RMetS
Royal Meteorological Society

The ERA-Interim reanalysis: configuration and performance of the data assimilation system

D. P. Dee[a]*, S. M. Uppala[a], A. J. Simmons[a], P. Berrisford[a], P. Poli[a], S. Kobayashi[b], U. Andrae[c], M. A. Balmaseda[a], G. Balsamo[a], P. Bauer[a], P. Bechtold[a], A. C. M. Beljaars[a], L. van de Berg[d], J. Bidlot[a], N. Bormann[a], C. Delsol[a], R. Dragani[a], M. Fuentes[a], A. J. Geer[a], L. Haimberger[e], S. B. Healy[a], H. Hersbach[a], E. V. Hólm[a], L. Isaksen[a], P. Kållberg[c], M. Köhler[a], M. Matricardi[a], A. P. McNally[a], B. M. Monge-Sanz[f], J.-J. Morcrette[a], B.-K. Park[g], C. Peubey[a], P. de Rosnay[a], C. Tavolato[e], J.-N. Thépaut[a] and F. Vitart[a]

[a] *European Centre for Medium-Range Weather Forecasts, Reading, UK*
[b] *Japan Meteorological Agency, Tokyo, Japan*
[c] *Swedish Meteorological and Hydrological Institute, Norrköping, Sweden*
[d] *European Organisation for the Exploitation of Meteorological Satellites, Darmstadt, Germany*
[e] *Department of Meteorology and Geophysics, University of Vienna, Austria*
[f] *School of Earth and Environment, University of Leeds, UK*
[g] *Korea Meteorological Administration, Seoul, Korea*
*Correspondence to: D. P. Dee, ECMWF, Shinfield Park, Reading RG2 9AX, UK. E-mail: dick.dee@ecmwf.int

ERA-Interim is the latest global atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). The ERA-Interim project was conducted in part to prepare for a new atmospheric reanalysis to replace ERA-40, which will extend back to the early part of the twentieth century. This article describes the forecast model, data assimilation method, and input datasets used to produce ERA-Interim, and discusses the performance of the system. Special emphasis is placed on various difficulties encountered in the production of ERA-40, including the representation of the hydrological cycle, the quality of the stratospheric circulation, and the consistency in time of the reanalysed fields. We provide evidence for substantial improvements in each of these aspects. We also identify areas where further work is needed and describe opportunities and objectives for future reanalysis projects at ECMWF.

Key Words: ERA-40; 4D-Var; hydrological cycle; stratospheric circulation; observations; forecast model

Received 21 January 2011; Revised 22 March 2011; Accepted 23 March 2011; Published online in Wiley Online Library 28 April 2011

Citation: Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**: 553–597. DOI:10.1002/qj.828

1. Introduction

ERA-Interim is the latest global atmospheric reanalysis produced by the ECMWF*. ERA-Interim covers the period from 1 January 1989 onwards, and continues to be extended forward in near-real time. An extension from 1979 to 1989 is currently in preparation. Gridded data products include a large variety of 3-hourly surface parameters, describing weather as well as ocean-wave and land-surface conditions, and 6-hourly upper-air parameters covering the troposphere and stratosphere. Vertical integrals of atmospheric fluxes,

*All acronyms used in this article are given in the appendix.

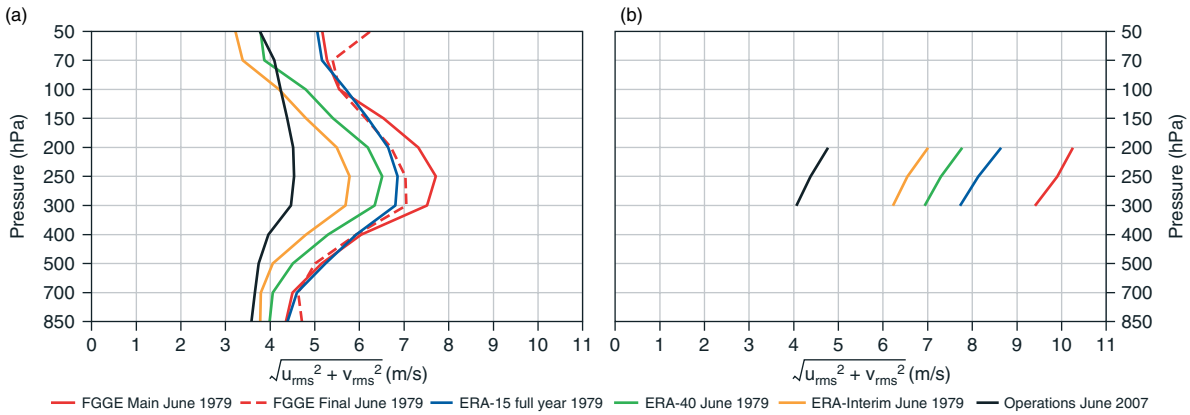

Figure 1. Globally averaged RMS errors in upper-air winds from short-range forecasts produced in ECMWF reanalyses, relative to (a) radiosonde observations and (b) aircraft reports. Both panels show curves for two successive FGGE reanalyses, ERA-15, ERA-40, and ERA-Interim, which is currently being extended back to 1979. Data are from June 1979. For comparison, background errors in wind estimates from ECMWF operations for June 2007 are also shown. Figure adapted from Uppala *et al.* (2008).

monthly averages for many of the parameters, and other derived fields have also been produced. Berrisford *et al.* (2009) provide a detailed description of the ERA-Interim product archive. Information about the current status of ERA-Interim production, availability of data online, and near-real-time updates of various climate indicators derived from ERA-Interim data, can be found at http://www.ecmwf.int/research/era

The purpose of this article is to describe the model, data assimilation method, and observations used to produce ERA-Interim, and to present some basic measures of performance. Our assessments focus on progress made since the completion of the ERA-40 reanalysis in 2002 (Uppala *et al.*, 2005). ERA-Interim was conceived in part to prepare for a future, more ambitious reanalysis project at ECMWF, which will span the entire twentieth century. The primary goal for ERA-Interim has been to address several difficult data assimilation problems encountered during the production of ERA-40. These are mainly related to the representation of the hydrological cycle, the quality of the stratospheric circulation, and the consistency in time of reanalysed geophysical fields. A second objective was to improve on various technical aspects of reanalysis such as data selection, quality control, bias correction, and performance monitoring, each of which can have a major impact on the quality of the reanalysis products.

Reanalysis is a relatively young field, with origins in the exploitation of meteorological data collected for the FGGE in 1979. Those data were reanalysed several times, mainly in order to learn how to make better use of observations to initialise numerical weather forecasts. However, it was soon realised that the datasets generated by such a reanalysis can be of great value for atmospheric research. Reanalysis data provide a multivariate, spatially complete, and coherent record of the global atmospheric circulation. Unlike archived weather analyses from operational forecasting systems, a reanalysis is produced with a single version of a data assimilation system –including the forecast model used –and is therefore not affected by changes in method.

Successive generations of atmospheric reanalyses produced at various institutes have improved in quality as a result of better models, better input data, and better assimilation methods. These include the global reanalyses from NCEP (Kalnay *et al.*, 2006; Saha *et al.*, 2010), from JMA (Onogi *et al.*, 2007), and NASA (Schubert *et al.*, 1993; Rienecker *et al.*, 2011), in addition to those from ECMWF (Gibson *et al.*, 1997; Uppala *et al.*, 2005). The reanalyses have generated a growing variety of useful data products, spanning longer time periods at increasing spatial and temporal resolutions. A global reanalysis extending back to the late nineteenth century was recently produced by NOAA in collaboration with CIRES, using only surface pressure observations and prior estimates of SST and sea-ice distributions (Compo *et al.*, 2011).

With a large and diverse user base, the quality requirements for reanalysis products have evolved accordingly. The primary requirement, clearly, is that a reanalysis represents available observations. Many users regard reanalysis products as equivalent to observations, even if this is not always justifiable. Nevertheless, substantial progress has been made over the years in producing global estimates of the basic dynamical fields that are consistent with observations given their estimated uncertainties. This progress is nicely illustrated by Figure 1, which shows the fit to upper-air wind reports from radiosondes and aircraft, for each of the reanalyses produced at ECMWF to date. The curves show global RMS errors of the background wind estimates at observation locations, obtained from the short-range forecasts used for the data assimilation in each reanalysis. The data are from June 1979, when the global observing system included a single polar-orbiting satellite (TIROS-N) carrying HIRS, MSU, and SSU instruments. Also included for reference are the accuracy levels attained with the ECMWF operational forecasting system in June 2007, with the benefit of a vastly improved global observing network.

A further requirement for a multivariate reanalysis is physical coherence, meaning that estimated parameters must be consistent with the laws of physics as well as with observations. This is a defining property of reanalysis, which differentiates it from other methods for estimating geophysical parameters from observations. It is achieved by using a forecast model as the unifying context in which to assimilate and compare observations of various types and from multiple sources. A sufficiently realistic model is able to extrapolate information from locally observed parameters to unobserved parameters at nearby locations, and it can also propagate this information forward in time. In this way it is possible, for example, to obtain meaningful precipitation

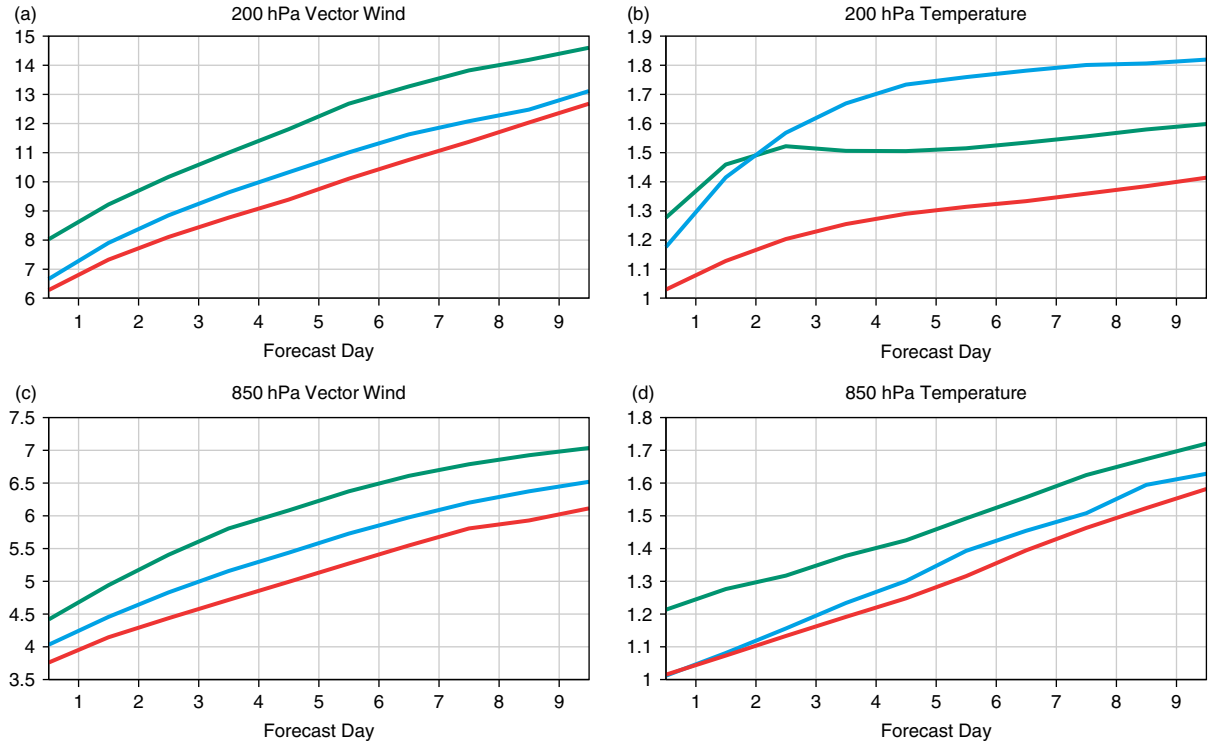

Figure 2. RMS forecast errors for (a) tropical wind vectors ($m\,s^{-1}$) and (b) temperatures (K) at 200 hPa, averaged over all forecasts issued daily at 1200 UTC in 1989, for ERA-Interim (red), ERA-40 (blue), and for the version of the ECMWF forecasting system operational at the time (green). Forecast errors for each system are relative to a fixed set of radiosonde observations. (c, d) are as (a, b), but for 850 hPa.

estimates from a reanalysis of temperature, humidity and wind observations.

The ability to predict future observations with the assimilating forecast model, using initial conditions from the reanalysis, provides a very useful measure of success in this respect. The skill of the reforecasts depends on the accuracy, completeness, and physical coherence of the reanalysed fields, as well as on the quality of the model used. To illustrate, Figure 2 shows the average skill of daily forecasts of tropical temperatures and winds issued in 1989, obtained with ERA-Interim, ERA-40, and with the ECMWF forecasting system operational at the time. RMS errors relative to available radiosonde observations in the lower and upper troposphere are shown for latitudes between 20°S and 20°N. The differences between pairs of curves in each panel reflect incremental improvements in data assimilation achieved from 1989 to 2001, when ERA-40 production began, and from 2001 to 2006, when the configuration of the ERA-Interim data assimilation system was established.

An emerging requirement for climate applications of reanalysis data is the accurate representation of variability on interannual and decadal time scales, leading, for example, to the ability to estimate trends with confidence. A goal of reanalysis has always been to produce an homogeneous record of past atmospheric evolution that is free of shifts and other spurious signals introduced by changes in the assimilation system. Nevertheless, the representation of climate signals in reanalysis is inevitably affected by changes in the global observing system and by the presence of time-varying biases in models and observations. This circumstance is not unique to reanalysis; its manifestations can be found in all existing historic analyses of observations. Direct measurements of many of the most basic global properties of the atmosphere, such as its average temperature, do not exist. Estimation of climate change indicators must therefore always involve statistical modelling and analysis, and this invariably requires some type of extrapolation of information from incomplete and fundamentally uncertain measurements.

The extent to which climate-quality requirements for reanalysis data (or any other data derived from observations) can be realised is open to some debate (e.g. Thorne and Vose, 2010; Dee *et al.*, 2011). Nevertheless, the use of information from modern reanalysis systems for monitoring and assessment of climate change is clearly taking hold (Willett *et al.*, 2010). Uncertainties in these assessments are very difficult to quantify; it is therefore useful and important to consider reanalysis data in tandem with the more traditional, observation-only climate datasets (e.g. Simmons *et al.*, 2010).

In any case, to build confidence in climate change information derived from reanalysis data requires that all ingredients are made fully transparent. These include, broadly, the data assimilation methodology, the forecast model, and the input observations. The organisation of this paper reflects these three categories, with a section devoted to the description of each. In section 5 we summarise our evaluation of the performance of ERA-Interim, with a particular focus on progress made since ERA-40. We then conclude and discuss future prospects in section 6.

2. Data assimilation

The ERA-Interim reanalysis is produced with a sequential data assimilation scheme, advancing forward in time using 12-hourly analysis cycles. In each cycle, available observations are combined with prior information from

a forecast model to estimate the evolving state of the global atmosphere and its underlying surface. This involves computing a variational analysis of the basic upper-air atmospheric fields (temperature, wind, humidity, ozone, surface pressure), followed by separate analyses of near-surface parameters (2 m temperature and 2 m humidity), soil moisture and soil temperature, snow, and ocean waves. The analyses are then used to initialise a short-range model forecast, which provides the prior state estimates needed for the next analysis cycle.

The forecast model has a crucial role in the data assimilation process. Use of the model equations makes it possible to extrapolate information from locally observed parameters to unobserved parameters in a physically meaningful way, and also to carry this information forward in time. The skill and accuracy of the forecast model determines how well the assimilated information can be retained; better forecasts mean that smaller adjustments are needed to maintain consistency with observations as time evolves.

Additionally, while producing a forecast, the model estimates a wide variety of physical parameters such as precipitation, turbulent fluxes, radiation fields, cloud properties, soil moisture, etc. Even if not directly observed, these are constrained by the observations used to initialise the forecast. The accuracy of these model-generated estimates naturally depends on the quality of the model physics as well as that of the analysis.

The data assimilation thus produces a coherent record of the global atmospheric evolution constrained by the observations available during the period of reanalysis. The ERA-Interim archive currently contains 6-hourly gridded estimates of three-dimensional (3D) meteorological variables, and 3-hourly estimates of a large number of surface parameters and other two-dimensional (2D) fields, for all dates from 1 January 1989. The complete contents of this archive are described in Berrisford *et al.* (2009). Data for dates from 1 January 1979 will soon be added to the archive.

2.1. *Atmospheric analysis*

The core component of the ERA-Interim data assimilation system is the 12-hourly 4D-Var of the upper-air atmospheric state. The defining feature of 4D-Var is that it uses the forecast model to constrain the state evolution within each analysis window. The version of 4D-Var used for ERA-Interim also updates a set of parameter estimates that define bias corrections needed for the majority of satellite-based radiance observations.

Mathematically, the analysis can be described as the minimisation of

$$\begin{aligned} J(\mathbf{x},\boldsymbol{\beta}) =& (\mathbf{x}^{\mathrm{b}}-\mathbf{x})^{\mathrm{T}}\mathbf{B}_{\mathbf{x}}^{-1}(\mathbf{x}^{\mathrm{b}}-\mathbf{x}) \\ &+ (\boldsymbol{\beta}^{\mathrm{b}}-\boldsymbol{\beta})^{\mathrm{T}}\mathbf{B}_{\boldsymbol{\beta}}^{-1}(\boldsymbol{\beta}^{\mathrm{b}}-\boldsymbol{\beta}) \\ &+ [\mathbf{y}-\mathbf{h}(\mathbf{x},\boldsymbol{\beta})]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y}-\mathbf{h}(\mathbf{x},\boldsymbol{\beta})] \end{aligned} \quad (1)$$

jointly with respect to the control variables $(\mathbf{x},\boldsymbol{\beta})$. In 4D-Var the control $\mathbf{x}$ is typically the model initial state (at the beginning of the analysis window), which, in view of the model constraint, defines the state at any other time within the window. The additional control $\boldsymbol{\beta}$ contains parameters for the variational bias corrections (VarBC) applied to the radiance observations.

Input data for the analysis consist of prior (background) estimates $(\mathbf{x}^{\mathrm{b}},\boldsymbol{\beta}^{\mathrm{b}})$ for the controls, and a set of observations $\mathbf{y}$ that are valid within the current analysis window. Additional information needed to solve the minimisation includes specifications for the covariances $(\mathbf{B}_{\mathbf{x}},\mathbf{B}_{\boldsymbol{\beta}})$ of errors in the background estimates, and covariances $\mathbf{R}$ of errors in the observations. The latter account for measurement errors as well as the inability of the model to represent small-scale information contained in some of the observations. The background state estimate $\mathbf{x}^{\mathrm{b}}$ is obtained from a short-range forecast, which is initialised using the analysis produced in the previous cycle. Background estimates $\boldsymbol{\beta}^{\mathrm{b}}$ for the bias parameters are simply the estimates produced by the previous analysis.

The observation operator $\mathbf{h}(\mathbf{x},\boldsymbol{\beta})$ can be regarded as an extension of the forecast model; it is used to simulate observations given a model state, possibly making use of bias parameters to adjust for systematic errors. Its implementation involves integration of the model equations to advance the state estimate to the time of observation, followed by interpolation to the actual observation location, followed by simulation of the observable (e.g. surface pressure, temperature, humidity, wind speed and direction, emitted radiance, atmospheric refraction, etc.). The ability of the observation operator to accurately model observations affects the quality of the analysis; errors or inaccuracies in the observation operator result in incorrect or suboptimal interpretation of the available data. However, this source of error is in principle factored in the definition of $\mathbf{R}$, and may also be compensated by bias corrections.

The 4D-Var analysis in ERA-Interim is obtained by successive linearisations of the model and observation operator (Courtier *et al.*, 1994; Veersé and Thépaut, 1998). The algorithm consists of a pair of nested loops. The outer loop integrates the nonlinear forecast model, producing a 4D state estimate and simulated observations. The inner loop then solves a linearised version of Eq. (1) for the control variable increments, using the tangent linear and adjoint of a simplified version of the forecast model at lower horizontal resolution (Trémolet, 2004). For ERA-Interim, the spectral resolution of the outer loop is T255 (∼79 km), and two successive inner loops at resolutions T95 (∼210 km) and T159 (∼125 km) are used for the minimisation.

2.1.1. *Benefits of 4D-Var*

The use of 4D-Var for the atmospheric analysis in ERA-Interim is a major step forward from ERA-40. Data assimilation in ERA-40 was based on a 6-hourly 3D variational analysis scheme (3D-Var; Courtier *et al.*, 1998), using the so-called FGAT configuration. Consistent use of the model equations in 4D-Var (e.g. in the implementation of the observation operator in Eq. (1)) can result in more effective use of observations, for example, by extracting information about tendencies in the mass field (Rabier *et al.*, 1998, 2000).

A key feature of 4D-Var is the flow-dependent influence of observations that results from using a forecast model to constrain the analysis (Thépaut *et al.*, 1996). The ability of the data assimilation system to exploit physical information implicit in the model equations can be very beneficial, especially where observations are sparse. Thépaut (2006) and Whitaker *et al.* (2009) have shown that 4D-Var outperforms 3D-Var in such situations, and that it is capable of producing accurate analyses of the large-scale tropospheric circulation based only on observations of surface pressure.

2.1.2. *Variational bias correction*

Another important advance in ERA-Interim is the inclusion of a completely automated scheme for correcting biases in satellite radiance observations, which makes use of the bias control variable in the variational analysis (Eq. (1)). The data assimilation system detects data events such as the appearance of new satellite data streams, and then initialises, updates, and keeps track of bias parameters for radiance data from all available satellites. Bias corrections for individual sensor channels are expressed in terms of a small set of predictors, which can depend on the atmospheric state at the observed location or on the state of the instrument itself. The bias parameters determine the linear combination of predictors used for correcting each radiance observation. They are continuously adjusted by the variational analysis to minimise inconsistencies among the available sources of information, including observations from conventional instruments such as radiosondes and aircraft.

2.1.3. *Background-error covariances*

The background-error covariances used in Eq. (1) determine how the analysis spreads locally observed information to nearby locations, and how it uses this information to adjust estimates of unobserved variables. The real significance of these covariances is that they completely define the spatial scales and multivariate constraints for all possible adjustments to the background state that the variational analysis is able to produce. The capacity of the data assimilation system for extracting and propagating useful information from observations is strongly affected by these constraints.

