling was considered appropriate. In addition, the time-step of the assimilating T159 model was reduced to 30 minutes from its default setting of one hour, due primarily to deficiencies in the analysis of the atmospheric tides. Both adjustments were shown to be beneficial to general analysis and forecast quality in pre-production tests for ERA-40. A 30-minute time-step is also used for the T159 inner-loop integrations of the operational 4D-Var assimilation.

The raw-radiance assimilation developed and refined for ECMWF's current operational data assimilation places heavy reliance on the use of data from the AMSU-A instrument available only since the summer of 1998, and the only other TOVS/ATOVS radiances assimilated operationally since its introduction have been from MSU and channel 12 of HIRS. The operational assimilation in particular relies on the uppermost AMSU-A channels for the analysis of the upper stratosphere. For ERA-40, the raw -radiance assimilation has had to be extended in several ways. Data from SSU are utilized for analysis of the pre-AMSU stratosphere. Data from more HIRS channels are utilized to compensate for the lower resolution of MSU compared with AMSU and for the absence of humidity data from SSM/I in the first part of the TOVS era. Data from the earlier VTPR instruments are utilized from 1973 until TOVS data become available in late 1978. Bias correction of these satellite data has been an ongoing challenge (as indeed has bias correction of radiosonde data). Discussion of these developments has been given in other talks at the workshop.

IPCC trends for the specified, radiatively-active gases have been introduced for ERA-40 as recommended by a Workshop on Boundary Conditions and Atmospheric Composition held in July 1998. Aerosol variability and a trend in the stratospheric water vapour source due to methane oxidation (see later) have not been included.

Enhanced sets of post-processed products have also been developed for ERA-40, in collaboration with members of the user community. Examples are vertically-integrated fluxes, fields from the physical parametrizations for support of chemical-transport modelling, and special grid-point and catchment-basin diagnostics. They are specified in the Archive Plan included in the project's web pages.

4. Some aspects of trends

One aspect of the monitoring of ERA-40 production has been the examination of fields for which a long-term trend might occur, either real or spurious. Three examples are presented here.

Near-tropopause temperatures

A cooling trend in temperatures near the tropical tropopause and in the lower stratosphere is a feature of the instrumental record and of analyses, both from reanalysis projects and from operations, at least for the satellite era. Discussion in the context of the ECMWF system has been given by Simmons et al.(1999). The ERA-40 analyses have thus been monitored in this respect as production has proceeded.

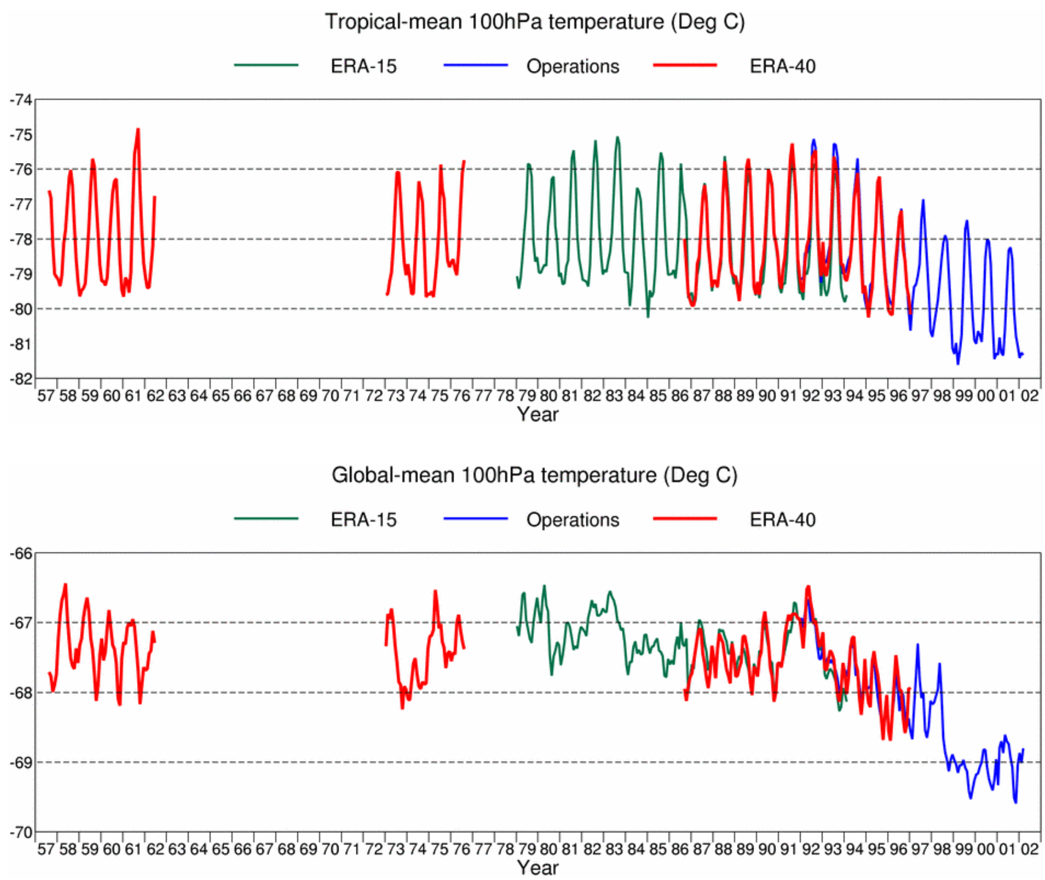

Fig. 5 Tropical- and global-mean 100hPa temperature for months from August 1957 to March 2002, derived from ERA-15 (green), ECMWF operations (blue) and ERA-40 (red).