Most aspects of the background-error covariance model for ERA-Interim are similar to ERA-40. Multivariate covariances are defined by incorporating linear balance relations in the formulation (Derber and Bouttier, 1999). Vertical correlations can vary with wavenumber, e.g. to support deep correlations for broad horizontal structures and shallower correlations for narrow structures. Newly introduced in ERA-Interim is the use of wavelet-like weighting functions, to allow variations in spatial correlation scales that depend on both wavenumber and location (Fisher, 2004, 2005).

Within the constraints of this formulation, background-error correlations are calculated from statistics of an ensemble of 4D-Var assimilations (Fisher, 2003). Background-error variances are estimated during the data assimilation, using a simple updating scheme (Fisher and Courtier, 1995). The scheme uses the leading eigenvectors of the Hessian of Eq. (1) to approximate analysis errors, and includes a simple inflation model to account for error growth between analysis cycles. The variances obtained in this way weakly depend on the underlying atmospheric flow, due to the role of the forecast model in the 4D-Var Hessian. A more significant dependence of the background-error variances on the flow is produced by transforming the model state control variables for the variational analysis, using a nonlinear balance equation linearised about the background state (Fisher, 2003).

The methodology for specifying background-error covariances used in ERA-Interim may not be suitable for a reanalysis extending to earlier decades, since it is not designed to represent the effects of major changes in data coverage on the accuracy of the background estimates.

2.1.4. *Humidity and ozone analysis*

Inspired by difficulties with the representation of the hydrological cycle in ERA-40 (Uppala *et al.*, 2005), a completely revised humidity analysis scheme was developed by Hólm (2003). The new scheme involves a nonlinear transformation of the humidity control variable to render the humidity background errors more nearly Gaussian. The transformation normalizes relative humidity increments by a factor that depends on background estimates of relative humidity and vertical level. No humidity increments are allowed in the stratosphere; this is achieved by prescribing very small humidity background errors above the diagnosed tropopause. The modifications to the formulation of the humidity analysis and its impact on the assimilation of humidity-sensitive observations, from conventional as well as satellite instruments, are described by Andersson *et al.* (2005).

Ozone, which is a prognostic variable of the forecast model, is analysed simultaneously with the other model state variables in the 4D-Var analysis. The background-error covariance model for ozone in ERA-Interim is univariate, meaning that errors in the model-generated ozone estimates are assumed independent of those for other state variables. In a 3D-Var analysis this assumption implies that ozone observations can only affect the analysis of the ozone field itself. However, a 4D-Var analysis can change any aspect of the model initial state in order to improve the fit to ozone observations later within the analysis window, e.g. by redirecting the flow, provided this change does not deteriorate the fit to other observations (Peuch *et al.*, 2000). In theory, this feature can be beneficial –observations of trace gas concentrations may contain useful dynamical information –but in practice it requires highly accurate ozone information.

We found, in fact, that the assimilation of ozone profile data in ERA-Interim systematically caused large and unrealistic changes in temperature and winds near the top of the model, where the flow is not well constrained by other observations (Dee, 2008). These artificial changes to the circulation represented the solution of least cost (in the sense of Eq. (1)) for accommodating the large and systematic inconsistencies between the ozone observations and their model predictions. Spurious increments often occur in data assimilation when information from different sources is in conflict (Dee, 2005). The problem encountered with the assimilation of ozone profile data cannot be fully addressed until adequate bias corrections for the data have been developed (Dragani and Dee, 2008). As a temporary solution for ERA-Interim, the symptoms of the problem were removed by modifying the analysis scheme in order to prevent any direct influence of ozone observations on the dynamical fields.

2.1.5. *Radiative transfer modelling*

The observation operator used for simulating satellite radiance observations in ERA-Interim is based on RTTOV version 7, which incorporates the fast transmittance model described by Matricardi *et al.* (2004). ERA-40 used RTTOV-5 (Saunders *et al.*, 1999).

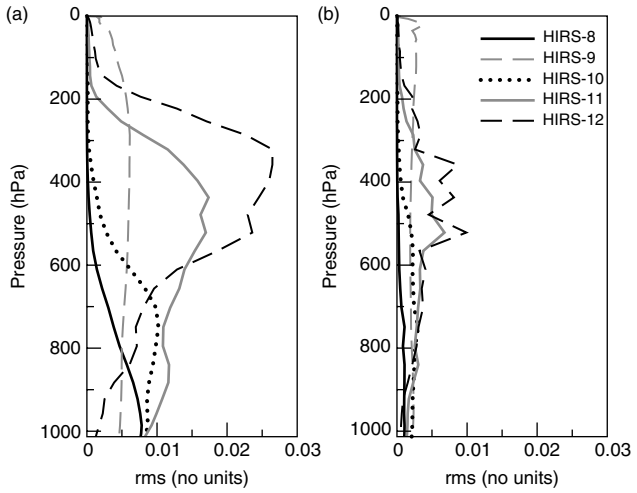

Figure 3. RMS of the difference between line-by-line HIRS channels 8, 9, 10, 11 and 12 layer-to-space transmittances for NOAA-14 for 117 independent profiles and six viewing angles, and (a) RTTOV-5, (b) RTTOV-7.

Matricardi *et al.* (2004) have shown that differences between fast-model and line-by-line calculations of transmittances are greatly reduced at all pressure levels when using RTTOV-7. Figure 3 presents RMS errors of transmittance profiles for channels 8–12 of the HIRS, obtained with RTTOV-5 and RTTOV-7. These results are based on a set of atmospheric profiles independent of those used for regressing the new fast transmittance model. The figure shows a ten-fold error reduction for the simulation of the HIRS water vapour channels (10–12), and a five-fold reduction for the ozone and window channels (8, 9).

A further modification to RTTOV-7 was implemented in ERA-Interim to remove an incorrect representation of the splitting of spectral lines due to the terrestrial magnetic field (Kobayashi *et al.*, 2009). The so-called Zeeman effect is significant at the low pressures that prevail in the upper stratosphere and therefore affects the interpretation of high-peaking stratospheric channels of AMSU-A, which constitute a critical source of information for upper-stratospheric temperatures since the introduction of AMSU-A in 1998.

2.1.6. 1D+4D-Var rain assimilation

Significant progress has been made in recent years in the ability to use passive microwave data in areas affected by cloud and precipitation. Initial work carried out by Marécal and Mahfouf (2000, 2002, 2004) was based on an off-line 1D-Var retrieval scheme for TCWV using derived rain rates as observations. The retrieved TCWV was then assimilated in 4D-Var as a pseudo-observation, similar to the 1D+4D-Var approach previously used for clear-sky radiances (Phalippou, 1996; Gérard and Saunders, 1999).

The scheme implemented in ERA-Interim uses rain-affected radiances rather than derived rain rates. The 1D-Var retrieval step, which is embedded in the first outer loop of the 4D-Var analysis, makes use of the linearised large-scale condensation and convection schemes of the forecast model. Data selection, quality control, and bias correction of the radiances are handled in this step as well. The TCWV retrievals thus obtained are then analysed, together with all other available observations, in the 4D-Var analysis (Bauer *et al.*, 2006a,b).

Table I. Sea-surface temperature and sea-ice concentration datasets used in ERA-Interim.

ERA-Interim dates	SST and SIC product used
January 1989–June 2001	NCEP 2D-Var sea surface temperature (NCEP 2D-Var)
July 2001–December 2001	NOAA Optimum Interpolation Sea Surface Temperature v2 (NCEP OISST v2)
January 2002–January 2009	NCEP Real-Time Global sea surface temperature (NCEP RTG)
From February 2009	Operational Sea Surface Temperature and Sea-Ice Analysis (OSTIA)

2.2. Surface analysis

As in ERA-40, the analysis of surface parameters over land and ocean is performed separately from the main atmospheric analysis, in several steps. First, an OI scheme produces 6-hourly estimates of screen-level temperature and dewpoint, combining synoptic observations over land with background estimates derived from the latest atmospheric analysis (Douville *et al.*, 1998). The analysis increments for screen-level temperature and humidity are subsequently used to update soil moisture and soil temperature estimates for each of the four layers of the land-surface model, by a simple empirical approach (Douville *et al.*, 2000; Mahfouf *et al.*, 2000). Snow depth, snow water equivalent, and snow density estimates generated by the forecast model are then updated based on a Cressman analysis of station observations of snow depth and (when available) snow cover data from satellites (Drusch *et al.*, 2004).

Finally, an OI analysis of ocean wave height is performed using wave-height data from space-borne radar altimeters, when available. Background estimates for this analysis are generated by the forecast model, which contains a fully coupled wave model describing the evolution of two-dimensional wave spectra at the sea surface. Analysed wave heights are used to adjust the model-predicted wave spectra based on assumptions about the contributions of wind-sea and swell spectra (Lionello *et al.*, 1992). SST and SIC are prescribed boundary conditions for the atmospheric model (Table I).

3. Forecast model

The ERA-Interim reanalysis is produced with the ECMWF IFS, which incorporates a forecast model with three fully coupled components for the atmosphere, land surface, and ocean waves. A history of changes introduced in the IFS since 1985 is maintained at http://www.ecmwf.int/products/data/operational_system/evolution. ERA-Interim is based on IFS release Cy31r2, used for operational forecasting at ECMWF from 12 December 2006 until 5 June 2007. Table II summarises relevant upgrades to the forecast model introduced during the five-year period since release Cy23r4, which became operational at ECMWF in December 2001 and was used to produce ERA-40. Several changes

Table II. Relevant changes to the ECMWF forecast model from Jan 2002 to September 2006, listed by IFS release. ERA-40 used release Cy23r4, and ERA-Interim uses Cy31r2, which is identical to Cy31r1 except for technical changes.

Release	Date	Description
Cy24r3	22 Jan 2002	Minor changes to convective precipitation and supersaturation checks
Cy25r1	09 Apr 2002	Revised short-wave radiation scheme
		Interactive radius of cloud droplets
		Retuning of land surface scheme
		Improved wind gust post-processing
		Bug fix for convective momentum transport
Cy25r3	14 Jan 2003	Improved cloud scheme numerics
		Revised ice settling
		Mixing of total water in cloud top entrainment
		Major convection changes (switching, initiation, entrainment)
		Increased precipitation efficiency of convection
Cy26r3	07 Oct 2003	HALO radiation sampling
		New aerosol climatology
		Relaxation of mass flux limiter for long time steps
Cy28r1	09 Mar 2004	Minor fixes to convection scheme
		New snow analysis, using NESDIS snow cover
Cy28r3	28 Sep 2004	Revised convection scheme numerics and calling of cloud scheme
		Hourly radiation (instead of 3-hourly)
		Improved numerics of surface tile coupling
Cy29r1	05 Apr 2005	New moist boundary-layer scheme
		Revision to tile coupling of snow cover
Cy30r1	01 Feb 2006	Minor bug fixes in convection
		Revised ozone chemistry scheme
		Increase of vertical resolution from 60 to 91 layers*
Cy31r1	12 Sep 2006	Revised cloud scheme, including treatment of ice supersaturation
		Implicit computation of convective transports
		Modified orographic drag
		Salinity effect on saturation at ocean surface
		Gust fix for orography

*Note that the increased vertical resolution introduced in the ECMWF operational forecasting system via Cy30r1 is not used in ERA-Interim.

to the model physics with substantial impact on reanalysis quality are described in the following sections (also Beljaars *et al.*, 2006). Complete technical documentation of the IFS can be found at http://www.ecmwf.int/research/ifsdocs

3.1. Atmospheric model

The dynamical core of the atmospheric model is based on a spectral representation for the basic dynamical variables, a hybrid sigma-pressure vertical coordinate, and a semi-Lagrangean semi-implicit time stepping scheme. The ERA-Interim configuration uses a 30 min time step and has a spectral T255 horizontal resolution (compared to T159 for ERA-40), which corresponds to approximately 79 km spacing on a reduced Gaussian grid (125 km for ERA-40). The vertical resolution is unchanged, using 60 model layers with the top of the atmosphere located at 0.1 hPa.

3.1.1. Clouds and convection

Several modifications to the model physics were introduced with potentially significant impact on the representation of the hydrological cycle. The cloud scheme was substantially revised in Cy25r3. Changes were made to the formulation of ice sedimentation, cloud-top entrainment, cumulus subsidence, conversion of supersaturated profiles, cloud erosion, and the numerics of the cloud scheme. At the same time, changes to the convection scheme were implemented to improve the distinction of shallow, deep and mid-level convection, computation of the cloud base, modelling of cumulus rising parcels, night-time onset of convection over land, convective precipitation efficiency, and the initiation of cloud-base winds.

The main impact of these changes was to increase the activity of the convection scheme (e.g. more convective precipitation by allowing it to trigger at night). As a consequence the model atmosphere is less unstable, has less vertical motion and the large-scale precipitation scheme produces less precipitation. Figure 4 shows a reduction in 12 h forecast errors in 200 hPa height for May 2002 due to increased heating from convection, affecting forecast skill downstream over the Atlantic and Europe. Improvements also included a reduction in tropical wind errors, and better phasing of precipitation events (Bechtold *et al.*, 2004).

Further revisions to the cloud scheme were introduced in Cy31r1. Most importantly, a new parametrization was implemented to allow supersaturation with respect to ice in the cloud-free part of a grid box at temperatures lower than 250 K (Tompkins *et al.*, 2007). As expected, this led to a substantial increase of relative humidity in the upper

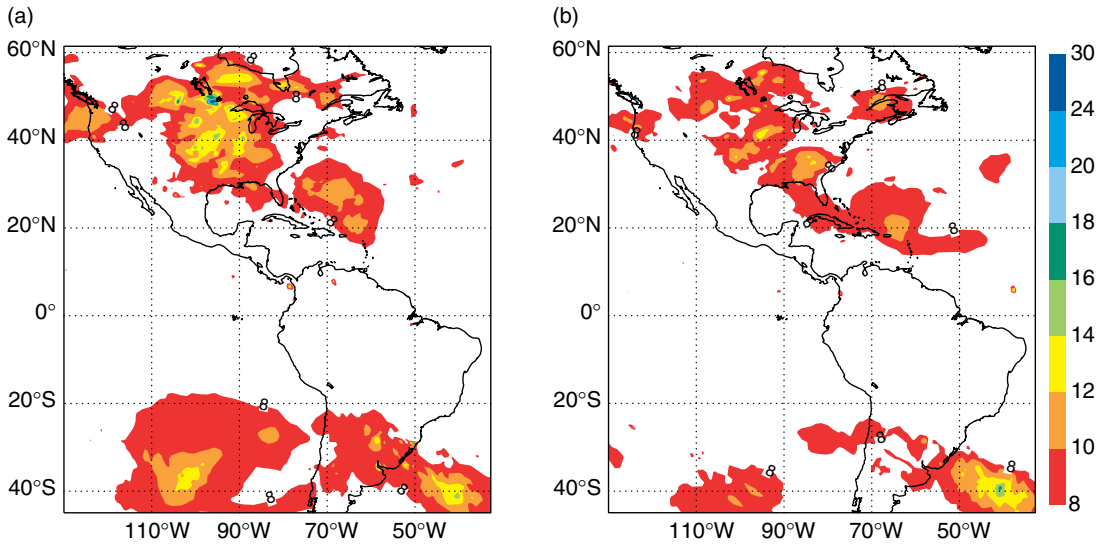

Figure 4. Average 12-hour forecast errors of geopotential height at 200 hPa, from daily forecasts in May 2002, using IFS (a) Cy25r1 and (b) Cy25r4 (after convection and cloud changes). Forecasts are verified against analyses produced by the same system.

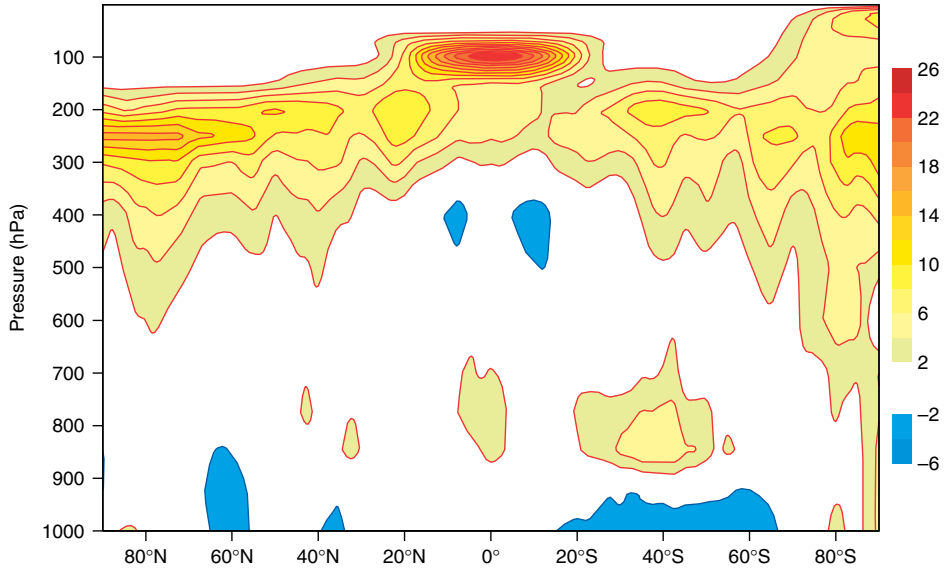

Figure 5. Zonally averaged differences in relative humidity (%) in IFS Cy31r1 model climates with and without supersaturation with respect to ice. Relative humidity is defined with respect to water when $T > 0\,^{\circ}\mathrm{C}$, ice when $T < -23\,^{\circ}\mathrm{C}$, and a mixed phase function for temperatures in between.

troposphere (Figure 5). Verification of these changes is difficult because radiosonde humidity measurements are known to be biased at high altitudes. However, studies based on research data had suggested that the ECMWF model and analysis fields were too dry in these parts of the atmosphere (Tompkins *et al.*, 2007). As a result of the introduction of supersaturation, which was not allowed in ERA-40, relative humidity in the stratosphere has increased over the southern polar cap, for reasons that are not fully understood.

3.1.2. Moist boundary layer

A moist boundary-layer scheme based on eddy diffusion combined with mass-flux transport (Köhler, 2005, Köhler *et al.*, 2011) was introduced in Cy29r1. Stratus and stratocumulus had been systematically underpredicted in previous versions of the model, mostly due to insufficient moist mixing in the boundary layer. The new scheme uses the moist conserved variables liquid water static energy (which reduces to dry static energy in dry conditions) and total water (vapour + condensed water). The mixed-layer height is determined using an entraining parcel, selecting the top of stratocumulus, or cloud base in shallow convection situations. The distinction between stratocumulus and shallow convection is based on inversion strength (Klein and Hartman, 1993). A similarity profile of diffusion coefficients is prescribed over the depth of the boundary layer. The diffusion coefficient profile consists of a surface-driven component that scales with surface fluxes (Troen and Mahrt, 1986; Holtslag, 1998) and a cloud-top-driven component that scales with cloud-top radiative cooling (Lock, 1988). Boundary-layer-top entrainment is explicitly prescribed in terms of buoyancy flux with a surface buoyancy component and a cloud-top radiative cooling component.

The mass flux term, which represents the effects of the large coherent eddies that mix throughout the boundary layer, simulates countergradient transports. An updraught model is used for the updraught properties and for vertical velocity (Siebesma and Teixeira, 2000). The latter is part

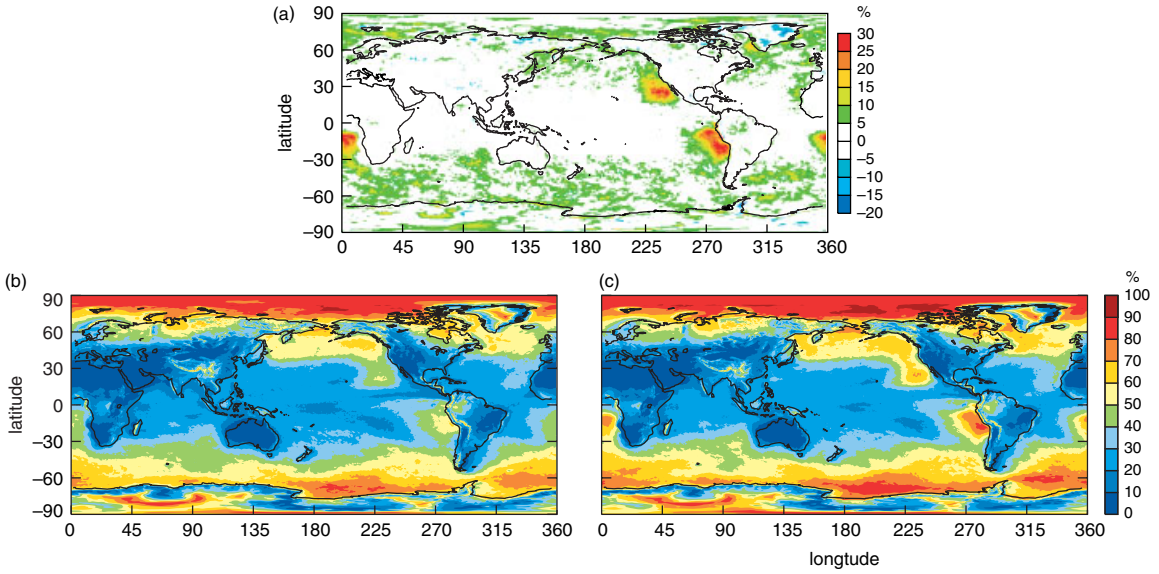

Figure 6. Model climate of low cloud cover from an ensemble of one year T159L60 integrations with (b) the dry boundary-layer scheme and (c) the new boundary-layer scheme. The difference between new and old schemes is shown in (a).

of the plume entrainment parametrization and is also used to locate the boundary-layer top. Full details of the implementation of the new scheme in the IFS are given in Köhler *et al.* (2011).

The new moist boundary-layer scheme has the desired effect of producing more stratocumulus in previously underpredicted areas. This is illustrated in Figure 6 which shows the impact of the new scheme on the model climate for low cloud cover. The stratocumulus areas located near the west coasts of North America, South America and Africa show a substantial increase in boundary-layer cloud cover. Liquid water content of the stratocumulus clouds appears more realistic (not shown). Köhler *et al.* (2011) show that for the EPIC experiment, the new boundary-layer scheme makes the inversion slightly sharper. However the model inversion is still about one model level (25 hPa) below the observed inversion at around 870 hPa.