Fig. 5 shows time series of monthly-mean 100hPa temperature analyses averaged over the tropics and globally. Results are shown for ERA-15, ECMWF operations since mid-1992 and for the ERA-40 analyses available at the time of writing. In the period of overlap there is good agreement between these sets of analyses. A marked cooling over most of the 1990s is evident, though it remains to be seen whether ERA-40 will reproduce the very sharp fall from 1997 to 1998 seen in the operational analyses.

Results for the years before the ERA-15 period must be treated with more caution because of the absence or limited capability of the satellite observing system, and because of changes in radiosonde coverage and accuracy. The ERA-40 analyses for these years do not produce values that are grossly different from those produced in the early years of ERA-15, but this is not, of course, a guarantee of accuracy. Careful validation of the early analyses will be needed to establish their utility.

Near-stratopause temperatures

Although the ERA-40 system has model levels that extend into the mesosphere, the model exhibits a quite pronounced systematic error at upper levels. In data assimilation this error tends to be corrected when data from the highest-sounding satellite channels are assimilated, but the reanalysis is nevertheless vulnerable to any gaps in coverage or drifts in bias of these sounding channels, and is likely to exhibit jumps when satellite data first become available and (to a lesser extent) when major changes in instrument type occur.

To illustrate this, we present results from a pre-production run for autumn 1986 in which data from the uppermost stratospheric sounding channel, SSU-3, were inadvertently blacklisted from 24 October onwards. Fig. 6 presents time series of temperature at 1hPa for the background and analysis, showing how the data systematically produce warming increments, and how temperatures fall sharply by about 5K over a few days after 24 October. Moreover, this assimilation was started from an analysis for 1 September produced by an earlier assimilation that used no radiance data, and a rapid warming can be seen for the first days that satellite data were assimilated.

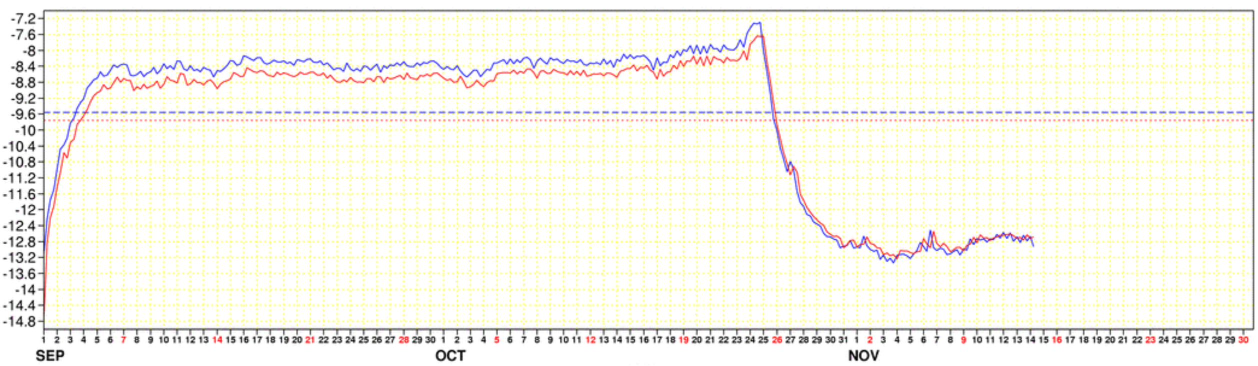

Fig. 6 Global- mean background (red) and analysed (blue) temperature at 1hPa (C) plotted 6 hourly for a prototype ERA-40 assimilation from 1 September to 14 November 1986.

Stratospheric specific humidity

Fig. 7 compares seven-year (1989-1995) zonal-mean ERA-40 stratospheric water-vapour distributions with the water-vapour climatologies for January and July derived from UARS measurements by Randel et al. (1998). The ERA-40 analyses are some 10-15% drier than the UARS retrievals in the upper stratosphere, and also in the lower stratosphere at higher latitudes where air moistened by methane oxidation has descended. An exception is the winter pole near the stratopause, where the model lacks the resolution to represent correctly the descent of air from the drier mesosphere.

The stratospheric humidity in ERA-40 is clearly a major step forward from the simple prescription of a specific humidity of $2.5x10^{-6}$ (equivalent to a 4ppmv volume mixing ratio) used in ERA-15, and the pattern of isolines in the upper stratosphere and lower down at high latitudes suggests a quite realistic pattern of advection of the humidity field. In the lower stratosphere, however, there is a too-rapid upward progression of the annual cycle of drying and moistening in the tropics, and generally too homogeneous and low humidities in middle latitudes. The latter may result from excessive mixing of dry air into middle latitudes from both the tropics and the cold, dry Antarctic polar night, and/or from inadequate descent from higher levels. Summertime moistening above the tropical tropopause appears to be much too large.

The parametrized upper-level source of water vapour due to methane oxidation used in the assimilating model for ERA-40 is based on relaxation towards a value of 6 ppmv for water vapour around the stratopause, and this is held fixed following the recommendation of the Working Group on Boundary Conditions and Atmospheric Composition that a trend over the more than 40-year reanalysis period should not be included. The value of 6 ppmv was based on publications earlier than that of Randel et al. (1998), whose results clearly indicate that a higher value would be more appropriate for recent years. The 1996 IPCC report estimated that methane had increased by 145% since pre-industrial times and by 6% over the decade 1984-1994. This suggests that the value of 6 ppmv used for ERA-40 is a reasonable compromise, being appropriate to the middle years of the reanalysis period. The results from ERA-40 presented here have nevertheless been used to justify an increase in the source term in the operational assimilation system from April 2002 onwards.