3.1.3. *Surface drag*

The representation of orographic effects on the atmospheric flow was revised in Cy31r1 of the IFS. The 'effective roughness length' concept, in which the aerodynamic roughness length is enhanced due to drag of subgrid orographic features, was replaced by the TOFD scheme (Beljaars *et al.*, 2004). TOFD represents the drag due to orographic scales between 5000 m and 10 m and is implemented as a tendency profile on model levels. This has the advantage that it is independent of the surface boundary condition. The roughness length represents the effects of vegetation and land use only, which limits its magnitude to a few metres.

Other modifications related to surface effects introduced in Cy31r1 include: revision of the vegetation roughness length specification based on land-use information rather than climatology; exclusion of the blocked layer in the forcing of gravity waves by subgrid orography; changes to improve the numerical balance between dynamical forcing and near-surface drag; a modification to the atmospheric boundary condition for moisture at the ocean surface to account for salinity effects.

The new TOFD scheme, combined with modified surface roughness length characteristics, changes the drag over land in a complicated pattern. There is an increase in areas with tropical forest and deserts, which is due to the relation between roughness length and land use. Overall, the TOFD scheme leads to a reduction of the drag coefficient, although this is not the case everywhere because the geographical distribution of the effect is controlled by the standard deviation of small-scale orographic features. The latter are computed from the 30 arc-second GTOPO30 orographic dataset, using scales between 2 and 20 km. Wind speed at 10 m shows a general increase, which is consistent with the introduction of TOFD and the reduction of vegetation roughness length in many locations (Figure 7). Wind speed verifies better with respect to SYNOP observations in mountainous regions (not shown).

3.1.4. *Impact of model physics changes on cloud cover*

Figure 8 shows differences between total and low cloud cover in ERA-Interim and ERA-40. Using ISCCP observations as a reference, the following features have improved:

- Marine stratocumulus cloud cover increases by 15–25% due to the implementation of the new moist boundary-layer scheme (Cy29r1) (Köhler, 2005; Köhler *et al.*, 2011).
- Tropical ocean total cloud cover decreases by 5–15%. This is the result of an overall improved hydrological cycle resulting in a drier lower troposphere, and the introduction of ice supersaturation which delays the formation of ice clouds (Tompkins *et al.*, 2007).
- Tropical land cloud cover increases by 20–30%. This impact is due to a 10–25% increase in high cloud resulting from improved deep convective triggering (Cy25r3) (Bechtold *et al.*, 2004), combined with additional low cloud produced by the new boundary-layer scheme.
- Midlatitude ocean (low- and medium-height) cloud cover increases by about 5% due to improved numerical solution of the cloud production and decay equations (Cy31r1).

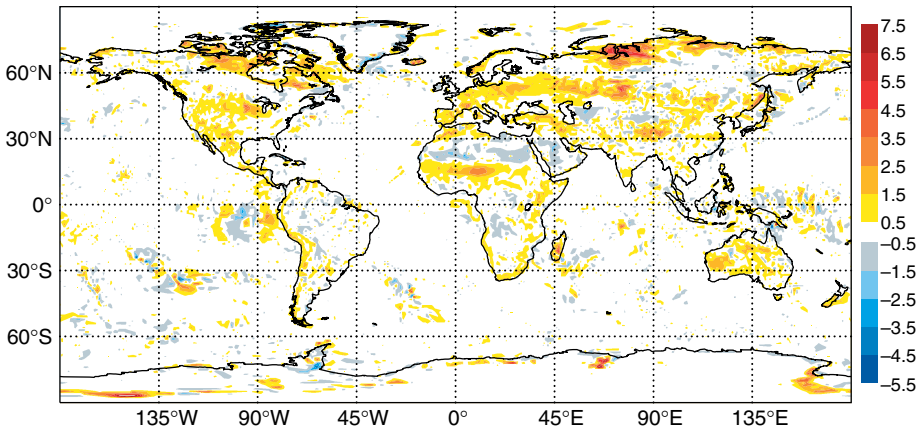

Figure 7. Difference in 10 m wind speed (m s^{-1}) between simulations with the TOFD scheme (including new roughness length related to land use) and the control model (using effective roughness). Results shown are for the period 10 March 2004–10 April 2004, based on daily 12–24 h forecast segments for the drag coefficient and 24 h forecasts for the 10 m wind.

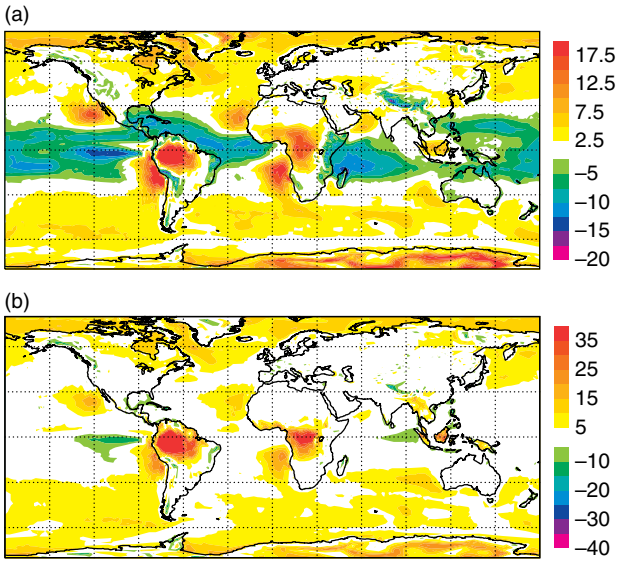

Figure 8. Mean differences (%) in (a) total cloud cover and (b) low cloud cover between ERA-Interim and ERA-40, based on 30 h and 36 h forecasts issued twice daily during the period 1990–2001.

3.1.5. *Prognostic ozone*

Dethof and Hólm (2004) provided a detailed description of the prognostic ozone model used for ERA-40, which was based on an updated version of the Cariolle and Déqué (1986) scheme and still forms the basis for the model used in ERA-Interim. The ozone continuity equation is expressed as a linear relaxation towards a photochemical equilibrium for the local value of the ozone mixing ratio, the temperature, and the overhead ozone column. An ozone destruction term is used to parametrize the heterogeneous chemistry as a function of the equivalent chlorine content for the actual year. Several recent upgrades of the chemistry parametrization scheme (Cariolle and Teyssèdre, 2007) were included in the ERA-Interim system.

3.1.6. *Incoming solar radiation*

For solar irradiance, ERA-Interim uses a constant value of 1370 W m^{-2} throughout, i.e. no account is taken of the solar cycle. Variations due to the varying distance between the Earth and the Sun are incorporated as described in Paltridge and Platt (1976).

Due to a programming error in the calculation of incident solar radiation as a function of solar zenith angle, the global solar radiation in ERA-Interim is overestimated by about 2 W m^{-2}. This error, which has been present in the ECMWF forecast model for a very long time, has also affected ERA-40. Model simulations have shown that the systematic impact of this error on the model climate is confined to the upper stratosphere, which is warmed by approximately 1 K globally.

3.1.7. *Radiative transfer*

Solar radiation is attenuated by absorbing gases, consisting of water vapour, uniformly mixed gases (oxygen, carbon dioxide, methane, nitrous oxide) and ozone, and scattered by molecules (Rayleigh scattering), aerosols and cloud particles. For water vapour, the radiation scheme uses prognostic information produced by the forecast model. For aerosols, carbon dioxide, trace gases, and ozone, climatological information is used instead. Note that prognostic ozone is available but not used in the radiation scheme; fully activating the coupling between ozone and radiation in the forecast model has been found to increase temperature errors in the stratosphere.

The effect of aerosols on radiative transfer in the atmosphere is modelled based on prescribed climatological aerosol distributions. ERA-40 used an aerosol climatology originally developed by Tanré *et al.* (1984), which provides annual mean geographical distributions for maritime, continental, urban and desert aerosol types, in addition to a uniformly distributed tropospheric and stratospheric 'background' aerosol loading. A new climatology for the annual cycle of various tropospheric aerosol types (sea-salt, dust, organic and black carbon, sulphate) was introduced in Cy26r3, derived from Tegen *et al.* (1997) based on simulation by chemical transport models. It describes the annual cycle in terms of monthly mean aerosol optical depth distributions.

Well-mixed (vertically and horizontally) tropospheric background aerosols with an optical thickness of 0.03 and stratospheric background aerosols with an optical thickness of 0.045 are added to the previous amounts, with a rate of change of optical thickness with pressure of 3.7×10^{-7}Pa^{-1} and 0.23×10^{-7}Pa^{-1}, respectively. There is no evolution of volcanic aerosols. Therefore, the radiation scheme does not

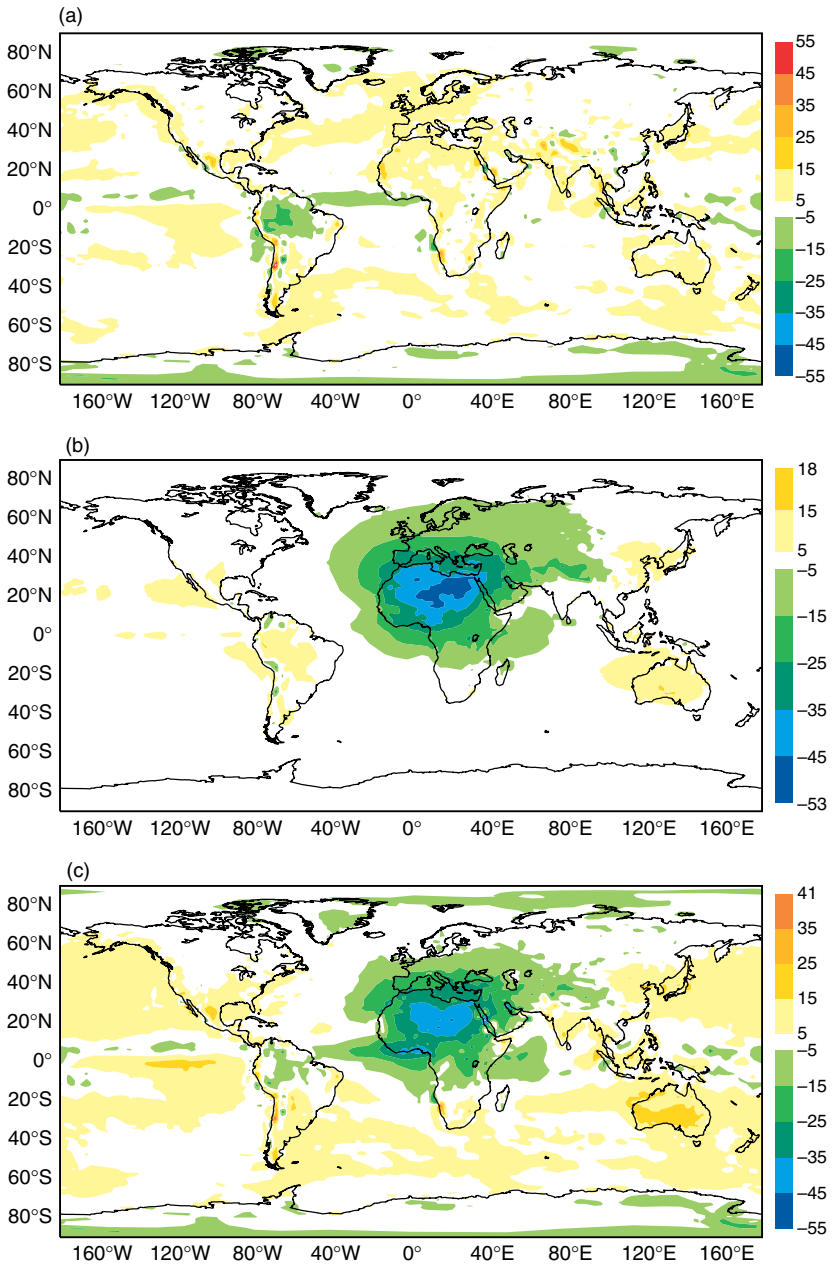

Figure 9. Difference in (a) long-wave, (b) short-wave and (c) total radiative divergence (W m^{-2}) between ERA-Interim and ERA-40, averaged for 1993.

account for changes in stratospheric sulphate as induced, for example, by the Pinatubo eruption in 1991.

Relevant radiative properties (extinction coefficient, single scattering albedo, asymmetry factor) are consistently derived for each aerosol type and for the various spectral intervals used in the radiation scheme, following Hess *et al.* (1998). The number of intervals for short-wave radiation was increased to six in Cy25r1, and both the spatial sampling and temporal frequency of radiative transfer computations were increased in Cy26r3.

Tompkins *et al.* (2005) have demonstrated that the direct radiative effect of the new aerosol climatology has improved simulations of the African Easterly jet. Rodwell (2005) and Rodwell and Jung (2008) further evaluated the impact of the new climatology on features of the general circulation, both near and far from the location of the main changes in aerosol optical depth, which is centred over the Sahara. They found improvements in the representation of the North African monsoon, in tropical precipitation, and a reduction in mean extratropical circulation errors.

The impact of the new climatology on long-wave and short-wave radiative divergence (i.e. the imbalance between the net fluxes at the top of the atmosphere and surface) was not explicitly addressed in these studies. Figure 9 presents differences in long-wave, short-wave and total radiative divergence between ERA-Interim and ERA-40, averaged for the year 1993. The two reanalyses differ in the details of the radiation schemes used, the horizontal and vertical distributions of cloud, and in the optical properties of aerosols. However, the main signal in the short-wave radiative divergence clearly visible over North Africa is a direct result of the change in aerosol climatology.

The greenhouse gases carbon dioxide, methane, nitrous oxide, CFC-11 and CFC-12 are assumed to be globally well-mixed. The concentrations for these gases are set to observed 1990 values plus a linear trend as specified in the IPCC Second Assessment Report (IPCC, 1996).

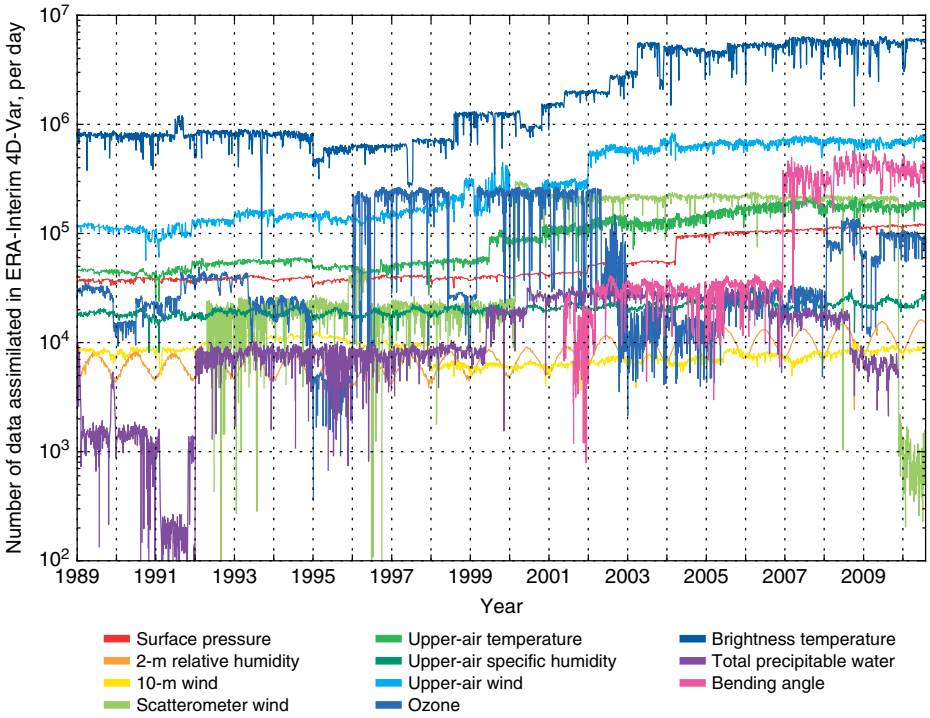

Figure 10. Daily counts, on a logarithmic scale, of observations assimilated in the atmospheric analysis component of ERA-Interim.

The ozone climatology used in the radiation scheme distributes the ozone mixing ratio as a function of pressure, latitude and month following Fortuin and Langematz (1994).

3.2. *Land surface*

Apart from minor changes as listed in Table II, the land-surface component of the forecast model used in ERA-Interim is essentially identical to that in ERA-40. It uses the TESSEL scheme (Viterbo and Beljaars, 1995; Viterbo *et al.*, 1999) to evolve the thermal and water storage in four layers of soil and snow during the forecast. Within a model grid box, the coverage of each tile depends on the dominant type and area fraction, for high and for low vegetation. These quantities, which are kept fixed in time, are based on the GLCC dataset derived from one year of AVHRR data and ancillary information (Loveland *et al.*, 2000). The snow scheme is based on Douville *et al.* (1995), with separate treatment of open-area snow and snow shielded by high vegetation (Viterbo and Betts, 1999).

The newly introduced dependence of vegetation roughness length on the dominant high vegetation type, already mentioned in section 3.1.3, has affected the behaviour of the land-surface model. An increase in vegetation roughness increases aerodynamic resistance, and also reduces winter evaporation and sublimation from snow (Beljaars and Viterbo, 1994).

3.3. *Ocean waves*

The wave-model component of the forecast model represents the impact of ocean waves on airflow via transfer of energy and momentum across the interface. This is achieved by a two-way coupling, passing wind fields and other atmospheric parameters that influence wave growth to the wave model, and returning information about the impact of the sea state on surface roughness via the Charnock parameter (Janssen, 2004).

The wave model incorporated in the IFS is based on the WAM approach (Komen *et al.*, 1994). The version used in ERA-Interim includes several enhancements, both in physics and numerics, over the version that was used in ERA-40 (Janssen *et al.*, 2005; Janssen, 2008). The most significant for climate applications are the introduction of a scheme for treating unresolved bathymetry effects and a reformulation of the dissipation source term (Bidlot *et al.*, 2007).

The horizontal resolution of the wave model in ERA-Interim is 110 km; wave spectra are discretised using 24 directions and 30 frequencies.

3.4. *Sea-surface temperature and sea-ice concentration*

Global estimates of SST and SIC are required as boundary conditions for the atmospheric forecast model. The IFS does not incorporate its own analysis of these fields, but rather relies on estimates produced elsewhere. These are then aggregated/interpolated to the ECMWF reduced-Gaussian model grid as needed. For dates prior to 2002, ERA-Interim used the same SST and SIC input data used for ERA-40, as described by Fiorino (2004). Starting in January 2002, a switch was made to data used in the ECMWF operational forecasting system, beginning with the daily operational NCEP product and most recently from the OSTIA (Stark *et al.*, 2007).

4. Observations

The number of observations assimilated in ERA-Interim has increased from approximately 10^6 per day on average in 1989, to nearly 10^7 per day in 2010. Figure 10 shows, on a logarithmic scale, the daily counts for all observations used in the atmospheric 4D-Var analysis. The overwhelming majority of data, and most of the increase over time, originate from satellites. This includes clear-sky radiance

measurements (quantified as brightness temperatures) from polar-orbiting and geostationary sounders and imagers, atmospheric motion vectors derived from geostationary satellites, scatterometer wind data, and ozone retrievals from various satellite-borne sensors. Also derived from satellite observations are the total precipitable vapour estimates produced within the 1D+4D-Var scheme described in section 2.1.6. Measurements of atmospheric refraction (quantified as bending angles) obtained from GPS radio occultation began to be used in ERA-Interim in 2001, growing to significant numbers by the end of 2006.

The conventional observing system, in spite of much lower data volumes, still serves as an indispensable constraint to the atmospheric reanalysis. *In situ* measurements of upper-air temperatures (T), wind (u/v), and specific humidity (q) were available from radiosondes, pilot balloons, aircraft, and wind profilers. Data counts for these sources are more or less steady during the reanalysis period, with the exception of aircraft reports whose numbers increased greatly after 1998. Observations of surface pressure (Ps), 2 m temperature, 2 m relative humidity (RH), and near-surface (10 m) winds (u/v) from ships, drifting buoys, and land stations were also assimilated in steady numbers.

4.1. *Input datasets*

Observations assimilated in ERA-Interim for all dates prior to 2002 consist mainly of input data originally prepared for ERA-40. These data and their sources are described in Uppala *et al.* (2005). The most notable differences in their use in ERA-Interim are:

- All satellite clear-sky radiance data used in ERA-Interim are subject to variational bias correction, as described in section 2.1.2;
- Clear-sky radiances from SSM/I, used as 1D-Var retrievals of TCWV and surface wind speed in ERA-40, are now directly assimilated in ERA-Interim;
- Rain-affected SSM/I radiances, which were not used in ERA-40, are assimilated in ERA-Interim using the 1D+4D-Var approach described in section 2.1.6;
- Surface pressure observations are subject to an automated bias-correction scheme (section 4.3.1);
- ERA-Interim uses newly derived radiosonde temperature bias adjustments (section 4.3.2);
- Ozone data usage in ERA-Interim has been considerably expanded (section 4.5.1);
- Additional scatterometer ocean surface wind data were used in ERA-Interim, including recalibrated data from ERS-1 and ERS-2, and data from QuikSCAT (section 4.5.2).

Various other changes in data usage relative to ERA-40 are described below.

For dates after 2002, ERA-Interim has used observations from ECMWF's operational archive, which were originally received on the GTS. The main differences in data usage from the ECMWF operational forecasting system include:

- Data from TMI, IASI, and ASCAT used in the operational forecasting system have not been assimilated in ERA-Interim, due to technical limitations in the IFS version used;
- Ozone data from MLS-Aura, not used for forecasting, have been assimilated in ERA-Interim;
- Radiance data from AIRS, used for forecasting since October 2003, have been assimilated in ERA-Interim since April 2003 (section 4.4.4).

The following datasets were acquired especially for ERA-Interim:

- Reprocessed altimeter wave-height data from ERS-1 and ERS-2, acquired from ESA, were calibrated using buoy measurements, and then merged into a single consistent dataset (section 4.6.2);
- AMV wind data from geostationary satellites reprocessed by EUMETSAT, from Meteosat-3 (January–-May 1995), Meteosat-4 (January–December 1993), Meteosat-5 (January 1995 –February 1997), and Meteosat-6 (February–December 1997) (Delsol *et al.*, 2008);
- Reprocessed ozone profiles from the GOME instrument on ERS-2, provided by the Rutherford Appleton Laboratory (Siddans *et al.*, 2002) for the period January 1996 –December 2002 (Dragani, 2010a);
- Reprocessed CHAMP GPS radio occultation data obtained from UCAR, used in ERA-Interim beginning May 2001 (section 4.5.3).

The ERA-Interim reanalysis continues to be updated in near-real time, closely following data usage of the ECMWF operational forecasting system.

4.2. *Quality control and data selection*

All observations used in ERA-Interim are subject to a suite of quality control and data selection steps. Various preliminary checks serve to detect errors that can occur when measurements are recorded or transmitted. These include checks for completeness of reports, physical feasibility, integrity of ship routes and aircraft flight tracks, hydrostatic consistency of radiosonde profiles, and the occurrence of duplicate reports. Observations that fail any of these checks are flagged for exclusion from further analysis. Quality information generated prior to and during the analysis, along with data departures, are stored with the observations and can be made available for later investigation.

Depending on the type of observation, a thinning procedure may be applied to reduce data density to a level commensurate with the resolution supported by the analysis method. Thinning of satellite observations, including high-density AMVs and scatterometer wind data as well as radiances, also serves to prevent effects of spatially correlated observation errors that are not explicitly accounted for by the analysis method. Aircraft reports are thinned along flight tracks and vertically during ascent and descent, and data from wind profilers are similarly thinned in the vertical direction.

Data selection rules specified in so-called blacklists are applied in order to exclude observations that are expected to have a negative impact on the reanalysis. The selection may be based on prior knowledge about instrument performance, for example, to exclude station data that are known to be unreliable, or to remove observations from failing satellite sensors. Data may also be excluded because they cannot be usefully interpreted by the data assimilation system. This is the case for certain types of weather observations (e.g. surface wind over land, visibility) and for many satellite measurements that are not accurately

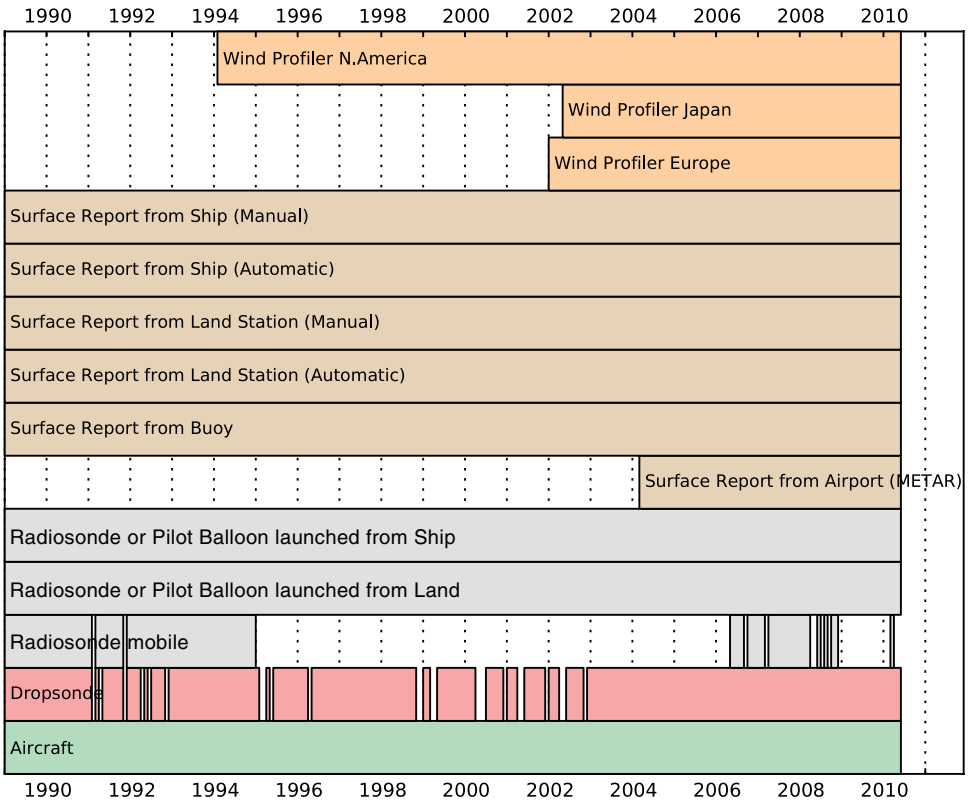

Figure 11. Timeline of conventional observations assimilated in ERA-Interim.

represented by the forecast model (e.g. due to effects of cloud, rain, or land-surface properties).

Blacklist information is vitally important to the quality of the analysis, and requires continuous maintenance and reassessment to reflect the accumulated experience from previous use and analysis of the observations. A blacklist is often used to exclude observations that have not previously been analysed; such data may be passed through the assimilation system to be monitored and evaluated in the context of all other available observations. Quality feedback information generated in this manner can be used to develop quality-control and bias-correction procedures for the new data, which may ultimately lead to their useful assimilation.

Subsequent quality-control tests involve integrated information about the likely state of the atmosphere, as estimated by the data assimilation system. These tests are essentially statistical in nature. They aim to exclude (or reduce the influence of) observations that are considered implausible, given all other information available to the system, including assumptions about natural uncertainties in both the background estimates and the observations themselves.

The first of these tests is the so-called background check, which is applied to all observations that pass the preliminary screening and thinning steps, including those that are blacklisted. The background check relies on the ability to predict each individual observation, using the forecast model together with the appropriate observation operator (section 2.1). This test eliminates any observation whose departure from the background exceeds a prescribed threshold, which is proportional to the expected departure based on error statistics for background and observations. The proportionality factor generally depends on the reliability of the instrument, as well as on a judgement of risk associated with assimilating outlier observations of a given type. For example, criteria for selecting radiance data from infrared sensors can be quite strict, in order to avoid contamination of the temperature analysis by assimilating cloud-affected radiance data from these sensors. On the other hand, the background check must be generous enough to prevent discarding observations with genuine information about unexpected events.

The final test for all remaining observations is the variational quality control (Andersson and Järvinen, 1999), which is embedded in the iterative minimisation of the 4D-Var cost function (1). The variational quality control effectively reduces the influence of outlier observations as the minimisation converges. Whether a particular observation becomes an outlier during convergence depends on the degree of consistency with other nearby observations. This test can therefore be regarded as a refinement of the background check, as it takes into account additional information from newly available observations as they are being assimilated.

4.3. *Use of conventional data*

The timeline in Figure 11 summarises the sources of *in situ* observations used in ERA-Interim. Surface observations from land stations (Ps, RH), ships (Ps, u/v at 10 m), and drifting buoys (Ps, u/v at 10 m) were used throughout the reanalysis period, as were reports from radiosondes (T, u/v, q) and pilot balloons (u, v) launched from land stations and ships, dropsondes (T, u/v, q), and aircraft reports (T, u/v). Wind profiler data (u/v) from North American sites became available in 1994, with European and Japanese profilers added in 2002. Hourly METAR airport weather reports (Ps) began being used in 2004. Pseudo-surface pressure observations, used in ERA-40, were not used in ERA-Interim.

Detailed data selection rules for conventional observations are specified in blacklists, which exclude data from land stations, and some buoys and ships that are known to be

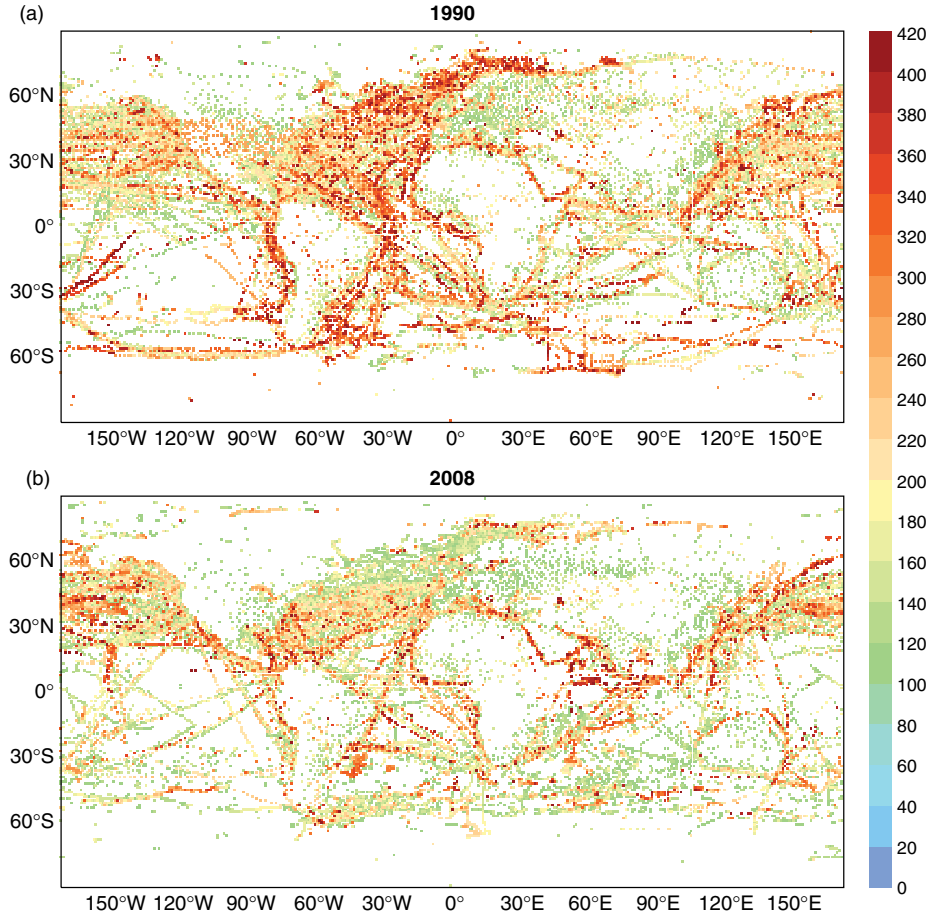

Figure 12. Annual RMS bias corrections (Pa) for surface pressure observations, averaged for (a) 1990 and (b) 2008.

problematic. Many of these selection rules depend on time. In addition, the following data are always excluded:

- All near-surface wind observations over land;
- Surface pressure observations over high terrain;
- Radiosonde observations below the model surface;
- Near-surface relative humidity observations at night-time or over high terrain;
- Radiosonde specific humidity observations in extreme cold conditions ($T < 193$ K for RS-90 sondes, $T < 213$ K for RS-80 sondes, $T < 233$ K otherwise);
- Radiosonde specific humidity observations at high altitude ($p < 100$ hPa for RS-80 and RS-90 sondes, $p < 300$ hPa for all other sonde types).

High terrain typically means elevations higher than 1500 m, but this can vary depending on the instrument. Station-dependent data selection rules for wind profilers, e.g. to exclude low-altitude observations in coastal regions, were inadvertently omitted from the ERA-Interm blacklists prior to November 2007.

To illustrate the impact of quality-control decisions on data usage, Tables III and IV show cumulative data counts and rejection rates for all conventional observations ingested in ERA-Interim during the years 1995 and 2005, respectively. Observation types and observed parameters are indicated in the first two columns. For upper-air observations, the counts and statistics are approximately separated into near-surface, tropospheric and stratospheric categories. For comparison, the tables include statistics for ERA-40 in 1995 and for the ECMWF operational forecasting system in 2005. The thresholds used for the background check are shown, as are the outlier limits used to define the variational quality control (Andersson and Järvinen, 1999). All percentages are relative to the total data counts listed in the left-most numeric column. Data reductions due to blacklisting and thinning (not shown) account for the fact that rates do not add to 100%. Observations subject to variational quality control are counted as rejected if they are downweighted by more than 75% during the variational analysis.

A more detailed discussion of conventional data usage in ERA-40, ERA-Interim, and the ECMWF operational forecasting system is presented by Tavolato and Isaksen (2010).

4.3.1. *Surface pressure bias correction*

Surface pressure observations are adjusted for systematic errors by means of a simple adaptive updating scheme, which is described in detail by Vasiljevic (2006) and Vasiljevic *et al.* (2006). Biases are assumed local and independent for each station, ship, or buoy, and are constrained to vary slowly in time. At the beginning of each analysis cycle, a bias estimate is obtained at each observation location by averaging the departures at that location over the most recent 30-day period. A statistical test is then performed to check whether the bias estimate is likely independent of estimates at nearby locations. Since spatially coherent departures may be indicative of large-scale model errors, the scheme will not apply a bias adjustment to the observed surface pressure value in such cases.

Figure 12 shows RMS bias corrections applied during the years 1990 and 2008. Globally averaged values are 262 Pa

Table III. Table of data counts and quality control statistics for conventional data in ERA-40 (E4) and ERA-Interim (EI) for the year 1995.

			All observations		FG rej (%)		VarQC rej (%)		Used obs (%)		BG QC limits		Var QC limits		Units	Rejection ratio (%)	
Obs type	Obs value	Level	E4	EI	E4	EI	E4	EI	E4	EI	E4	EI	E4	EI		E4	EI
SYNOP	*Ps*	surf	11 168 251	18 041 532	1.85	1.25	1.02	0.25	86.15	53.61	320.0	240.0	var	200.0	Pa	2.87	1.50
SYNOP	*RH*	surf	11 021 987	11 045 165	0.02	0.11	0.03	0.00	30.67	18.16	52.0	n.a.	28.0	n.a.	%	0.05	0.11
SHIP	*Ps*	surf	2 418 295	2 412 324	1.12	1.76	0.70	0.88	78.00	79.79	380.0	260.0	var	200.0	Pa	1.82	2.64
SHIP	*u*/*v* (10 m)	surf	2 633 672	2 628 644	0.57	0.67	0.10	0.30	41.62	50.42	11.6	11.4	10.8	10.8	m s^{-1}	0.67	0.97
DRIBU	*Ps*	surf	2 183 944	3 075 928	1.88	2.68	1.14	0.29	64.52	50.81	380.0	320.0	var	240.0	Pa	3.02	2.97
DRIBU	*u*/*v* (10 m)	surf	1 419 484	1 424 686	0.46	0.43	0.25	0.52	43.11	48.84	10.9	10.7	7.4	7.4	m s^{-1}	0.71	0.95
TEMP	*T*	all	14 173 985	14 101 291	1.20	1.06	0.74	1.00	93.04	93.37	3.5	3.0	2.5	2.5	K	1.95	2.06
TEMP	*T*	0–100	3 341 643	3 315 190	1.10	1.52	1.43	1.53	94.07	94.05	4.8	4.7	3.6	3.6	K	2.53	3.06
TEMP	*T*	100–900	9 617 377	9 574 844	1.02	0.90	0.52	0.71	93.98	94.06	3.5	3.0	2.5	2.5	K	1.54	1.61
TEMP	*T*	900–1000	1 214 961	1 211 253	2.95	1.10	0.60	1.77	82.74	86.02	3.9	4.7	3.6	3.6	K	3.55	2.87
TEMP	*u*/*v*	all	10 821 439	10 751 160	0.71	0.70	0.33	0.63	91.60	91.24	10.5	11.0	9.1	9.1	m s^{-1}	1.04	1.34
TEMP	*u*/*v*	0–100	2 789 733	2 768 702	1.06	1.00	0.55	0.74	96.26	95.67	15.7	13.0	10.2	10.1	m s^{-1}	1.60	1.75
TEMP	*u*/*v*	100–900	6 756 365	6 710 221	0.61	0.62	0.26	0.59	95.77	95.27	11.7	11.5	9.1	9.1	m s^{-1}	0.87	1.21
TEMP	*u*/*v*	900–1000	1 275 243	1 272 139	0.48	0.49	0.24	0.62	59.34	60.39	10.5	11.0	9.1	9.1	m s^{-1}	0.73	1.11
TEMP	*q*	all	11 519 732	11 477 986	0.71	0.46	0.06	0.17	58.91	58.62	varying		varying		–	0.77	0.63
TEMP	*q*	100–900	8 216 284	8 182 711	0.67	0.54	0.08	0.21	70.92	69.97	varying		varying		–	0.75	0.75
TEMP	*q*	900–1000	1 624 186	1 621 401	1.61	0.53	0.05	0.16	59.08	61.91	varying		varying		–	1.67	0.70
AIREP	*T*	all	12 023 988	12 066 985	1.21	1.18	0.09	0.08	41.94	39.56	4.6	4.2	3.7	3.8	K	1.31	1.26
AIREP	*T*	100–900	10 495 862	10 533 445	1.20	1.24	0.11	0.08	45.14	40.90	4.6	4.2	3.7	3.8	K	1.30	1.32
AIREP	*T*	900–1000	1 525 185	1 530 593	1.30	0.74	0.02	0.06	19.92	30.33	5.5	6.2	5.1	5.2	K	1.32	0.79
AIREP	*u*/*v*	all	10 410 626	10 449 261	1.26	1.36	0.16	0.16	48.44	44.00	~ 15.0	~ 14.5	12.7	12.8	m s^{-1}	1.43	1.52
AIREP	*u*/*v*	100–900	9 048 325	9 082 108	1.38	1.49	0.19	0.17	52.37	45.75	~ 15.5	~ 15.5	12.8	12.8	m s^{-1}	1.57	1.66
AIREP	*u*/*v*	900–1000	1 359 187	1 364 033	0.46	0.48	0.02	0.02	22.30	32.37	~ 15.0	~ 14.5	11.7	12.8	m s^{-1}	0.48	0.50
PILOT	*u*/*v*	all	3 928 818	3 939 870	0.68	0.81	0.30	0.63	84.67	82.55	10.7	10.4	9.1	9.1	m s^{-1}	0.98	1.44
PILOT	*u*/*v*	0–100	447 563	450 149	1.07	1.63	0.36	0.60	63.10	61.19	15.7	13.0	10.5	10.2	m s^{-1}	1.44	2.23
PILOT	*u*/*v*	100–900	2 856 343	2 863 414	0.62	0.72	0.30	0.65	89.14	86.66	12.1	11.0	9.1	9.1	m s^{-1}	0.92	1.37
PILOT	*u*/*v*	900–1000	624 904	626 303	0.65	0.65	0.24	0.56	79.69	79.11	10.7	10.4	9.1	9.1	m s^{-1}	0.89	1.21
US profil	*u*/*v*	all	1 497 290	1 502 501	2.51	2.80	0.34	0.61	96.62	95.91	12.9	11.8	9.2	9.1	m s^{-1}	2.85	3.41
US profil	*u*/*v*	0–100	20 978	21 354	12.68	13.49	1.23	1.39	86.00	85.03	18.9	16.8	11.2	10.5	m s^{-1}	13.91	14.88
US profil	*u*/*v*	100–900	1 440 230	1 444 911	2.39	2.66	0.32	0.57	96.74	96.07	12.9	11.8	9.3	9.1	m s^{-1}	2.71	3.23
US profil	*u*/*v*	900–1000	36 082	36 236	1.60	2.02	0.45	1.74	97.72	95.95	12.9	12.4	9.2	5.5	m s^{-1}	2.05	3.75

'Obs type' definitions are given in the Appendix. 'FG rej' = rejection rate for the background check; 'VarQC rej' = rejection rate for variational QC. 'Bg QC limits' = QC limits used by the background check; 'VarQC limits' = outlier limits used by the variational QC. 'Units' apply to the previous four columns. 'Rejection ratio' represents the fraction of data removed by either the background check or the variational QC. All percentages are relative to the total data counts listed in 'All observations' columns.

Table IV. As Table III, but for the year 2005 and comparing with the ECMWF operational forecasting system (Oper), based on observations assimilated between 5 April 2005 and 4 April 2006.