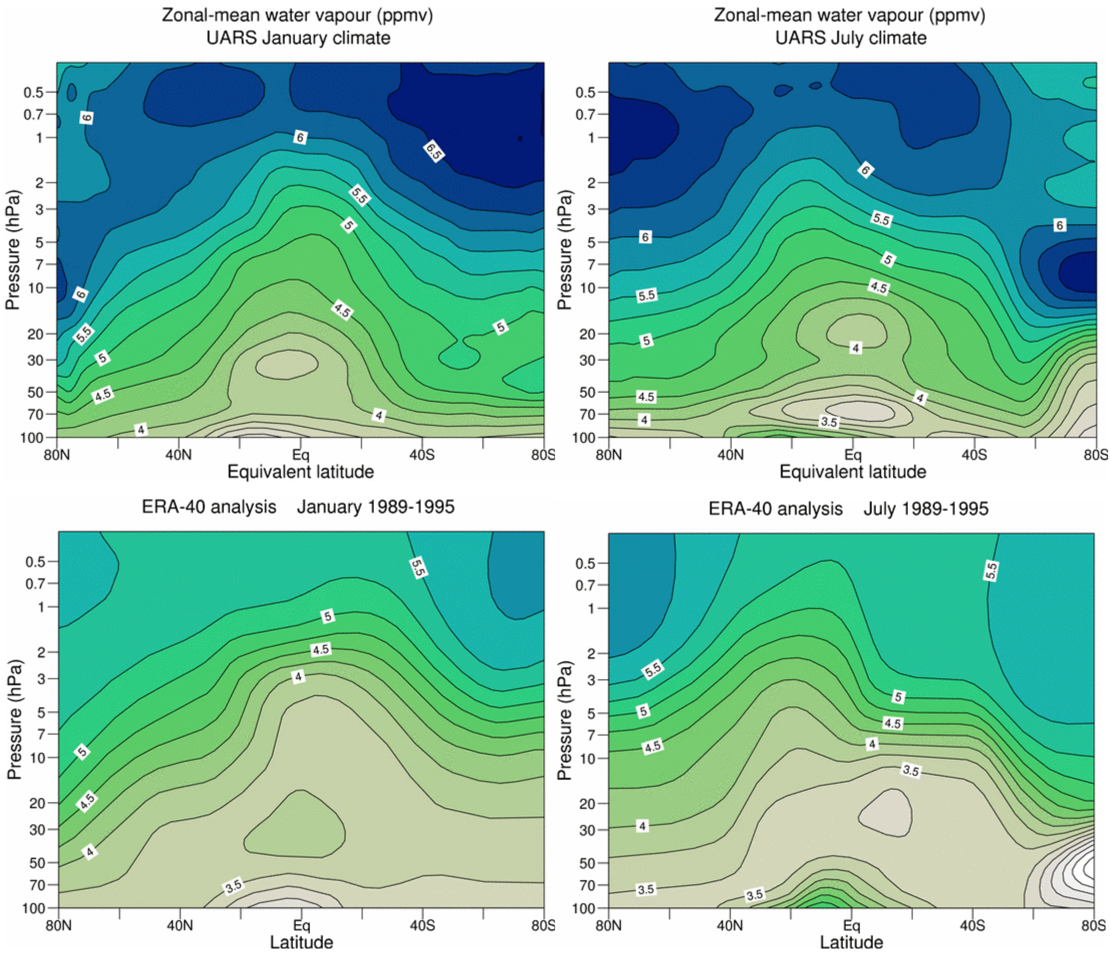

Fig. 7 Upper: Monthly-mean distribution of the mixing ratio of water vapour (ppmv) as a function of pressure and potential vorticity (expressed as equivalent latitude), based on UARS (HALOE, supplemented by MLS) data analysed by Randel et al. (1998). Lower: Monthly-mean distribution of the mixing ratio of water vapour (ppmv) as a function of pressure and latitude derived from ERA-40 analyses for the years 1989-1995. Results are shown for January in the left-hand panels and July in the right-hand panels. The contour interval is 0.25 ppmv, and darker shading denotes moister regions.

5. Short-range forecast verification

The ERA-40 system has also been monitored by regular verification of the short-range forecasts carried out as part of the production. A sample of results is presented here.

Fig. 8 shows root-mean-square (r.m.s.) one-day forecast errors from three systems for which results for 1989 are available. These are ECMWF operations at the time, ERA-15 and ERA-40. The forecasts from each system are verified against the analyses from the same system. Results are averaged over the extratropical northern and southern hemispheres and the tropics, and are shown for temperature and wind for standard pressure levels from 10 to 1000hPa. It is clear from Fig. 8 that the short-range forecasts of tropospheric and stratospheric winds and tropospheric temperature from ERA-40 are much more consistent with the corresponding verifying analyses than is the case for ERA-15 and operations in 1989. The ERA-40 forecasts also have the lowest errors for temperature in the extratropical stratosphere below 10hPa.