			All observations		FG rej (%)		VarQC rej (%)		Used obs (%)		BG QC limits		Var QC limits		Units	Rejection ratio (%)	
Obs type	Obs value	Level	Oper	EI	Oper	EI	Oper	EI	Oper	EI	Oper	EI	Oper	EI		Oper	EI
SYNOP	*Ps*	surf	38 193 760	37 237 480	0.68	0.795	0.13	0.20	47.05	20.87	260.0	260.0	200.0	200.0	Pa	0.81	0.99
SYNOP	*RH*	surf	21 919 122	21 540 624	0.08	0.08	0.00	0.00	21.74	17.55	n.a.	n.a.	n.a.	n.a.	%	0.08	0.08
SHIP	*Ps*	surf	2 404 746	2 376 857	0.77	1.18	0.69	0.94	88.87	91.87	280.0	280.0	200.0	200.0	Pa	1.46	2.12
SHIP	*u*/*v* (10 m)	surf	2 327 477	2 297 341	0.64	0.59	0.32	0.28	59.84	46.31	10.8	11.6	10.8	10.8	m s^{-1}	0.96	0.87
DRIBU	*Ps*	surf	7 078 349	6 349 051	1.81	2.26	0.29	0.31	38.81	40.42	340.0	340.0	240.0	240.0	Pa	2.10	2.57
DRIBU	*u*/*v* (10 m)	surf	419 727	396 648	1.76	1.68	1.70	1.56	86.30	80.93	10.3	9.9	7.4	7.4	m s^{-1}	3.46	3.24
METAR	*Ps*	surf	155 551 667	15 453 014	0.06	0.23	0.04	0.16	89.39	91.02	380.0	360.0	340.0	340.0	Pa	0.10	0.38
TEMP	*T*	all	16 042 720	15 940 932	0.84	0.82	0.92	1.05	91.50	93.51	2.9	3.0	2.5	2.5	K	1.76	1.87
TEMP	*T*	0 100	4 118 642	4 087 865	1.37	1.08	1.50	1.48	90.34	95.37	4.4	4.6	3.6	3.6	K	2.87	2.56
TEMP	*T*	100–900	10 667 340	10 602 075	0.62	0.68	0.65	0.78	92.28	93.77	2.9	3.0	2.5	2.5	K	1.27	1.46
TEMP	*T*	900–1000	1 256 738	1 250 990	1.01	1.12	1.23	1.94	88.73	85.17	4.6	4.9	3.6	3.6	K	2.24	3.06
TEMP	*u*/*v*	all	14 570 453	14 340 834	0.47	0.54	0.42	0.60	93.50	92.77	9.7	10.8	9.1	9.1	m s^{-1}	0.89	1.14
TEMP	*u*/*v*	0–100	4 708 088	4 589 855	0.58	0.58	0.49	0.61	96.45	96.18	11.9	12.9	10.1	10.1	m s^{-1}	1.07	1.19
TEMP	*u*/*v*	100–900	8 534 351	8 430 361	0.41	0.51	0.37	0.57	95.79	95.23	10.7	11.4	9.1	9.1	m s^{-1}	0.78	1.08
TEMP	*u*/*v*	900–1000	1 327 990	1 320 602	0.49	0.57	0.53	0.78	68.30	65.25	9.7	10.8	9.1	9.1	m s^{-1}	1.02	1.35
TEMP	*q*	all	13 533 114	13 418 561	0.47	0.49	0.16	0.19	59.07	59.88	varying		varying		–	0.64	0.68
TEMP	*q*	100–900	9 364 838	9 280 919	0.58	0.59	0.21	0.24	73.54	75.29	varying		varying		–	0.79	0.83
TEMP	*q*	900–1000	1 668 555	1 660 797	0.58	0.62	0.17	0.19	66.45	63.12	varying		varying		–	0.75	0.81
AIREP	*T*	all	69 664 448	66 671 628	0.22	0.37	0.06	0.11	58.39	63.20	4.0	4.1	3.8	3.8	K	0.28	0.48
AIREP	*T*	100–900	58 595 984	56 113 624	0.19	0.33	0.06	0.11	59.51	66.33	4.0	4.1	3.8	3.8	K	0.25	0.44
AIREP	*T*	900–1000	11 067 772	10 557 288	0.33	0.54	0.10	0.13	52.47	46.55	5.7	5.9	5.1	5.0	K	0.43	0.67
AIREP	*u*/*v*	all	69 749 344	66 602 336	0.40	0.45	0.10	0.13	64.91	63.69	13.6	14.7	12.5	12.5	m s^{-1}	0.50	0.58
AIREP	*u*/*v*	100–900	58 935 220	56 307 792	0.40	0.46	0.10	0.14	64.93	66.43	13.6	15.7	12.6	12.6	m s^{-1}	0.50	0.60
AIREP	*u*/*v*	900–1000	10 813 077	10 293 823	0.40	0.41	0.07	0.07	64.84	48.71	13.8	14.7	12.5	12.5	m s^{-1}	0.47	0.48
PILOT	*u*/*v*	all	6 233 195	6 123 265	0.52	0.65	0.46	0.66	78.42	83.46	9.9	10.9	9.1	9.1	m s^{-1}	0.99	1.31
PILOT	*u*/*v*	0–100	1 326 562	1 245 501	0.73	0.81	0.50	0.62	83.01	84.43	12.3	13.0	10.2	10.2	m s^{-1}	1.23	1.43
PILOT	*u*/*v*	100–900	4 114 306	4 083 615	0.48	0.63	0.48	0.70	79.82	85.94	9.9	11.6	9.1	9.1	m s^{-1}	0.96	1.33
PILOT	*u*/*v*	900–1000	792 327	794 149	0.37	0.49	0.33	0.50	63.49	69.17	10.7	10.9	9.1	9.1	m s^{-1}	0.71	0.99
EU profil	*u*/*v*	all	13 216 667	13 885 356	0.27	1.62	0.14	0.40	24.00	52.85	10.9	10.4	9.1	9.1	m s^{-1}	0.42	2.03
EU profil	*u*/*v*	0–100	404 936	424 900	3.72	3.70	2.05	1.85	32.40	31.79	12.2	14.5	10.1	10.2	m s^{-1}	5.77	5.54
EU profil	*u*/*v*	100–900	11 212 697	11 732 330	0.18	1.53	0.09	0.34	26.49	54.59	10.9	10.4	9.1	9.1	m s^{-1}	0.27	1.87
EU profil	*u*/*v*	900–1000	1 599 034	1 728 126	0.07	1.78	0.02	0.46	4.43	46.20	11.4	10.7	9.1	9.1	m s^{-1}	0.09	2.24
JP profil	*u*/*v*	all	3 503 045	3 481 487	0.21	0.53	0.36	0.80	71.72	98.44	11.6	11.6	9.1	9.1	m s^{-1}	0.58	1.33
JP profil	*u*/*v*	100–900	3 012 748	2 993 565	0.25	0.52	0.42	0.73	83.39	98.61	11.6	12.0	9.1	9.1	m s^{-1}	0.67	1.25
JP profil	*u*/*v*	900–1000	490 297	487 922	0.00	0.61	0.00	1.22	0.00	97.41	n.a.	11.6	n.a.	9.1	m s^{-1}	0.00	1.83
US profil	*u*/*v*	all	11 139 485	11 302 865	0.46	1.00	0.24	0.46	70.60	98.31	11.4	11.1	9.7	9.1	m s^{-1}	0.70	1.45
US profil	*u*/*v*	0–100	179 565	175 906	1.67	3.23	0.31	0.59	86.05	96.17	13.9	14.5	11.2	10.8	m s^{-1}	1.98	3.82
US profil	*u*/*v*	100–900	10 697 506	10 860 851	0.45	0.90	0.24	0.42	72.07	98.43	11.4	12.2	9.7	9.1	m s^{-1}	0.69	1.32
US profil	*u*/*v*	900–1000	262 414	266 108	0.00	3.28	0.00	1.97	0.00	94.62	n.a.	11.1	n.a.	9.1	m s^{-1}	0.00	5.25

in 1990 and 219 Pa in 2008. Bias corrections decrease with time in most areas, although some ship observations in the Indian Ocean with large biases remain in 2008. It is possible that information about station height has improved during the period considered, but this remains to be investigated further.

4.3.2. *Radiosonde temperature bias correction*

Spurious shifts in station time series of radiosonde temperature observations can result from equipment changes at the station, or from other systematic changes, e.g. in observing practice. ERA-Interim uses the RAOBCORE_T_1.3 set of temperature adjustments (Haimberger *et al.*, 2008) to account for these types of biases. The adjustments were derived from analysis departure statistics from ERA-40 (prior to 2001) and from ECMWF operational analyses (from 2001 onward).

Additionally, diurnally varying biases due to solar radiative effects at high altitudes can be substantial, and these also depend on equipment type. The annual cycle of these radiative biases is removed using adjustments obtained with a sequential estimation scheme originally developed for ERA-40 (Andrae *et al.*, 2004). This scheme acts on geographically defined groups of stations, in order to avoid relying on information about equipment type, which is often unavailable for historic observations. It produces adjustments to the observed temperature profiles that depend on solar elevation, based on a rolling 12-month record of averaged background departures for each group.

Figure 13 shows the combined bias corrections applied to reported temperature observations at 50 hPa from radiosondes launched at Bethel, Alaska, at 0000 and at 1200 UTC. The large adjustments in 1989 and 1995 are associated with changes in instrumentation.

4.4. *Use of satellite radiances*

Figure 14 summarizes the sources of radiance data, converted to brightness temperatures for assimilation, grouped by instrument and satellite. The total number of different instruments used in ERA-Interim is close to 50, with only 6 instruments simultaneously available in 1989 and nearly 20 in 2010.

An important change in the use of satellite radiances relative to ERA-40 is the use of a variational bias correction scheme, as briefly described in section 2.1.2. This scheme effectively intercalibrates the radiance data from different satellites, sensors, and channels, using all other observations as a reference. Mainly these consist of conventional observations from radiosondes, aircraft, ships, and surface stations. Where few such observations are available, the variational bias adjustments can be affected by biases in the model background. It may then be necessary to anchor the assimilation system by using uncorrected radiances. This is the case, for example, near the top of the model, where systematic model errors are often larger than typical biases in satellite measurements of stratospheric radiances. For this reason, data from the highest-peaking channels of SSU, followed by AMSU-A after its introduction in 1998, were assimilated without bias correction in ERA-Interim (Kobayashi *et al.*, 2009). Details of the implementation of the variational bias correction scheme in ERA-Interim and implications for the quality and consistency of the reanalysis are presented in Dee and Uppala (2008, 2009).

4.4.1. *TOVS and ATOVS*

Since ERA-40, data usage from the TOVS and ATOVS suites of instruments (HIRS, SSU, MSU, AMSU-A/B, MHS) has been enhanced in several respects. General improvements in radiative transfer modelling have been mentioned in section 2.1.5. Additionally, the simulation of AMSU-A lower-tropospheric channels over land has improved due to a more sophisticated treatment of surface emissivity. A great deal of work was done to improve the assimilation of upper-stratospheric radiances from SSU and AMSU-A, as described by Kobayashi *et al.* (2009).

The blacklist excludes the following data from the analysis:

- for HIRS: all observations from the first or last three fields-of-view of each scan line; all observations from channels other than 2–7, 11–12, and 14–15; all observations from channels 2, 3, or 12 over land; all observations from channels 5–7, 11–12, or 14–15 over high terrain; all observations over the Caspian Sea (because of occasional large uncertainties in the SST in that region and subsequent poor cloud determination).
- for MSU: all observations from the first or last field-of-view of each scan line; all observations over land from channels 2–3; all channel 1 observations; all channel 2 observations believed to be contaminated by rain.
- for AMSU-A: all observations from the first or last three fields-of-view of each scan line; all channel 5 and 6 observations over high terrain; all observations from channels other than 5–14; all observations believed to be contaminated by rain.
- for AMSU-B and MHS: all observations from the first or last nine fields-of-view of each scan line; all observations from a channel other than 3–5; all observations over land for channel 5 and over high terrain for channels 3 and 4; all observations over sea ice (using the condition SST $<$ 273.15 K); all observations believed to be contaminated by rain.
- for SSU: none (data from all three channels and eight fields-of-view were assimilated).

In addition, detailed information about erroneous geolocation data for various TOVS and ATOVS instruments was obtained from the JRA-25 project at JMA and incorporated in the ERA-Interim data selection rules.

4.4.2. *Clear-sky radiances from geostationary satellites*

CSRs from geostationary infrared imagers, not used in ERA-40, were introduced in ERA-Interim in 2001. These sensors were flown on GOES (8–13), Meteosat (5, 7, 8, 9) and MTSAT (1R) satellites. Depending on the instrument, one or two channels peaking in the water vapour absorption band (6.2–6.8 and 7.3 μm) were used, with maximum sensitivity in the mid-to-upper troposphere. The CSR products as disseminated by the space agencies represent spatial averages of radiances identified as cloud-free. These are then assimilated using the RTTOV radiative transfer model. All CSR data are subject to variational bias correction.

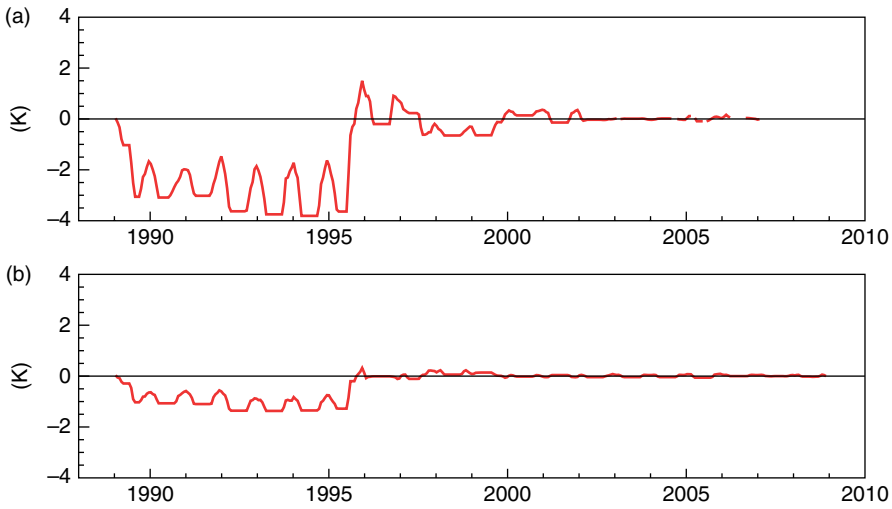

Figure 13. Temperature bias corrections at Bethel, Alaska for 50 hPa radiosonde reports (a) at 0000 UTC and (b) at 1200 UTC. Adjustments vary in time to take into account the pronounced annual cycle of the radiation error at high latitudes and low solar elevations. Shifts in 1989 and 1995 are caused by changes from VIZ to Space-Data and from Space-Data to Vaisala RS80-56 radiosondes. Night-time adjustments are generally weaker than daytime (0000 UTC at Bethel) adjustments.

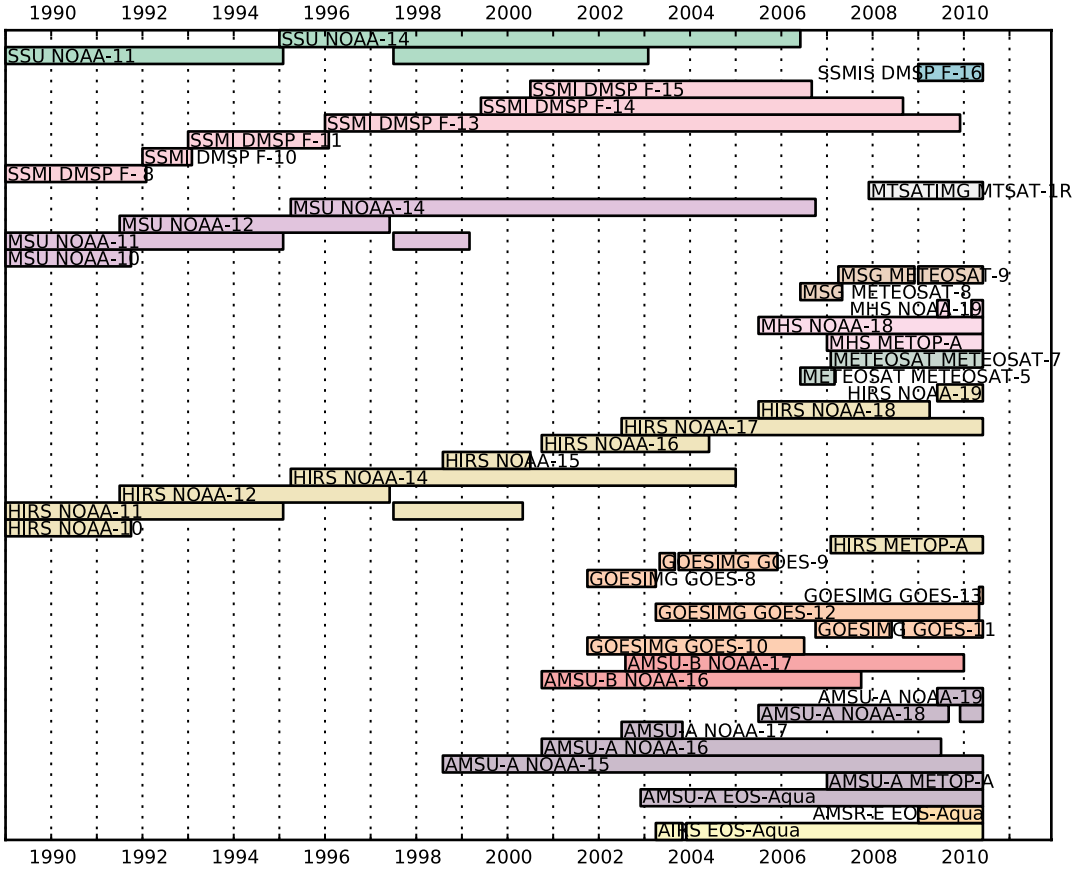

Figure 14. Timeline of clear-sky radiance observations assimilated in ERA-Interim.

The blacklist excludes CSR observations made during eclipse seasons, over high terrain, and with satellite zenith angles larger than 60°.

4.4.3. *Passive microwave imagers*

The third family of instruments used in ERA-Interim are the passive microwave imagers (SSM/I, SSMI/S, AMSR-E). These instruments exhibit a large sensitivity to surface emission and to integrated quantities such as TCWV, unlike microwave sounders such as AMSU-A/B and MHS, which primarily measure atmospheric layer quantities. Imagers are characterised by a conical scan pattern that ensures a constant viewing geometry across the scan to facilitate discrimination between surface and atmospheric contributions to the radiance measurements. Due to the difficulties in modelling land-surface emissivity with its large spatio-temporal variability, these data have been assimilated over oceans only.

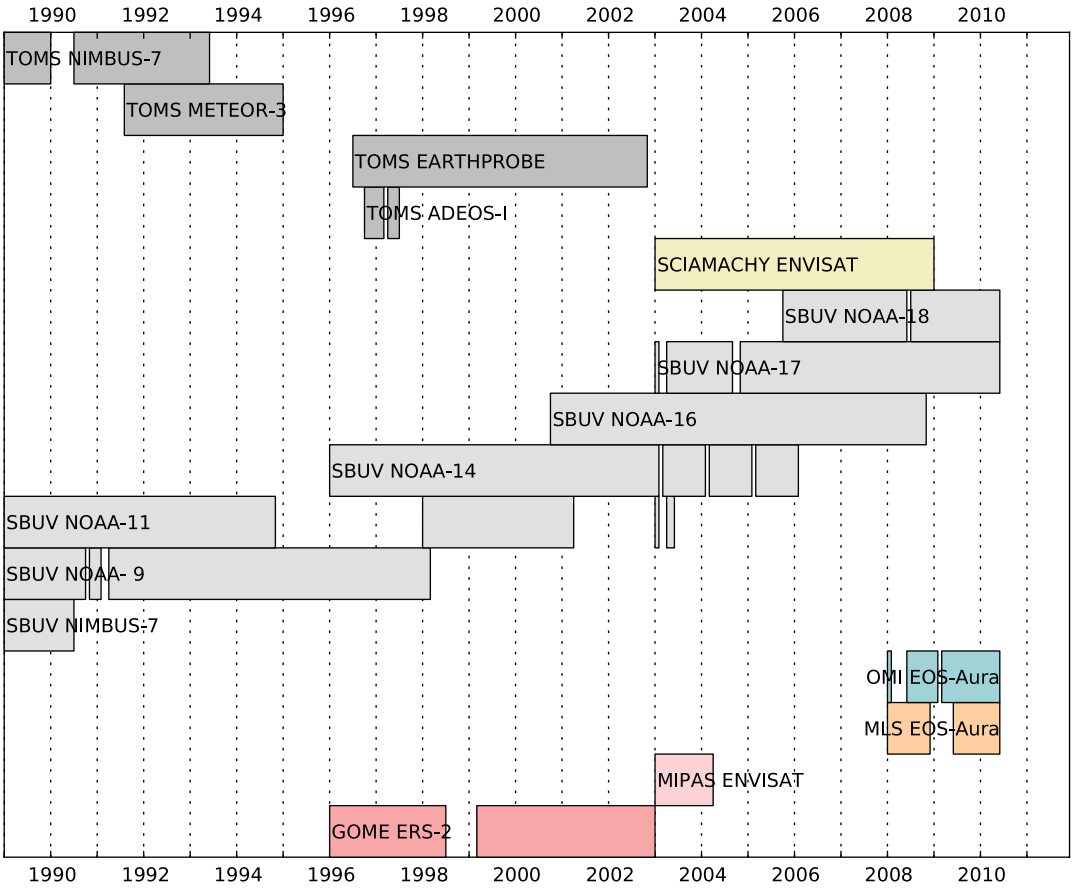

Figure 15. Timeline of ozone data assimilated in ERA-Interim.

Since ERA-40, major progress has been made in the ability to usefully assimilate imager data. ERA-40 used a 1D+3D-Var scheme to assimilate SSM/I data in clear-sky conditions (Phalippou, 1996; Gérard and Saunders, 1999). This scheme produced 1D-Var estimates of TCWV and 10 m wind speed, which were then used in the 3D-Var global analysis. Direct assimilation of clear-sky SSM/I radiances was introduced in the ECMWF operational model in 2002 (Bauer *et al.*, 2002). This method was used in ERA-Interim for SSM/I data from DMSP F-8/10/11/13/14/15, and also for the window channels of SSMI/S data from DMSP F-16 and AMSR-E data from EOS-Aqua.

Data from microwave imagers in cloudy and rain-affected areas are assimilated in ERA-Interim as TCWV estimates, obtained using the 1D+4D-Var scheme described in section 2.1.6. This scheme employs static bias corrections for each sensor, which are estimated from radiance background departures using short test reanalyses.

The blacklist selection rules exclude all passive microwave imager observations over land, all SSM/I-S observations from channels other than 12–18, and all AMSR-E observations from channels other than 5–10. From 2009 onward, 1D-Var TCWV retrievals were not used at latitudes greater than 45°.

4.4.4. *AIRS data*

Radiance measurements from the 2378-channel grating spectrometer AIRS on EOS-Aqua were introduced in ERA-Interim in April 2003. A subset of 324 channels is used, manually selected to sample the main features of the observed spectra. Only one out of nine footprints –the warmest field of view –is used in each observed scene. This maximises the likelihood of observing a clear scene. A cloud-detection scheme is applied as described by McNally and Watts (2003). This scheme exploits the relatively high vertical resolution of the instrument to identify only cloud-affected channels at each cloudy location. All AIRS channels flagged clear are subject to an additional random thinning operation that ensures a minimum horizontal spacing of 120 km.

Blacklist rules exclude the following AIRS data:

- Channels with wavelengths less than 4.46 μm;
- Channels in the ozone band;.
- Channels with significant surface sensitivity, at land locations;
- Channels sensitive to trace gases unaccounted for in RTTOV;
- All channels at the edge of each scan.

After these exclusions, up to 210 channels are available for assimilation in a typical cloud-free scene over ocean (McNally *et al.*, 2006).

4.5. *Use of other satellite data*

4.5.1. *Ozone retrievals*

ERA-Interim uses a greater variety of ozone data than ERA-40. During the overlap between the two reanalyses, from January 1989 to August 2002, ERA-Interim additionally assimilated partial column ozone retrieved from SBUV/2 on NOAA-14 as well as ozone profiles retrieved from GOME on ERS-2, produced by Rutherford Appleton Laboratory (Siddans *et al.*, 2002). Figure 15 shows a timeline of all ozone

instruments assimilated in ERA-Interim. Observations from SBUV, OMI, TOMS, GOME, and SCIAMACHY with solar elevation angles less than 6°, 10°, 10°, 15° and 6°, respectively, were excluded. Dragani (2010a) presents further details on these data and how they were used in ERA-Interim.

4.5.2. *Scatterometer ocean surface winds*

Space-borne scatterometers provide surface vector wind information over the global oceans on 25 km intervals in polar orbits with swaths between 550 km and 1800 km wide. Figure 16 shows the timeline of scatterometer data used in ERA-Interim, from ERS-1, ERS-2, and QuikSCAT. Due to the failure of several gyroscopes on board ERS-2 in January 2001, the data stream was interrupted until 21 August 2003. ERS-2 data coverage since then has been mostly limited to the North Atlantic region. QuikSCAT data, not used in ERA-40, were introduced in ERA-Interim on 24 February 2000, when rain-contamination flags began to be provided with the data. QuikSCAT failed at the end of 2009. ERA-Interim does not assimilate data from the ASCAT scatterometer on Metop-A because the relevant code developments in the IFS were made after the start of production. All scatterometer data are screened for land and sea ice. ERS data are thinned to 100 km, while information from QuikSCAT data is aggregated at 50 km resolution prior to analysis. For QuikSCAT, data flagged as contaminated by rain and data from the outer 200 km of the satellite swath are not used, and a consistency check on the inverted wind product is performed.

Prior to assimilation, a special effort was made to intercalibrate scatterometer wind data from ERS-1 and ERS-2. Observed backscatter values from scatterometers require conversion to wind information, which is achieved by a so-called GMF. For ERA-40, wind inversion was based on CMOD4 (Stoffelen and Anderson, 1997), and biases of this GMF were corrected in the wind-speed domain (Isaksen and Janssen, 2004). In addition, corrections in backscatter had been applied to compensate for known changes in calibration levels in the product as received from ESA over time. Nevertheless, close examination of the ERS wind product as used in ERA-40 showed significant residual dependencies on time and swath position. A recalibration was therefore performed prior to assimilation in ERA-Interim. An improved GMF (CMOD5.4) removed most of the need for wind-speed bias correction after wind inversion (Abdalla and Hersbach, 2007). Adjustments were made for small differences in calibration between ERS-1 and ERS-2, and suitable corrections made it possible to introduce ERS-1 data on 16 April 1992, rather than 1 January 2003 as in ERA-40.

4.5.3. *Global Positioning System radio occultation measurements*

Figure 16 also shows the use of atmospheric refractivity information (expressed as bending angle profiles) derived from GPSRO measurements. Relatively small numbers of profiles from CHAMP were assimilated beginning June 2001. A larger volume of data from the six-satellite constellation COSMIC was introduced in December 2006, supplemented by data from the GRAS instrument on MetOp-A after May 2008.

GPS bending angle profiles provide information at high vertical resolution about tropospheric temperature and humidity profiles as well as stratospheric temperatures. Measurement locations change from day to day but tend to be uniformly distributed over the globe. GPS data are used without bias correction and can help to constrain the variational bias adjustments made to radiance observations. Poli *et al.* (2010) provide additional details on these data and their assimilation in ERA-Interim.

4.5.4. *Atmospheric Motion Vectors*

As in ERA-40, a large number of satellite observations of the upper-air wind field in the form of AMVs are assimilated in ERA-Interim. The majority of these wind data are estimated from geostationary satellite imagery between about 55°N and 55°S. AMV data from Meteosat, GOES, GMS, and MTSAT were used as shown in Figure 16. Some early satellite sources were not correctly identified when originally received at ECMWF; these are labelled GOES-X and GMS-X in the figure. Segments of the AMV record for the Meteosat series were reprocessed for ERA-Interim by EUMETSAT (Delsol *et al.*, 2008). Data selection rules for geostationary AMVs make use of quality indicators as provided with the data, if available. In addition, low-level AMVs over land and all AMVs over the Himalayas are excluded.

AMV winds at high latitudes from MODIS instruments on the polar-orbiting EOS-Terra and EOS-Aqua satellites have been used since February 2007. These data had in fact been available from July 2002 but were inadvertently excluded from ERA-Interim. The MODIS AMVs are derived from subsequent orbit overpasses which allow the identification and tracking of cloud and water vapour features. Over land surfaces, only MODIS cloud and water vapour winds above 400 hPa are used due to inaccurate height assignment over high orography and ice. Over the ocean, cloud winds located higher than 700 hPa and water vapour winds above 550 hPa are added (Bormann and Thépaut, 2004).

4.6. *Surface observations*

As described in section 2.2, the analysis of near-surface parameters (2 m temperature and humidity), snow, and ocean waves is performed in separate steps following the upper-air atmospheric 4D-Var analysis. These components of the data assimilation system use relatively simple data interpolation schemes, and they do not generate comprehensive information about data quality control and usage. Many scientific and technical improvements in the surface analysis have been implemented in recent years, but these were not available at the end of 2006 when the configuration of the ERA-Interim data assimilation system was fixed.

4.6.1. *Land surface*

ERA-Interim uses an optimal interpolation scheme to analyse observations of screen-level temperature and relative humidity from surface stations. Global data counts for each parameter typically reach 20 000 per day throughout the reanalysis period. A Cressman-type interpolation is used to analyse station observations of snow depth, with daily counts ranging from a minimum of approximately 500 in

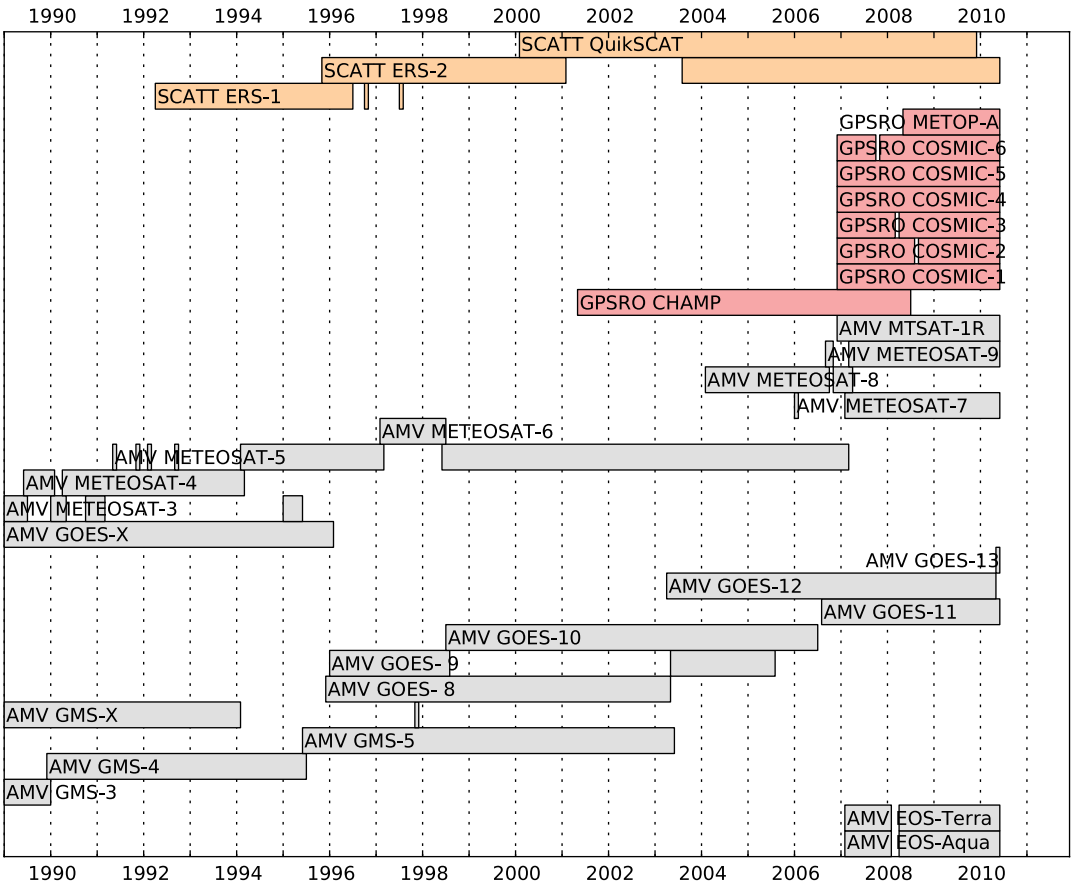

Figure 16. Timeline of satellite observations assimilated in ERA-Interim, other than clear-sky radiances or ozone.

boreal summer to 2000 in boreal winter. Snow observations between 1992 and 1994, inadvertently omitted in ERA-40, have been assimilated in ERA-Interim.

The NOAA/NESDIS daily IMS snow-cover product has been used to constrain the ERA-Interim snow analysis since July 2003. This satellite-based dataset covers the Northern Hemisphere at a spatial resolution of 24 km. At locations where the model background is free of snow but the IMS product indicates snow cover, 10 cm of snow is added to the background field. At locations that are snow-covered according to the model background but snow-free in the IMS product, a pseudo-observation of zero snow depth is presented to the analysis.

ERA-Interim snow analyses from 1 July 2003 to 23 February 2010 are affected by a geolocation error introduced during the processing of the IMS product at ECMWF. Location of data and auxiliary information used in the analysis was shifted in different directions, by about 100 km toward the southeast for the orography and land–sea mask, and by about 17 km toward the northwest for the data themselves. This has caused incorrect removal of snow in some coastal areas in the Northern Hemisphere during winter.

4.6.2. Ocean waves

The ocean wave analysis in ERA-Interim incorporates an optimal interpolation scheme to constrain predicted wave spectra using altimeter wave height observations. ERA-Interim used reprocessed ERS-1 and ERS-2 data from ESA, and near-real-time data from ENVISAT, JASON-1, and JASON-2, as received by ECMWF operations (Table V). Since ERS-2 and ENVISAT follow the same ground tracks with a separation of only 20 min, ERS-2 data were blacklisted when ENVISAT became available. Quality control for wave height observations includes a background check. In order to reduce the variability in the data to a level supported by the model resolution, an along-track averaging procedure is performed prior to the analysis (Abdalla and Hersbach, 2004).

Table V. Use of altimeter data in the ERA-Interim wave analysis.

ERS-1	01 Aug 1991 to 03 Jun 1996
ERS-2	03 May 1995 to 21 Jul 2003
ENVISAT	21 Jul 2003 to present
Jason-1	20 Oct 2003 to present
Jason-2	01 Feb 2010 to present

No observations of wave height were available to constrain the ocean wave spectra prior to the introduction of ERS-1 in August 1991. Daily coverage of a single satellite-borne altimeter represents approximately 10% of the number of grid points in the wave model. Consequently, over time, the percentage of the total domain for the wave model constrained daily by observations fluctuates between about 10% (when one satellite is available) and about 20% (when two satellites are available).

During the early part of the ENVISAT mission, the number of data received was erratic due to the unavailability of a telecommunication satellite. This issue was resolved

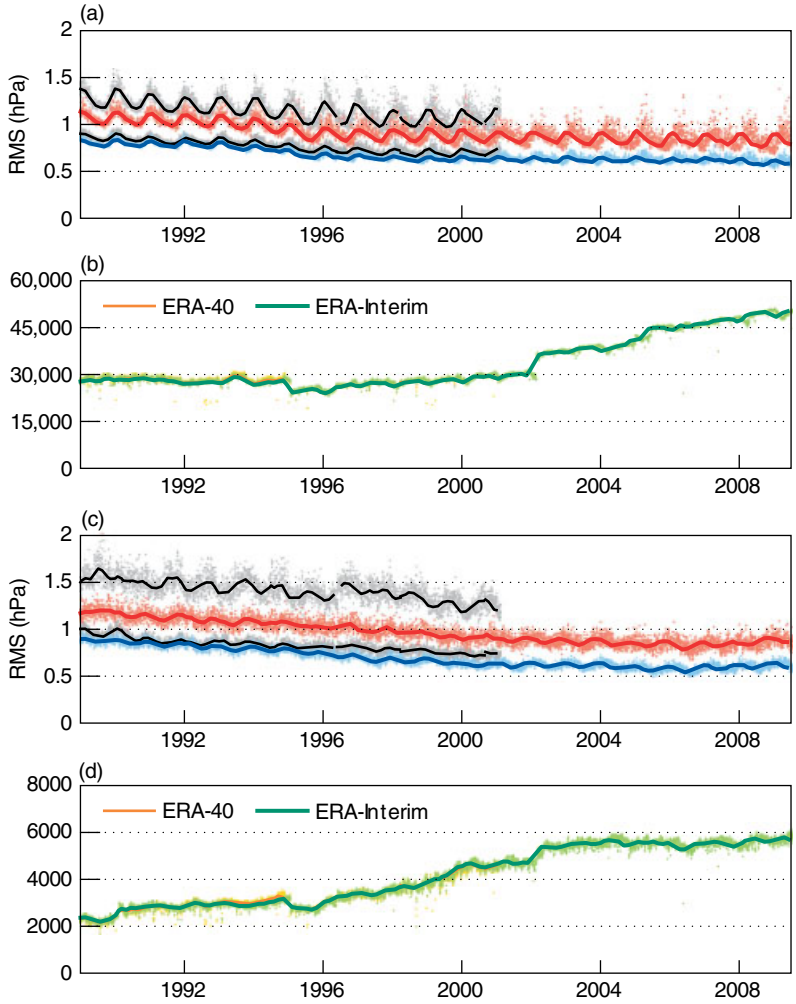

Figure 17. (a) RMS departures and (b) daily counts for surface pressure observations (hPa) assimilated in ERA-Interim and ERA-40, for the Northern Hemisphere. Solid curves are 3-month running averages, and lightly shaded dots show daily values. ERA-Interim background departures are shown in red, analysis departures in blue, and background and analysis departures for ERA-40 are shown in black. Gaps in 3-monthly averages occur whenever data are missing for at least one day. (c, d) are as (a, b), but for the Southern Hemisphere.

during the second half of June 2004. As a consequence, however, some orbits are missing in the near-real-time archive. Similar issues with data arriving late, or not at all, frequently occur in operational weather forecasting. Future reanalyses would benefit from using a back-filled and consistently reprocessed archive.

5. Performance

5.1. Data assimilation diagnostics

Data assimilation performance has been closely monitored during the production of ERA-Interim. An extensive suite of automated monitoring tools was developed to allow quick detection of possible problems with the selection and quality control of input observations, and to provide a first look at the overall quality of the reanalysis as it is being produced.

5.1.1. Departure statistics

Monitoring of the data assimilation involves keeping track of data usage for all components of the observing system. As explained in section 2.1, the assimilation system predicts every single observation before it is used, based on a short model forecast. The background departures generated in this way contain a great deal of information about the quality of the reanalysis, and how it evolves in time. Analysis departures produced at the end of each analysis cycle show how closely the reanalysis fits the data at the observation locations. However, the degree of fit is largely controlled by prescribed error statistics used in the analysis, and is therefore not very useful as a performance indicator.

Figure 17 shows hemispheric data counts and departure statistics for all assimilated surface pressure observations from land stations, for the period January 1989 to May 2010, for both ERA-Interim and ERA-40. RMS departures for both reanalyses reduce with time as the observing system improves. This is especially evident in the Southern Hemisphere. ERA-Interim background departures are significantly smaller than those for ERA-40, showing that the ERA-Interim assimilation system is better able to predict future observations. Consequently, smaller adjustments to the surface pressure field are needed in ERA-Interim, and this results in a much improved temporal consistency on synoptic time-scales. On the other hand, ERA-40 analysis departures are smaller than those for ERA-Interim, but this is simply the result of having used a weaker background constraint in the variational analysis of ERA-40.

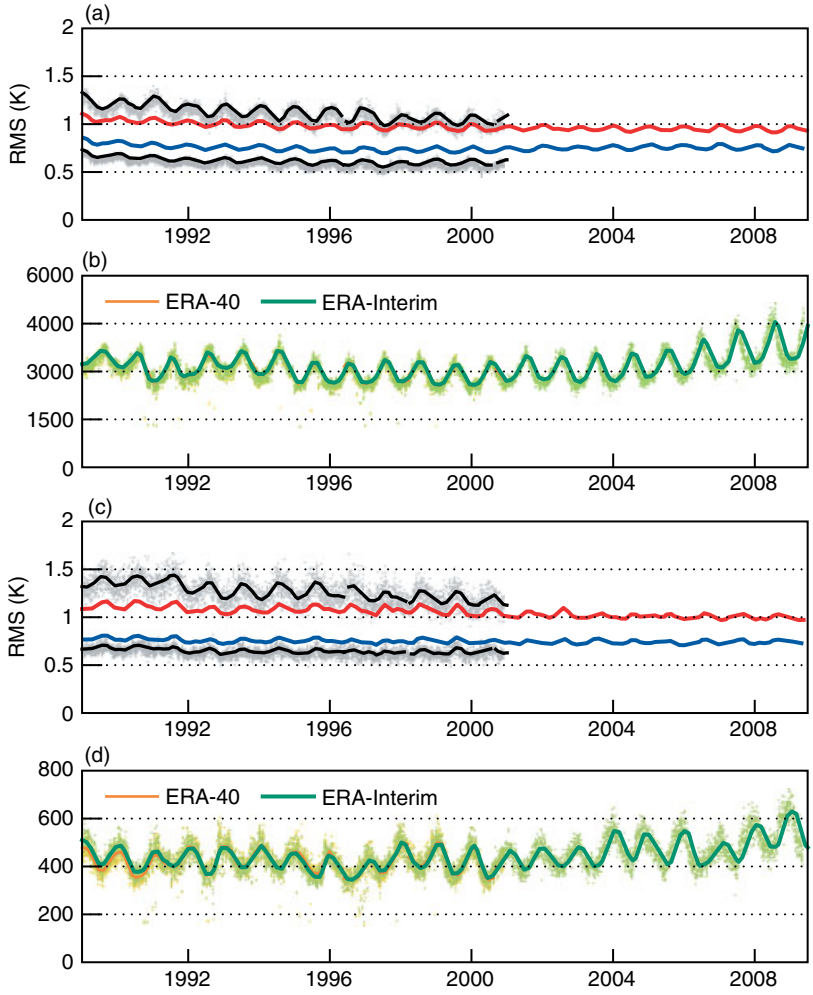

Figure 18. As Figure 17, but for temperature (K) observations from radiosondes at pressure levels between 450 and 600 hPa.

Similarly, Figure 18 shows hemispheric statistics of temperature departures from radiosonde observations at levels near 500 hPa. The strong seasonal cycle in departure statistics and data counts may be associated with increased synoptic variability in winter. It is likely that some data rejections in winter result from inadequate background-error specifications, which are used for quality-control decisions, but which do not properly account for changes in the atmospheric circulation. Improvements in the ERA-Interim background fit to observations are evident in this case as well, even though analysis departures are somewhat larger than those for ERA-40. As for surface pressure, the forecast model requires smaller corrections yet is better able to retain the information supplied by observations. In addition, the ERA-Interim departures are somewhat more uniform in time than those of ERA-40, especially in the Southern Hemisphere, probably as a result of better use of satellite observations.

For tropospheric humidity, Figure 19 shows data counts and RMS departures for radiosonde-observed specific humidity at levels in the vicinity of 700 hPa, averaged for tropical stations only. Similar to the temperature departures, ERA-Interim has smaller analysis increments yet better background departures than ERA-40, and these statistics do not vary greatly during the reanalysis. A weak seasonal cycle in the departures is still evident, since the majority of tropical stations are in fact located in the Northern Hemisphere.

The usage of ocean-surface wind data from scatterometers is summarised in Figure 20, in terms of global data counts and global departure statistics. Results are shown for ERA-Interim only. The ERS-2 outage mentioned in section 4.5.2 and its return to limited North Atlantic coverage in August 2003 are evident in the data counts and in the large seasonal variation of the departures. Compared to ERA-40 statistics (not shown here), long-term globally averaged standard deviations of background departures have improved considerably, from 1.62 m s^{-1} (Isaksen and Janssen, 2004) to 1.45 m s^{-1}. Statistics for ERS-1, ERS-2 and QuikSCAT overlap well. Beginning in 1999, departures gradually decline to a level approaching 1.32 m s^{-1} near the end of 2008, indicating an improvement over time in the quality of ERA-Interim surface winds. Similarly to the other variables discussed, standard deviations of ERA-Interim analysis departures (1.10 m s^{-1}) are larger than those for ERA-40 (0.97 m s^{-1}).

The ERS-1 and ERS-2 intercalibration described in section 4.5.2 has led to a coherent set of departure statistics for the combined ERS-1 and ERS-2 datasets (i.e. the blue curves in Figure 20). There is no visible sign of the transition from ERS-1 to ERS-2, other than the increased combined data volume around 1996. On the other hand, global mean background departures for ERS and QuikSCAT differ by approximately 0.15 m s^{-1}. Direct collocation of QuikSCAT and ERS data during the overlap period has confirmed a small wind-speed-dependent bias between the two scatterometers (not shown). Although the mismatch represents only 2% of the global mean wind speed, the issue of intercalibration between ERS and QuikSCAT

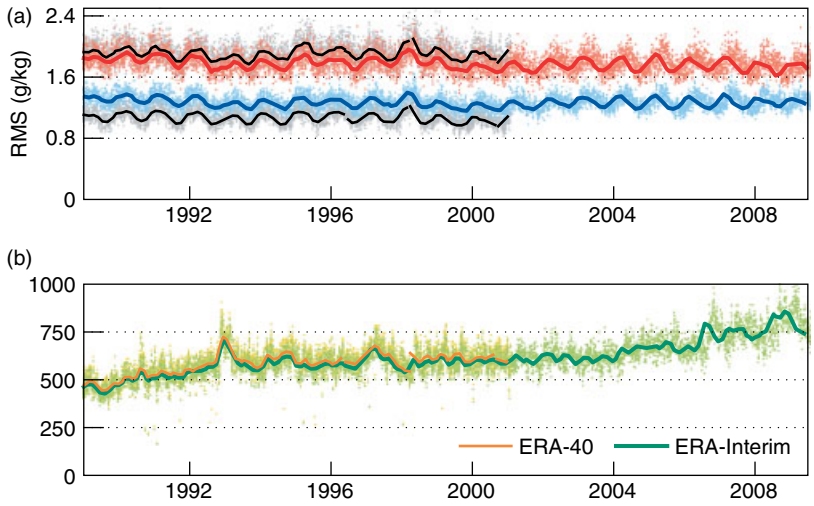

Figure 19. As Figure 17, but for humidity ($g\,kg^{-1}$) observations from tropical radiosondes, at pressure levels between 600 and 775 hPa.