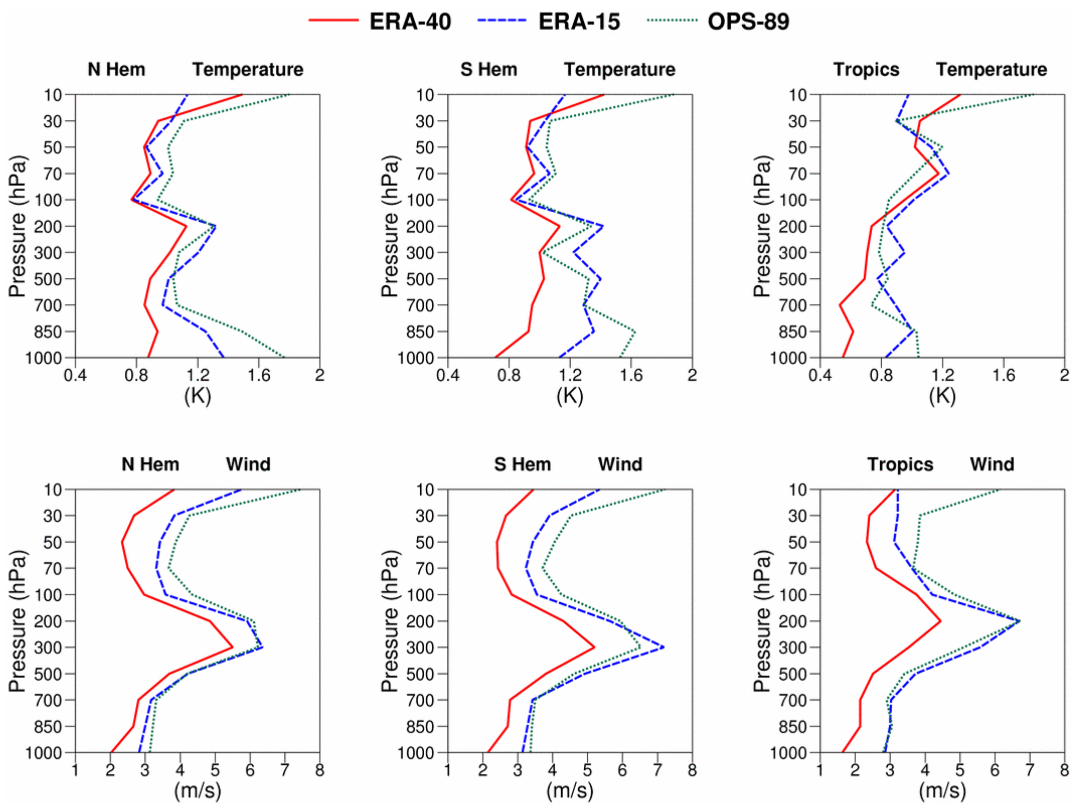

Fig. 8 Root-mean-square temperature and wind errors of one-day ERA-40, ERA-15 and operational forecasts, with verification against analyses. The errors of daily forecasts from 12UTC have been averaged for all dates in 1989.

Verification against analyses can give misleading results for short-range forecasts as it tends to favour a system in which little weight is given to observations in the data assimilation. Verification against radiosonde measurements provides a more independent validation of the forecasts, although variations in the spatial and temporal distribution of the verifying observations have to be taken into account in the detailed interpretation of results. Statistics corresponding to those shown in Fig. 8 are presented in Fig. 9 for the verification of the ERA-40, ERA-15 and operational forecasts against radiosondes. The general superiority of the short-range forecasts from ERA-40 is again evident. Differences are not as marked in the troposphere as seen in the verifications against analyses, but the ERA-40 forecasts of stratospheric temperature stand out much more clearly as the best in the verifications against radiosondes.

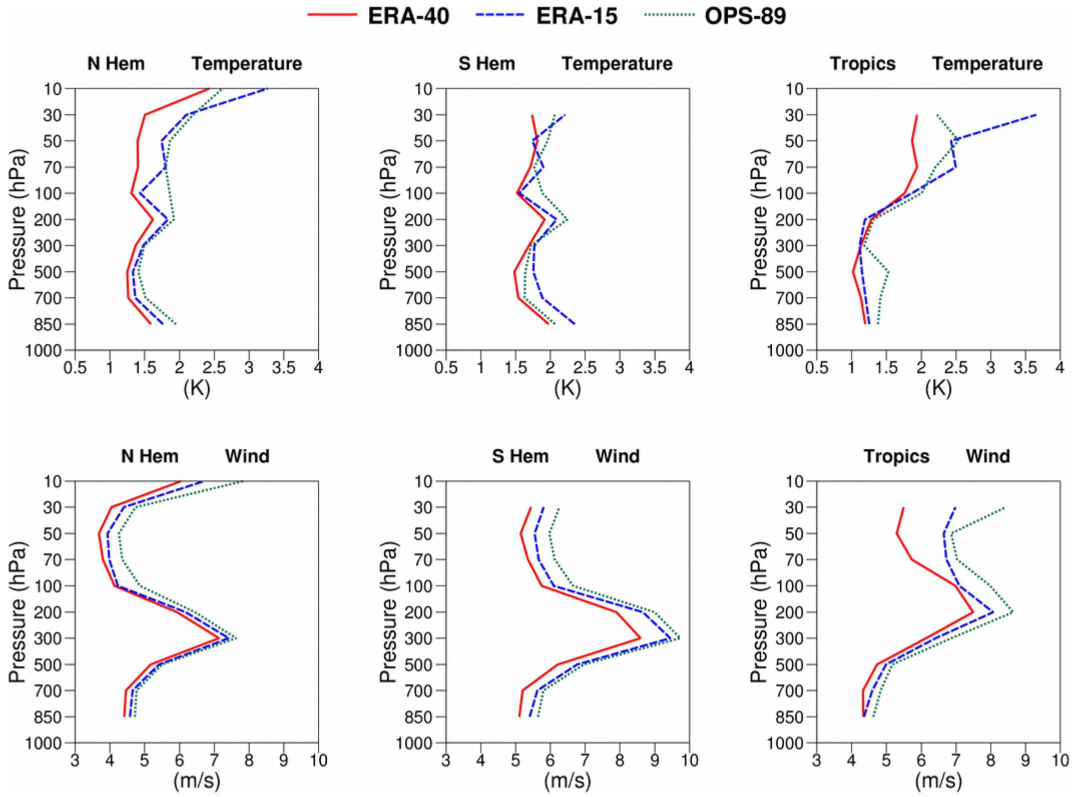

Fig. 9 Root-mean-square temperature and wind errors of one-day ERA-40, ERA-15 and operational forecasts for 1989, with verification against radoisonde measurements.