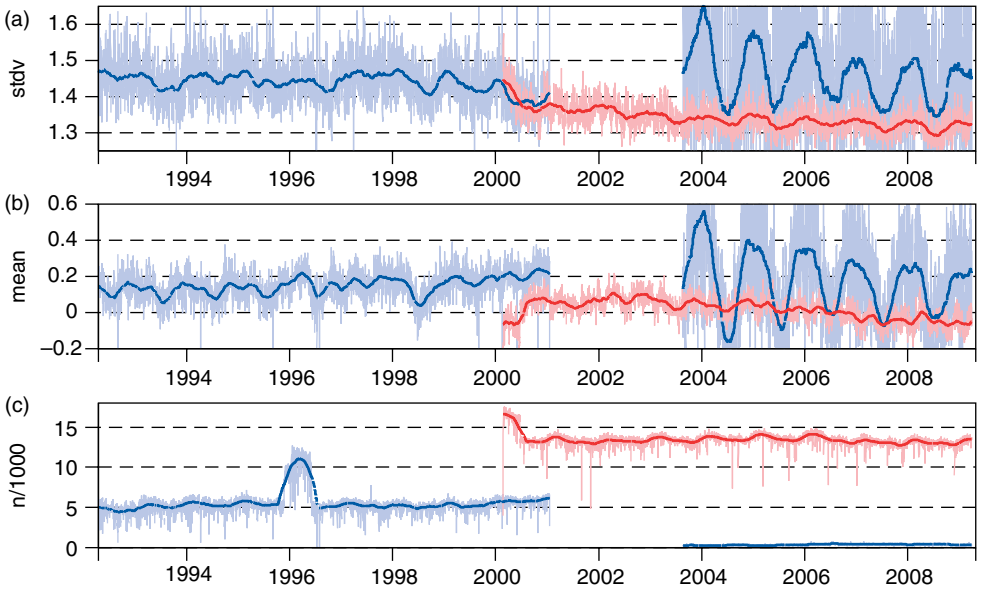

Figure 20. Global means and standard deviations of ERA-Interim wind-speed background departures ($m\,s^{-1}$), and the daily number of wind vectors used, from ERS-1 and ERS-2 starting on 16 April 1992 (blue), and from QuikSCAT starting on 24 February 2000 (red). The thin, lightly shaded curves show the daily averages, and the dark colours the three-monthly moving averages. (a) shows standard deviations of ERA-Interim background departures, (b) the global mean background departures, and (c) the daily data counts.

deserves more attention in a future reanalysis. The change in behaviour of QuikSCAT near the end of 2000 also merits further investigation.

The gradual increase in global mean departures between April 1992 and 2002 visible in Figure 20 indicates diminishing surface wind speeds in ERA-Interim compared to ERS and QuikSCAT, by approximately 0.1 $m\,s^{-1}$ over the 10-year period. After 2002, the trend is reversed. Increasing surface wind speeds over the oceans have also been seen in the ECMWF operational forecast system after 2002, where it was ascribed to changes in model resolution and physics. It now seems more likely to be related to the evolution of the observing system.

5.1.2. *Analysis increments*

As described in section 2, the ERA-Interim data assimilation advances in 12-hourly analysis cycles, each of which produces an adjustment to the prognostic model variables needed to maintain consistency of the model state estimates with the available observations. These state adjustments, which are usually referred to as analysis increments, represent the net response of the variational data assimilation to all observations used. They therefore provide sensitive diagnostics of the end-to-end performance of the system.

Generally, the variability of the analysis increments depends on the amount of information extracted from the input observations, but this relationship is not straightforward. For example, small increments can be the sign of a very good forecast model, but they can also be due simply to a lack of observations. In a sparsely observed situation, increased variability can be expected with increased data coverage, but may also indicate improper use of certain types of observations. A case in point is the assimilation of high-quality ozone profiles from GOME, which initially caused unrealistic temperature and wind increments in the upper stratosphere, pointing to an issue with the 4D-Var analysis of ozone (section 2.1.4).

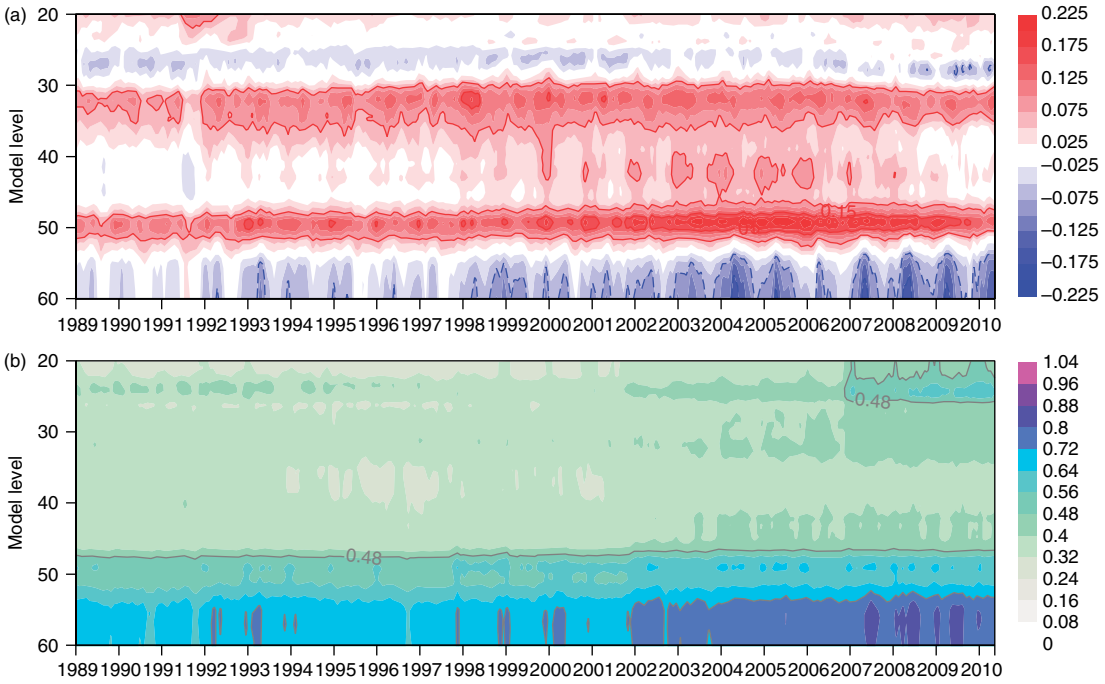

Figure 21. Twice-daily (a) global averages and (b) standard deviations of analysis increments for temperature (K). The vertical axis indicates model levels 60 (nearest to the surface) to 20 (near 40 hPa). Values are smoothed in time by 30-day averaging.

Systematic increments usually indicate the presence of residual biases in observations, the forecast model, or both. These generally introduce artificial sources and sinks of heat, energy, and water in the reanalysis, and hence affect the global budgets for these quantities. Changes in the mean increments introduce uncertainties in trend estimates for basic climate variables derived from reanalysis data. This is discussed further in section 5.2.3 below.

Figure 21 summarises the time sequence of twice-daily analysis increments for tropospheric temperature produced in ERA-Interim. Shown are 30-day averages of the global means and global standard deviations for each increment at each model layer. The vertical structure of the mean increments, which is quite persistent, indicates a systematic warming of the model troposphere by the observations, especially near model layers 50 (≈850 hPa) and 32 (≈250 hPa). The reverse is the case nearer the surface and just above model layer 30. The model has, on average, a cold bias in the troposphere, which is only partly corrected by observations. The degree of correction depends on the observation coverage, which changes in time.

The most relevant events affecting changes in tropospheric temperature increments in ERA-Interim are the growing numbers of aircraft reports in the late 1990s and the introduction of GPSRO data from COSMIC at the end of 2006. The effect of the latter is clearly visible in the variability of the upper tropospheric increments. The large increase in satellite observations between late 1998 and 2003 from ATOVS and various geostationary imagers is also evident.

The introduction of ocean-surface wind information from ERS in 1992 and QuikSCAT in 2000 adds to the variability of zonal wind increments over oceans near the surface, as shown in Figure 22(b). As a result of the intercalibration of scatterometer wind data described earlier, the effects on mean wind increments are minimal. Mid- and upper-tropospheric wind information is greatly increased in 2002, most likely due to the introduction of clear-sky radiance observations from GOES-8 and GOES-10. The impact of GPSRO data on upper-tropospheric winds is visible as well.

5.2. *Challenges from ERA-40*

Difficulties encountered in the production of ERA-40 included the representation of the hydrological cycle, the quality of the stratospheric circulation, and the temporal consistency of the reanalysed fields.

5.2.1. *Precipitation*

Accurate representation of the hydrological cycle in reanalysis presents a special challenge because it involves many parameters that are constrained only indirectly by observations. Estimates of precipitation associated with the reanalysis are produced by the forecast model, based on temperature and humidity information derived from the assimilated observations. Approximations used in the model's representation of moist processes strongly affect the quality and consistency of the hydrological cycle. Any imbalances in the analysed fields relative to the model equations can cause large initial changes in the forecast. These so-called spin-up/spin-down effects must be taken into account when using the output of the model to estimate hydrological parameters and other atmospheric fluxes.

Uppala *et al.* (2005) have explained the various difficulties encountered in ERA-40 with the assimilation of humidity information, which led to excessive rainfall over tropical oceans and a generally poor representation of the global transport of moisture in the atmosphere. These problems stemmed from a combination of factors, related to the formulation of the humidity analysis scheme and exacerbated by effects of the Pinatubo eruption on the bias adjustments for HIRS infrared radiances. Both the humidity analysis (section 2.1.4) and the method for correcting biases

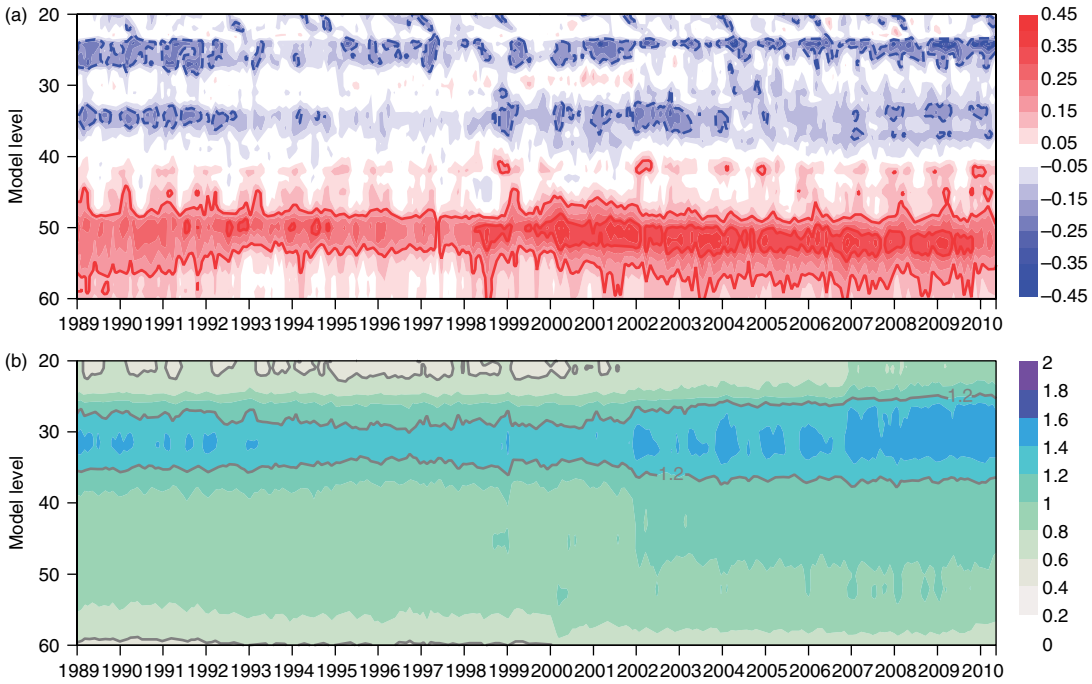

Figure 22. As Figure 21, but for u-wind increments ($\mathrm{m\,s^{-1}}$) and for ocean locations only.

in radiance data (section 2.1.2) have been completely revised in ERA-Interim. The role of variational bias correction in particular to the response of the data assimilation system to the Pinatubo eruption is discussed in some detail in Dee and Uppala (2009).

Figure 23 compares annual-mean precipitation rates from ERA-Interim and ERA-40 with observation-based estimates of the GPCP (Adler *et al.*, 2003). Figure 23(a,b,c) show differences between ERA-Interim and GPCP, between ERA-40 and GPCP, and between ERA-Interim and ERA-40. These results apply to the year 1990, prior to the excessive increase in precipitation seen in ERA-40. Figure 23(d, e, f) show the differences in TCWV based on the version-6 SSM/I retrievals over oceans produced by RSS (Wentz, 1997). Precipitation is higher in both ERA-Interim and ERA-40 than in GPCP over the tropical oceans. ERA-Interim is closer to GPCP, but ERA-40 and ERA-Interim are nevertheless in closer agreement with each other than either is to the GPCP estimate. At higher latitudes, ERA-Interim is in closer agreement with GPCP than ERA-40. TCWV from ERA-Interim is significantly lower than from ERA-40, and closer to RSS.

Figure 24(a, b, c) show time series of monthly averaged precipitation from ERA-Interim, ERA-40, and GPCP, for the entire globe, and restricted to land and ocean locations. The ERA-Interim estimates in this and subsequent figures represent model-generated rainfall accumulated during the first 12-hour segment of each forecast issued from the 12-hourly analyses. Figure 24(d, e, f) show the values obtained when using the first, second or third 12-hour forecast segments. Differences among those curves reflect the spatially averaged effects of spin-up/spin-down on precipitation during the forward integration of the model. Estimates from the initial forecast segments show the impact of observations most directly. Although less pronounced than in ERA-40, precipitation spin-up is still present in ERA-Interim, especially over oceans in midlatitude storm tracks. Spin-down typically occurs over tropical rain forests. These and other details of the spatial distribution and temporal behaviour of precipitation spin-up/spin-down are presented by Kållberg (2011), who also discusses the effects of forecast drift on estimates of cloud cover and radiative fluxes.

The ERA-Interim precipitation estimates do not show the excessive rainfall seen in ERA-40 following the eruption of Pinatubo, and they are much more stable throughout the reanalysis period. However, an abrupt reduction of approximately $1\,\mathrm{mm\,day^{-1}}$ in the global mean occurs in ERA-Interim at the beginning of 1992, followed by additional (but smaller) reductions in subsequent years. After 2006 the values gradually return to their pre-1992 levels. The shifts are evident only over oceans.

The shifts in global mean precipitation are spurious and can be entirely explained by a problem in the 1D+4D-Var rain assimilation scheme (section 2.1.6), which caused it to dry the analyses erroneously. In the early version of the scheme used for ERA-Interim, the linearised moist physics in the 1D-Var observation operator, which simulates rain from the model state, systematically overestimated rainfall by approximately a factor of two. Hence, the model appeared to produce excess rain, which was then corrected by observations. The resulting tendency to dry the model was mitigated but not entirely removed by bias correction of the radiances, and also by sampling effects caused by using data only when rain is observed, as explained in more detail by Geer *et al.* (2008).

The net drying effect of the 1D+4D-Var rain assimilation therefore depends on the volume and spatial coverage of the rain-affected SSM/I radiance data. Figure 24(d, e, f) indicate that the information from these data is not well retained by the model; the impact of the SSM/I data over the ocean, which is clearly evident in the shifts in mean precipitation during the first 12-hour segment of the forecast, has nearly disappeared in the second and third segments. It has been verified experimentally that the large change in global mean precipitation at the beginning of 1992 is not present when rain-affected SSM/I radiances are withheld from the assimilation. Subsequent shifts visible in

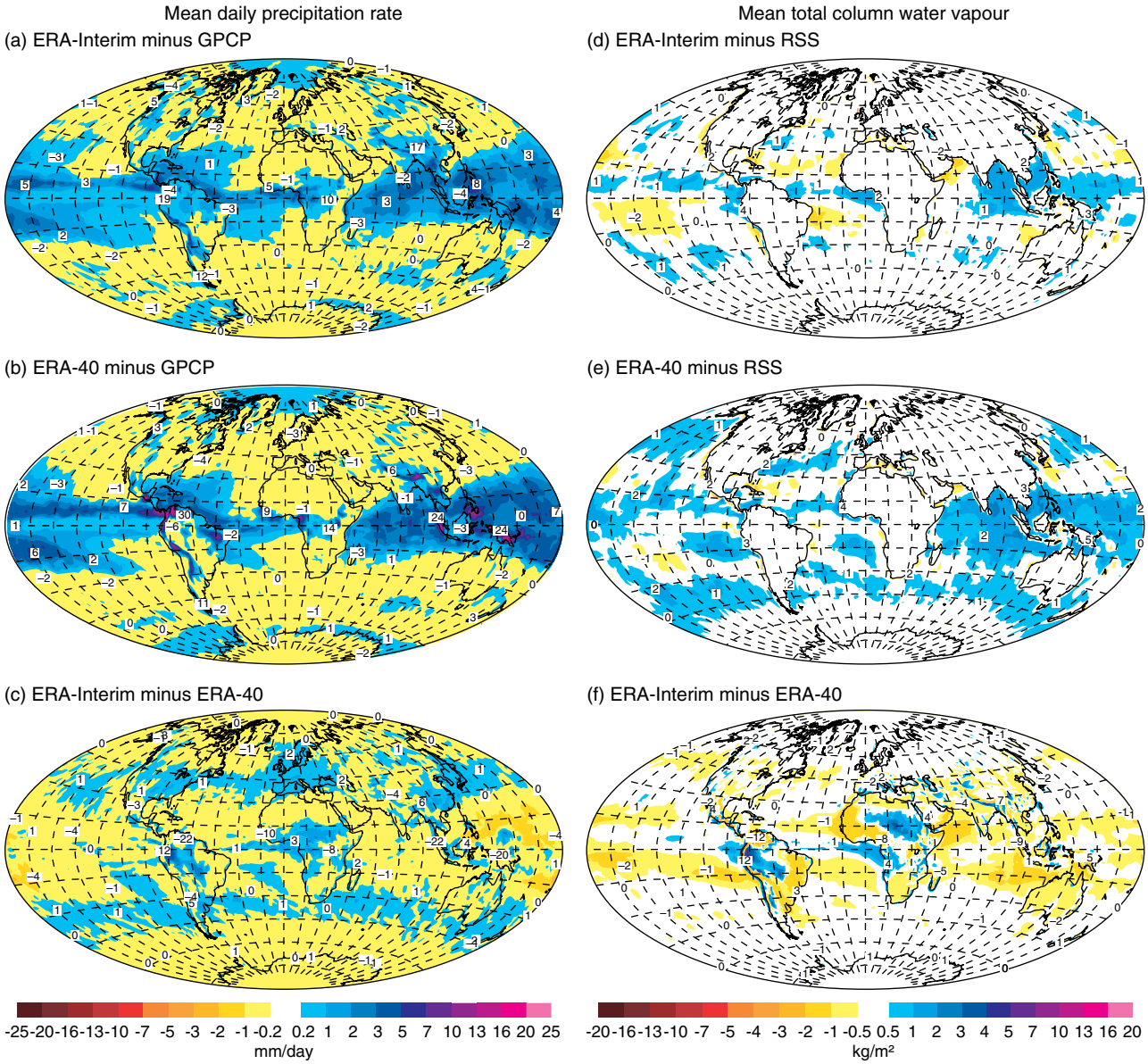

Figure 23. Differences in mean daily precipitation for 1990 (mm d^{-1}): (a) ERA-Interim minus GPCP, (b) ERA-40 minus GPCP and (c) ERA-Interim minus ERA-40. Differences in mean total column water vapour for 1990 (kg m^{-2}): (d) ERA-Interim minus RSS, (e) ERA-40 minus RSS and (f) ERA-Interim minus ERA-40. RSS denotes the version-6 retrievals from SSM/I produced by Remote Sensing Systems.

the time series also coincide with changes in data counts for TCWV assimilated with the 1D+4D-Var scheme, shown in Figure 10. The changes in data coverage are determined by the availability of SSM/I radiances from the DMSP series (Figure 14). It should be noted that, prior to June 1999, ERA-Interim assimilated SSM/I data from at most one instrument (from the ERA-40 input); beginning in June 1999, data from multiple instruments were used (from the ECMWF operational data stream).

The impact of satellite observations on the reanalysis of precipitation is most directly felt over the oceans. Over land, the information used by the model to generate rain is more strongly constrained by *in situ* measurements of temperature and humidity from radiosondes and land stations. The quality of precipitation estimates from reanalyses therefore tends to be better over well-observed land locations than over oceans. Similarly, observation-based estimates over land derived from rain-gauge data tend to be more reliable than those based on microwave products over ocean, although biases and uncertainties exist in both sources.

Simmons *et al.* (2010) have compared continental precipitation from ERA-Interim and ERA-40 with various observation-only data products. Figure 25(a, b) show 21-year average precipitation rates over land, obtained from the 1° × 1° GPCC product (Schneider *et al.*, 2008) and from ERA-Interim. Figure 25(c, d) show the corresponding estimates of decadal change in precipitation between 1990–1999 and 2000–2009. A similar comparison was done with GPCPv2.1, which differs from GPCC mainly due to corrections for under-catch of rain gauges. Differences between ERA-Interim and GPCPv2.1 (not shown) are generally smaller.