Root-mean-square one-day forecast errors for the ERA-40 system are displayed in Fig. 10 for each of the years 1989, 1991, 1993 and 1995. Verification is against analyses. The interannual differences in these statistics are very much smaller than the differences between ERA-40 and ERA-15 presented in Fig. 8 for 1989. Nevertheless, larger errors are seen for 1995 in tropical upper-tropospheric temperature forecasts, in tropical low-level wind forecasts and in 200hPa southern hemisphere wind forecasts. These larger forecast errors may be linked with problems in the assimilation of the raw HIRS radiances, which since the workshop has been shown to be primarily responsible for increased humidity increments and raised levels of precipitation over tropical oceans from 1991 onwards. The latter arose as a long-term consequence of the volcanic eruption of Mt Pinatubo in June 1991, which led to aerosol-effects on HIRS radiances that were inadequately handled in the ERA-40 production stream. The corresponding root-mean-square forecast errors measured against radiosonde data, presented in Fig. 11, show less of a problem in the tropics for 1995, but this may simply be because such verification cannot capture an increase in error that is predominantly located over the oceans.

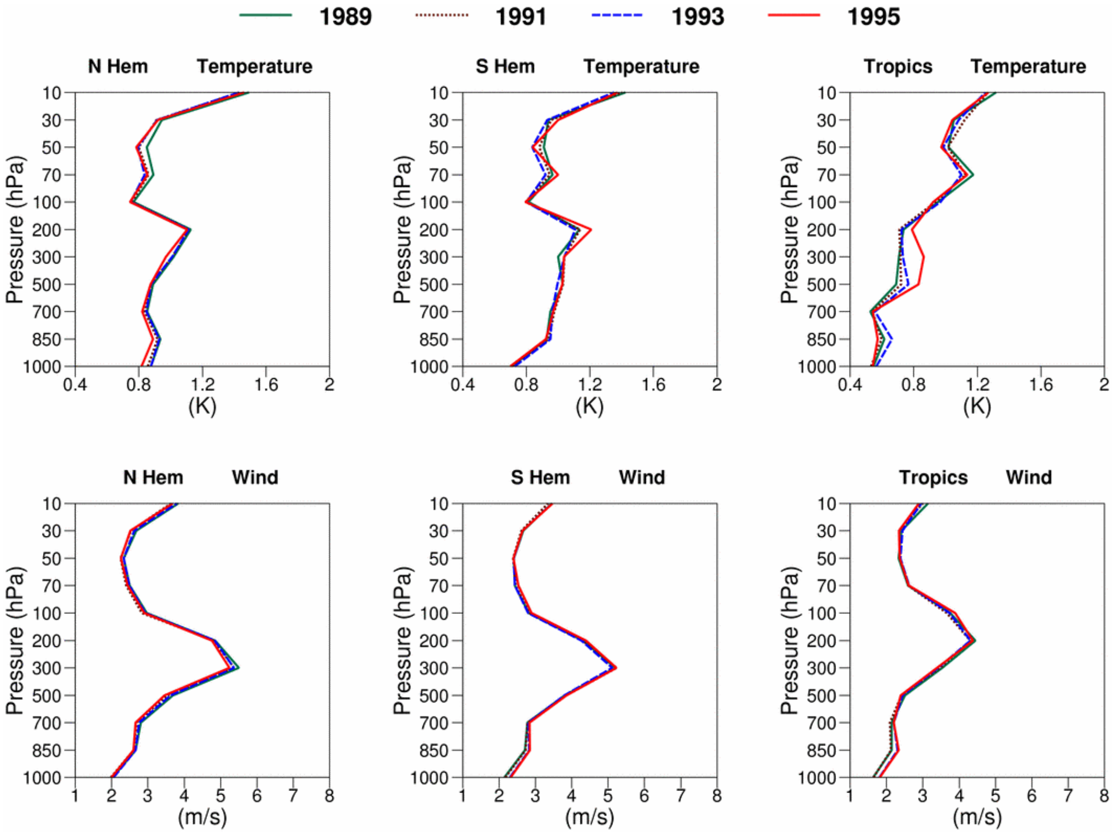

Fig. 10 Root-mean-square temperature and wind errors of one-day ERA-40 forecasts, with verification against analyses. The errors of daily forecasts from 12UTC have been averaged for all dates in 1989, 1991, 1993 and 1995.

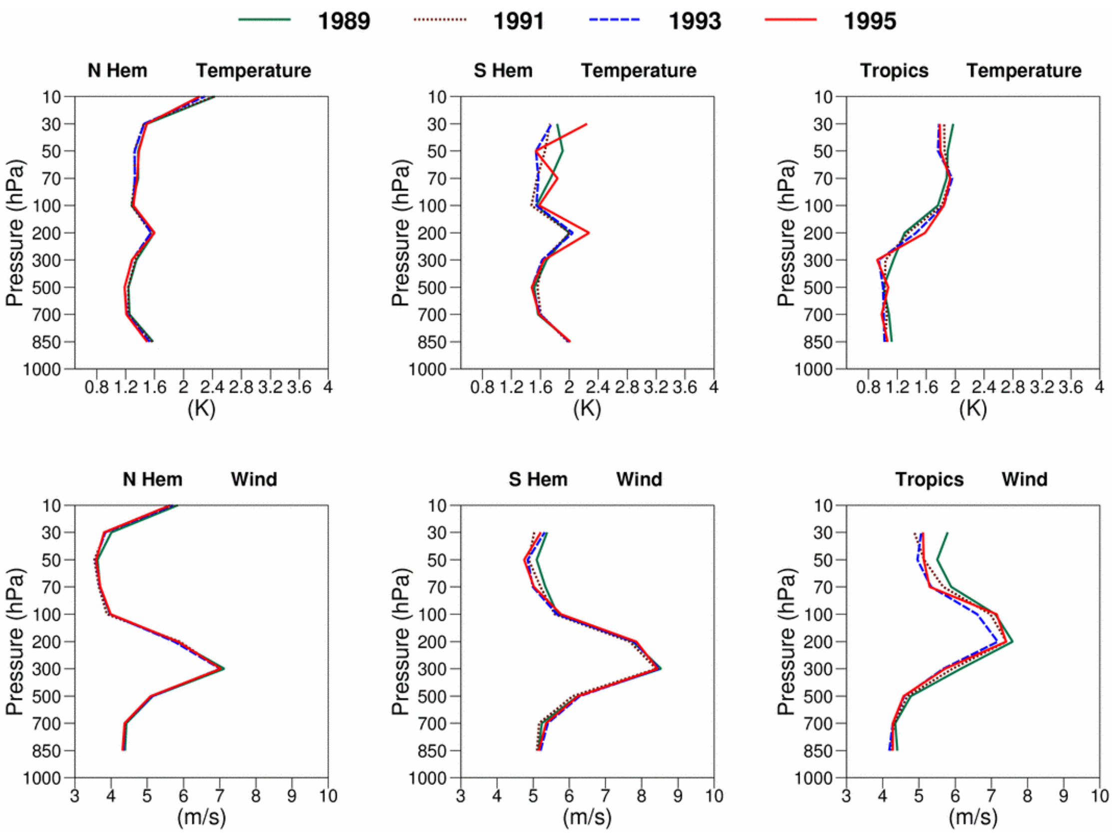

Fig. 11 Root-mean-square temperature and wind errors of one-day ERA-40 forecasts, with verification against radiosondes. The errors of daily forecasts from 12UTC have been averaged for all dates in 1989, 1991, 1993 and 1995.

Annual-mean root-mean-square errors of one-day 500hPa height forecasts for the extratropical northern and southern hemispheres are presented in Fig. 12. The ERA-40 forecast errors for the years from 1989 to 1996 are shown alongside the errors of ERA-15 and ECMWF operational forecasts for 1989, and also alongside the errors of the operational forecasts produced in 2001 by ECMWF, the Met Office and NCEP. Again, the ERA-40 results can be seen to be clearly better than the results from ERA-15 and from ECMWF's 1989 operations. They do not match those obtained operationally by ECMWF in 2001, when the T511 4D-Var system was in operational use and some enhancements to the observing system had occurred. They are, however, not much poorer than those from the 2001 operational systems of the Met Office and NCEP. Comparison with results for ECMWF operations presented by Simmons and Hollingsworth(2002) shows ERA-40 performance to be similar to that achieved by ECMWF operations in 1997 for the northern hemisphere and from mid-1998 to mid-1999 for the southern hemisphere, as judged by one-day 500hPa height forecast errors.

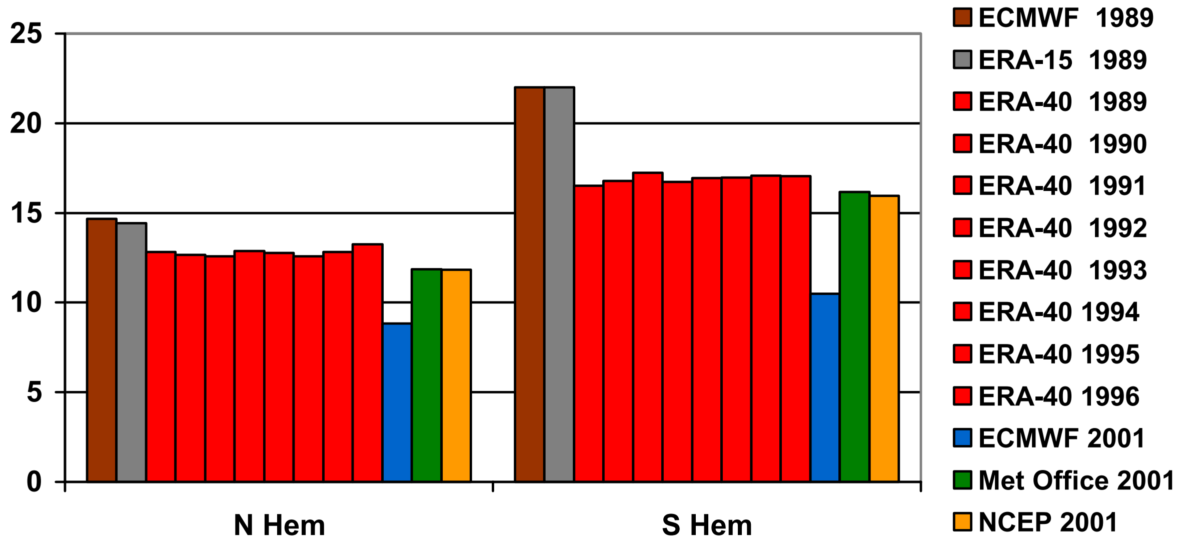

Fig. 12 Annual-average root-mean-square errors of 12UTC 500hPa height forecasts at one-day range from ERA-40 from 1989 to 1996, from ERA-15 and ECMWF operations for 1989, and from ECMWF, Met Office and NCEP operations for 2001, for the extratropical northern and southern hemispheres.

As a final example, Fig. 13 compares the one-day 500hPa height forecast errors, measured against analyses, from ERA-40 for the three periods 1958-60, 1973-75 and 1989-91. These periods were chosen to be representative of three distinct phases of the observing system. For the northern hemisphere the figure shows the expected reduction in forecast error as the observing system is improved from its pre-satellite state (1958-60), through the early satellite period (1973-75) when radiosonde numbers were especially high and aircraft data were beginning to become readily available, to the 1989-91 period by which time there had been a very substantial enhancement of the satellite component of the observing system in particular.

That the same is not evident for the southern hemisphere serves as a warning regarding the quality of the ERA-40 analyses in this hemisphere for the earlier years. The apparent one-day forecast errors shown in Fig. 13 are only a little larger in the southern than in the northern hemisphere in 1958-60, and the southern hemisphere errors are measured to be larger in 1973-75 than in 1958-60. The relatively low southern hemisphere errors for 1958-60 as measured here by verification against analyses almost certainly reflects the very limited coverage of upper-air (radiosonde) data at the time, resulting in verifying analyses that inherit more characteristics of the short-range background forecasts than would have been the case had accurate observational data been widely available to correct background-forecast error. The southern hemisphere errors measured for the 1973-75 ERA-40 forecasts are similar in magnitude to the southern hemisphere errors of operational ECMWF forecasts in 1986 and 1987.

Fig. 13 also shows results computed for a southern hemispheric subdomain encompassing Australia and New Zealand. More confidence can be placed in the accuracy of the verifying analysis for a region such as this where observational coverage is better than for the hemisphere as a whole, and it is indeed seen in this case that the short-range forecast errors for 1958-60 are larger than for 1973-75. The errors for 1989-91 are substantially smaller than those for both earlier periods over this region.

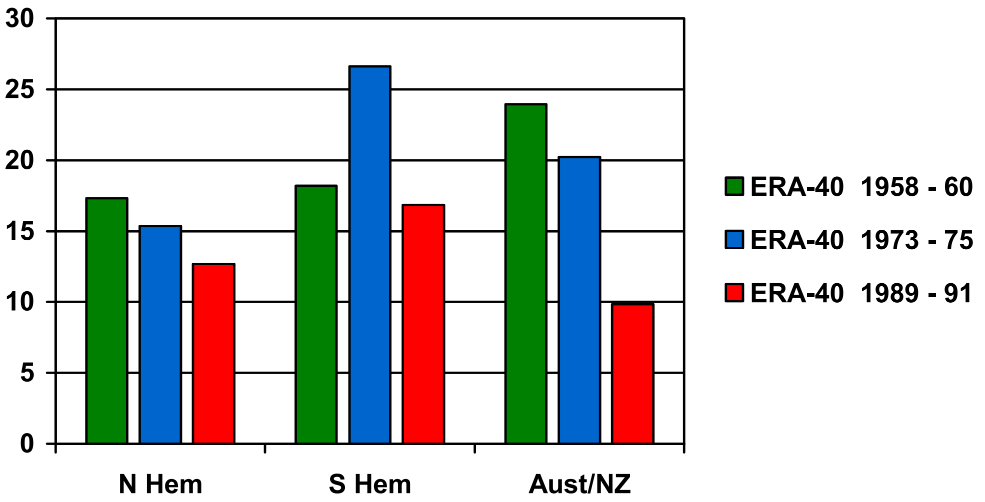

Fig. 13 Root-mean-square errors of 12UTC 500hPa height forecasts at one-day range from ERA-40, averaged over the periods 1958-60, 1973-75 and 1989-91, for the extratropical northern and southern hemispheres, and for a region covering Australia and New Zealand (from 12.5^{0}S to 45^{0}S and from 120^{0}E to 175^{0}E).

6. Concluding remarks

The data assimilation system that is being used for ERA-40 has been based primarily on the software version of the ECMWF data assimilation and forecasting system that was made operational in mid-2001, although it includes some changes and components that did not become operational until 2002. Some developments have had to be made specifically for ERA-40, to adapt the system for three- rather than four-dimensional variational data assimilation and for the lower horizontal resolution of the assimilating model, to cater for assimilation of older types of data and to introduce a more extensive range of output products. In this paper the principal differences from the system used for the earlier ERA-15 reanalysis have been identified.

Monitoring of ERA-40 production has been and remains a critical task. Several examples have been included in this paper. Some of the problems reported at the workshop have proven to be longer-lived and more serious than at first thought, especially as regards the tropical hydrology. Analysis quality for the modern satellite era nevertheless appears to be in most respects better than that of ERA-15, and indeed better than that of the operational ECMWF analyses for all but the most recent few years which benefit both from the establishment of the high-resolution 4D-Var assimilation system and from further improvements to the observing system. In addition, ERA-40 provides a much wider range of products, and products of greater vertical extent and over a longer period, than ERA-15.

The data assimilation system for ERA-40 has been implemented using standard ECMWF software facilities for running research experiments wherever possible. This has enabled the system to be made available for general research use within ECMWF and its Member States. The system has been run remotely from two Member States to perform observing-system studies. Prototype versions of the system proved invaluable in the testing of revisions that were in due course implemented both in ECMWF operations and in the production reanalyses.

Acknowledgements

The ERA-40 data assimilation system has benefited from the work of very many people, not only those active at ECMWF in the general development of the forecasting system and in its adaptation to ERA-40, but also those active elsewhere in the validation of experimental and production analyses, and those who earlier helped with advice and data during the planning of the project. The contributions of the project partners to the planning and validation are acknowledged in particular.

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. R. Meteorol. Soc.,* **124**, 1831-1860.

Andersson, E., and Järvinen, H. 1999: Variational quality control. *Q. J. R. Meteorol. Soc.,* **125**, 697-722.

Bouttier, F. 2001: The development of 12-hourly 4D-Var. *ECMWF Tech. Memo.*, **348**, 21pp.

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

Courtier, P., and Naughton, M. 1994: A pole problem in the reduced Gaussian grid. *Quart. J. R. Meteorol. Soc.*, **120**, 1389-1407

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.

Fisher, M., and Courtier, P. 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. *ECMWF Tech. Memo.*, **220**, 26pp.

Gérard, E. and Saunders, R.W. 1999: Four-dimensional variational assimilation of Special Sensor Microwave/Imager total column water vapour in the ECMWF model. *Quart. J. R. Meteorol. Soc.*, **125**, 3077-3101.

Gibson, J.K., Kållberg, P., Uppala, S., Nomura, A., Hernandez, A., and Serrano, E. 1997: ERA Description. ECMWF Re-Analysis Final Report Series, **1**, 71pp.

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

Harris, B.A. and Kelly, G.A. 2001 A satellite radiance bias correction scheme for data assimilation. *Quart. J. R. Meteorol. Soc.*, **127**, 1453-1468.

Hortal, M. and Simmons, A.J. 1991: Use of reduced Gaussian grids in spectral models. *Mon. Wea. Rev.,* **119**, 1057-1074.

Hortal, M. 2002: The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ECMWF forecast model. *Quart. J. R. Meteorol. Soc.*, **128**, *in press*.

Jakob, C. and Klein, S. A. 2000: A parametrization of the effects of cloud and precipitation overlap for use in general-circulation models. *Quart. J. R. Meteorol. Soc.*, **126**, 2525-2544.

Jakob, C., Andersson, A., Beljaars, A., Buizza, R., Fisher, M., Gérard, E., Ghelli, A., Janssen, P., Kelly, G., McNally, A.P., Miller, M., Simmons, A., Teixeira, J., and Viterbo, P. 2000: The IFS cycle CY21r4 made operational in October 1999. *ECMWF Newsletter*, **87**, 2-9.

Janssen, P.A.E.M., Doyle, J.D., Bidlot, J., Hansen, B., Isaksen, L., and Viterbo, P. 2002: Impact and feedback of ocean waves on the atmosphere. *Adv. Fluid. Mech.*, in press.

Järvinen, H., and Undén, P. 1997: Observation screening and background quality control in the ECMWF 3D-Var data assimilation system. *ECMWF Tech. Memo.*, **236**, 33pp

Järvinen, H., Andersson, E. and Bouttier, F. 1999: Variational assimilation of time sequences of surface observations with serially correlated errors. *Tellus*, 51A, 469-488.

Kållberg, P. 1997: Aspects of the re-analysed climate. ECMWF Re-Analysis Final Report Series, **2**, 89pp.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D., 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.

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.

McNally, A.P., Andersson, E., Kelly, G.A. and Saunders, R.W. 1999: The use of raw TOVS/ATOVS radiances in the ECMWF 4D-Var assimilation system. *ECMWF Newsletter*, **83**, 2-7.

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.

Marécal, V., Gérard, E., Mahfouf, J.-F., and Bauer, P. 2001: The comparative impact of the assimilation of the SSM/I and TMI brightness temperatures in the ECMWF 4D-Var system. *Quart. J. R. Meteorol. Soc.*, **127**, 1123-1142.

Miller, M. 1999: Resolution studies. *ECMWF Tech. Memo.*, **299**, 49pp.

Miller, M. and Hortal, M. 2001: The performance of 12-hourly 4D-Var in the T511/159 system, and the choice of timestep. Annex A of Bouttier (2001).

Morcrette, J.-J., Mlawer, E.J., Iacono, M.J. and Clough, S.A. 2001: Impact of RRTM in the ECMWF forecast system. *ECMWF Newsletter*, **91**, 2-9.

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

Randel, W.J., Wu, F., Russell, J.M.,III, Roche, A., and Waters, J. W. 1998: Seasonal cycles and QBO variations in stratospheric CH_4 and H_2O observed in UARS HALOE data. *J. Atmos. Sci.*, **55**, 163-185.

Rohn, M., Kelly, G.A., and Saunders, R.W. 2001: Impact of new cloud motion wind product from Meteosat on NWP analyses and forecasts. *Mon. Weather Rev.*, **129**, 2392-2403.

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.J., Untch, A., Jakob, C., Kållberg, P. and Undén, P. 1999: Stratospheric water vapour and tropical tropopause temperatures in ECMWF analyses and multi-year simulations. *Quart. J. Roy. Meteor. Soc.,* **125**, 353-386.

Teixeira, J. 1999: The impact of increased boundary layer resolution on the ECMWF forecast system. *ECMWF Tech. Memo.*, **268**, 55 pp.

Temperton, C., Hortal, H. and Simmons, A.J. 2001: A two-time-level semi-Lagrangian global spectral model. *Quart. J. Roy. Meteor. Soc.,* **127**, 111-128.

Untch, A. and Simmons, A.J. 1999: Increased stratospheric resolution in the ECMWF forecasting system. ECMWF Newsletter, 82, 2-8.

van den Hurk, B.J.J.M., Viterbo, P., Beljaars, A.C.M., and Betts, A.K. 2000: Offline validation of the ERA40 surface scheme. *ECMWF Tech. Memo.*, **295**, 42pp.

Viterbo, P., Beljaars, A., Mahfouf, J.-F. and Teixeira, J. 1999: The representation of soil moisture freezing and its impact on the stable boundary layer. *Quart. J. Roy. Meteor. Soc.,* **125**, 2401-2426.

Wedi, N.P. 1999: The numerical coupling of the physical parametrizations to the "dynamical"equations in a forecast model. *ECMWF Tech. Memo.*, **274**, 37pp.