Overall, both the 21-year means and the decadal changes are similar in the two datasets. ERA-Interim shows more rainfall than GPCC in most of the Northern Hemisphere, and in parts of South America. The large differences over central Africa are indicative of higher uncertainties

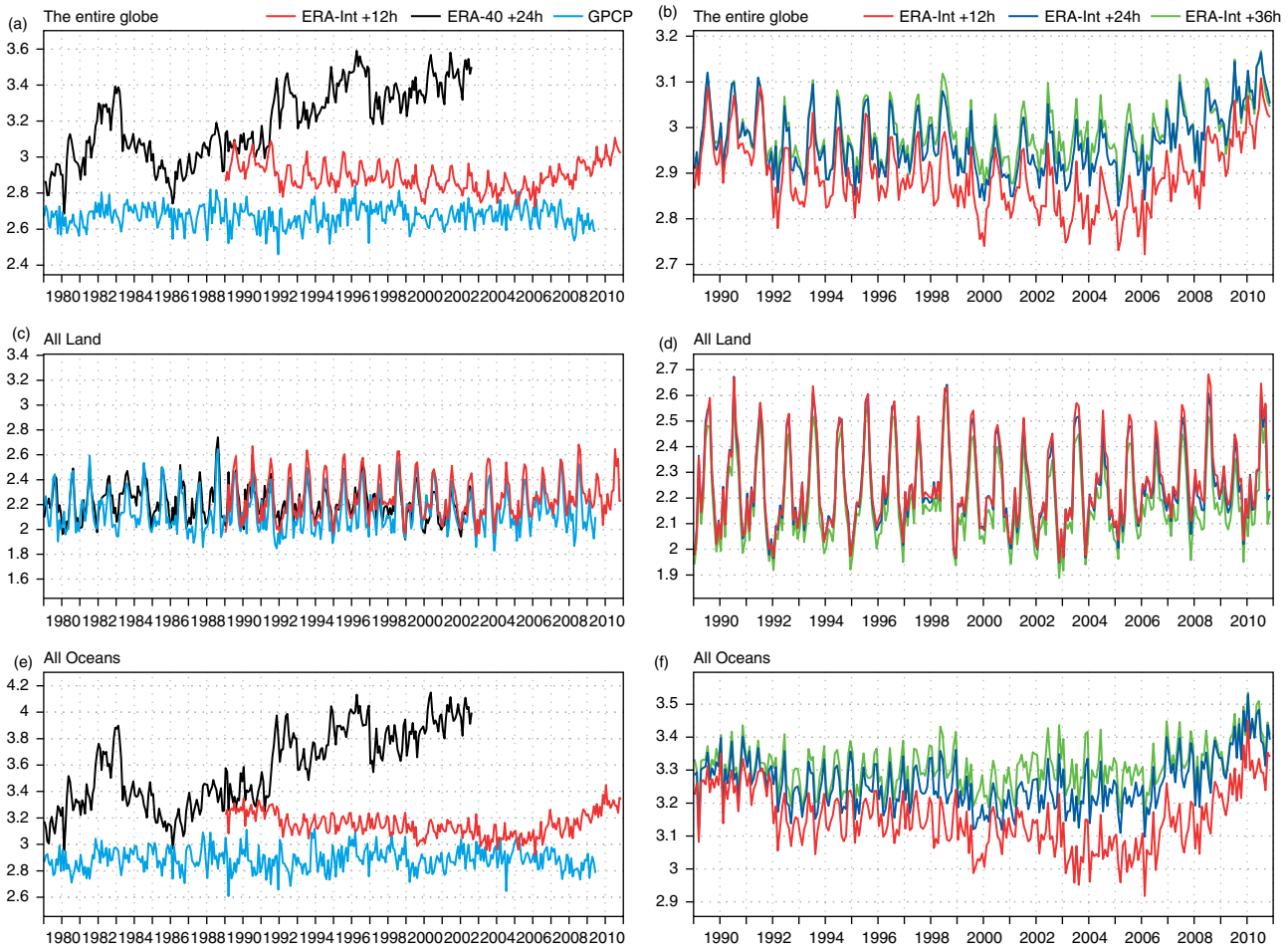

Figure 24. Monthly averaged precipitation and spin-up/spin-down effects: precipitation estimates (mm day^{-1}) for 1979–2010 from ERA-Interim (red), ERA-40 (black), and GPCP (blue), averaged for (a) the entire globe, (b) all land locations, and (c) all ocean locations. Results for ERA-Interim are based on accumulated rainfall in the initial 12-hour forecast segment; for ERA-40 the 12–24-hour segment was used. (d, e, f) show, for 1989–2010 and for ERA-Interim only, corresponding estimates obtained from the initial 12-hour segment (red; identical to (a, b, c)), the 12–24-hour segment (blue), and the 24–36 hour segment (green).

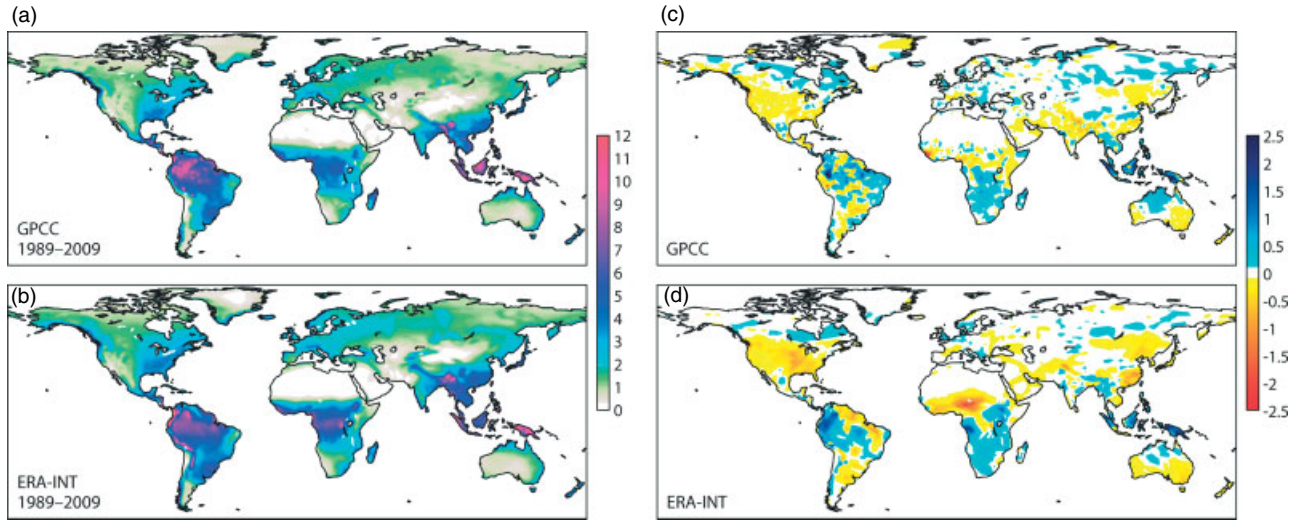

Figure 25. Average precipitation rates (mm day^{-1}) for the 21-year period 1989–2009 from (a) GPCC and (b) ERA-Interim, and decadal change in precipitation from (c) GPCC and (d) ERA-Interim, defined as the difference between the 2000–2009 average and the 1990–1999 average for each dataset.

due to the sparse radiosonde coverage there, particularly during the first decade of the averaging period. This is also evident in the decadal change estimates, where ERA-Interim probably overestimates the decrease in rainfall in the central African region. A possible explanation is the presence of a substantial warm bias in the model associated with underestimated aerosol optical depth in the region. This, combined with the improved data coverage during the second decade, results in apparent cooling accompanied by reduced precipitation.

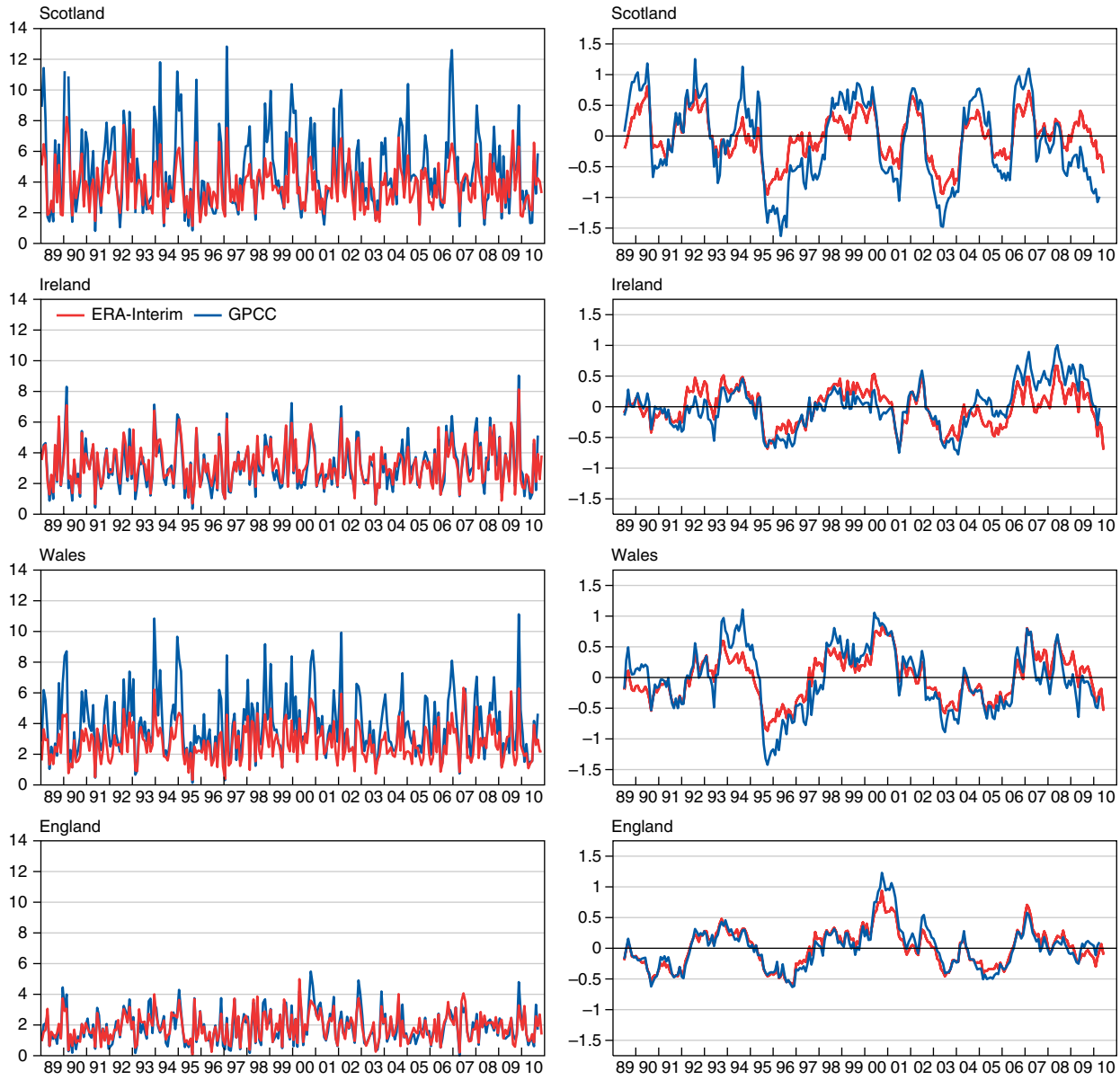

Figure 26. Monthly averaged precipitation rates (mm day^{-1}) (left) and corresponding anomalies (right), from ERA-Interim (red) and GPCC (blue), for four different 1° × 1° areas over the British Isles.

Figure 26 compares precipitation rates and anomalies from ERA-Interim with GPCC, at the 1° × 1° resolution of the GPCC gridded data product, at four different grid points located on the British Isles. The comparison uses monthly averages of ERA-Interim precipitation estimates interpolated to the GPCC grid. Anomalies were computed for each dataset relative to their own 21-year averages. Two locations, in Wales and Scotland, are in mountainous areas; the other two are in lowland regions of western Ireland and southern England. Month-to-month variability of ERA-Interim rainfall compares quite well with GPCC estimates, as do the anomalies. ERA-Interim produces lower values for rainfall maxima in mountainous regions.

Further evaluation of ERA-Interim precipitation estimates over land, including comparisons with GPCC as well as the 4 km-resolution PRISM dataset for the USA, is described in Balsamo *et al.* (2010). Bromwich *et al.* (2011) have made a detailed study of the realism and consistency of the hydrological cycle for the Antarctic region in ERA-Interim and several other recent reanalyses. While pointing to large uncertainties in this region, they conclude that ERA-Interim likely provides a realistic depiction of precipitation changes in the high southern latitudes during 1989–2009.

5.2.2. *Stratospheric circulation*

Difficulties with the stratospheric circulation in ERA-40 associated with an overly strong Brewer–Dobson circulation have been described in some detail by Uppala *et al.* (2005). The problem was clearly related to the assimilation of radiance observations, as could be seen, for example, from the intensification of wintertime ozone transport when stratospheric radiances began to be assimilated. Primarily as a result of biases in the model, the analysis produced large increments in the upper stratosphere with spurious oscillatory features propagating downward, especially during winter in the polar regions (Dee, 2005). In ERA-Interim, the underlying systematic differences among observations and the model background are quite effectively reduced by the variational bias correction scheme. This has

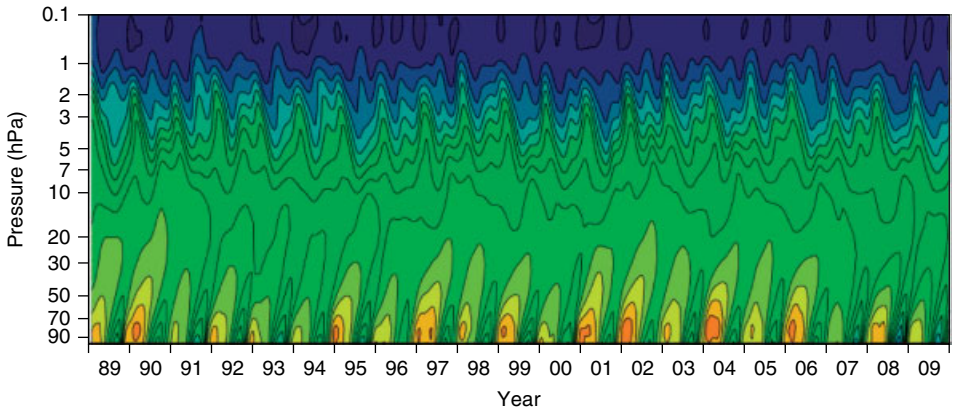

Figure 27. Stratospheric specific humidity averaged over the equatorial band (10°N–10°S) from ERA-Interim: 61-day running averages with contour interval of 0.2 mg kg^{-1}.

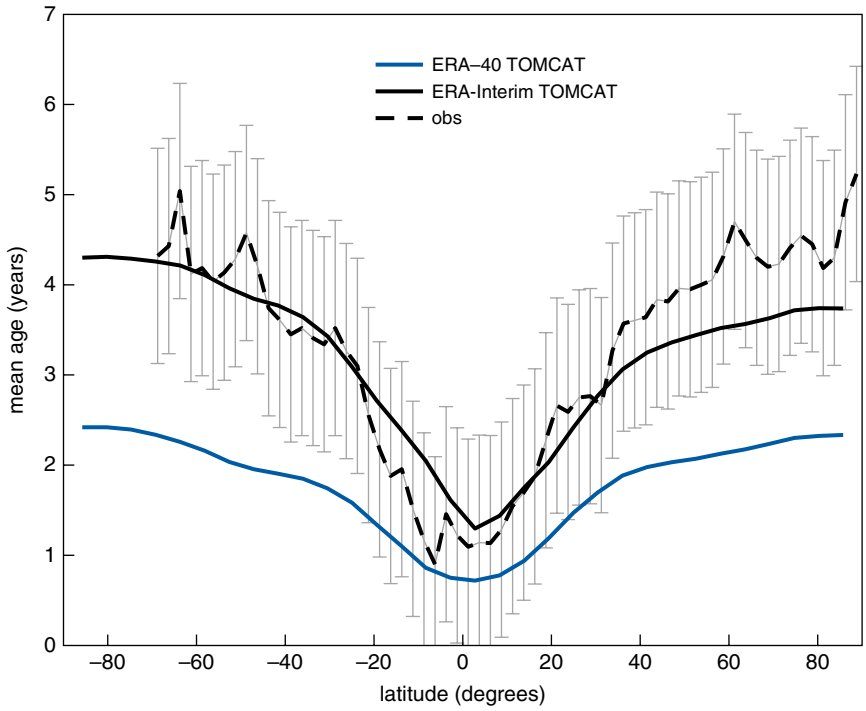

Figure 28. Mean age-of-air (years) at a height of 20 km, as a function of latitude, as determined from transport simulations with the TOMCAT model driven by winds from ERA-40 (solid blue) and ERA-Interim (solid black), in this case corresponding to the year 2000. Independent estimates from aircraft are also shown (dashed black) with error bars (grey).

led to smoother increments and a better vertical consistency in the temperature analysis, as shown, for example, in Figures 8 and 9 of Dee and Uppala (2008). The variational bias adjustments for stratospheric radiance observations have also led to a much improved temporal consistency on interannual time-scales, as briefly discussed in the next section.

As an example of the so-called *stratospheric tape-recorder*, Figure 27 shows the evolution in time of stratospheric humidity in ERA-Interim, for the tropical band between 10°N and 10°S. As in ERA-40, no change to the stratospheric humidity is made by the ERA-Interim analysis other than removal of any supersaturation. This means that the distribution of humidity is determined primarily in the forecast model, by tropospheric exchange, by upper-level moistening due to methane oxidation and by advection, with some loss due to precipitation in the cold polar night. In the tropical stratosphere, relatively dry air introduced at the tropical tropopause in boreal winter, and relatively moist air introduced in boreal summer, are transported slowly upwards. This transport was much too strong in ERA-40, with successive layers of moist and dry air reaching 30 hPa only some three months after they leave 100 hPa (Figure 19(a) in Uppala *et al.*, 2005). In ERA-Interim, the rate of upward transport is approximately halved, as indicated by the slope of contours in the lower stratosphere in Figure 27, and the attenuation of the signal as height increases is reduced. The upward transport remains, however, almost twice as fast as indicated by observations (e.g. Schoeberl *et al.*, 2008).

Mean age-of-air diagnostics, which measure the average time of residence in the stratosphere of an imaginary parcel of air, are quite sensitive to the strength of the Brewer–Dobson circulation. These diagnostics provide a useful measure of the quality of reanalysed wind fields and, specifically, of their suitability for driving models of stratospheric chemical transport and stratosphere/troposphere exchange. Figure 28

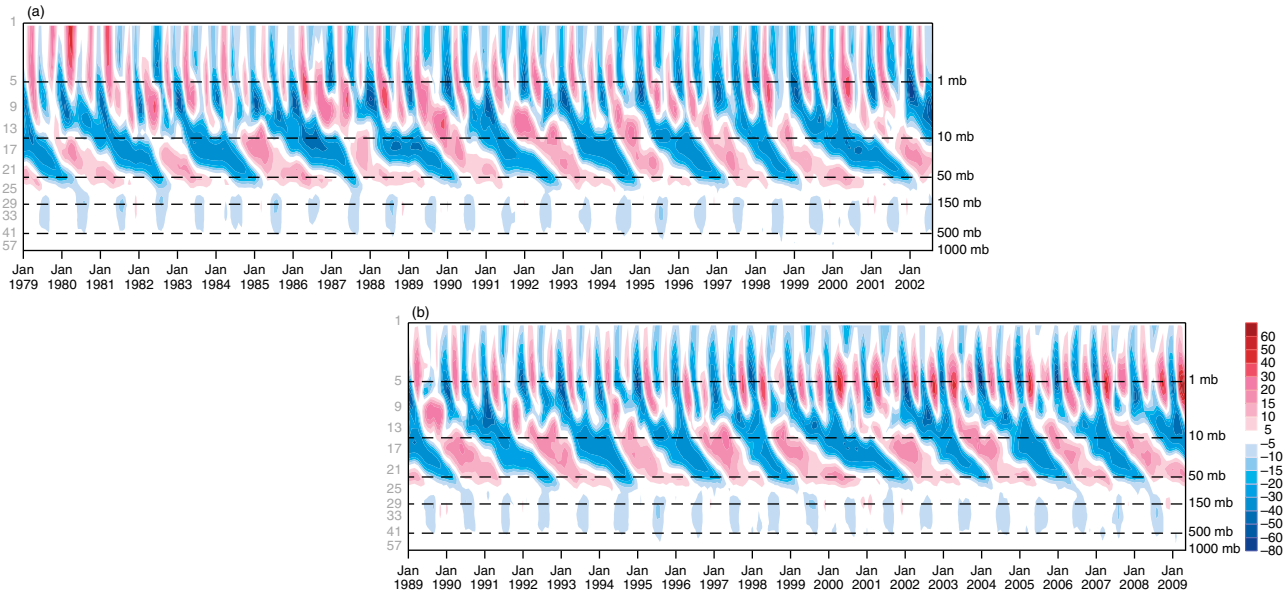

Figure 29. *u*-wind component ($m\,s^{-1}$) averaged monthly and between 2°N and 2°S, from (a) ERA-40 and (b) ERA-Interim.

demonstrates the improvements achieved in ERA-Interim in this respect. The curves represent estimates of mean age-of-air at 20 km altitude as a function of latitude, obtained with the TOMCAT chemical transport model (Chipperfield, 2006) when using reanalysed winds from ERA-40 and ERA-Interim. For reference, the figure includes observational estimates derived from *in situ* ER-2 aircraft measurements of CO_2 (Andrews *et al.*, 2001) and SF_6 (Ray *et al.*, 1999) along with indications of uncertainty. Monge-Sanz *et al.* (2007) further discuss this diagnostic and how it is computed. The close match between the TOMCAT age-of-air estimates and those obtained from observations reflects the improved quality of the mean stratospheric circulation in ERA-Interim –notwithstanding the remaining difficulties in representing the slow vertical transport of moisture, as discussed in the previous paragraph.

Representation of the QBO in ERA-40 was already quite good, as has been demonstrated using wind observations from radiosondes and independent rocketsonde data (Randel *et al.*, 2004). Figure 29 shows the zonal component of wind averaged for the equatorial band between 2°N and 2°S, for ERA-40 and ERA-Interim. Both reanalyses provide a similar and consistent representation of the temporal evolution and vertical structure of the QBO.

5.2.3. *Temporal consistency*

Reanalysis of the last few decades is especially challenging because of the rapid development of the observing system that has taken place during this period. The introduction of many new satellite instruments and their expanding spatial coverage present opportunities for constructing an increasingly accurate and complete description of the global atmosphere. However, the changing observing system can also lead to discontinuities and spurious low-frequency variations in the reanalysis. These may be caused by the assimilation of biased observations, or, alternatively, by introducing new observations that constrain previously unobserved components of model bias.

Good progress was made in ERA-Interim with the treatment of biases in observations, including surface pressure reports from land stations, ships, and buoys, and temperature data from radiosondes (section 4). Bias estimates for the vast majority of satellite data are generated by the variational analysis itself (section 2.1.2). The assimilation system makes adjustments to the radiance observations that depend on the sum total of information available, from all observations combined and from the forecast model. Dee and Uppala (2009) have demonstrated the effectiveness of this approach for correcting instrument biases in radiance data, and discussed implications for the representation of climate signals in ERA-Interim. The adaptive bias corrections produced in ERA-Interim illustrate the complex and time-dependent character of systematic errors present in the data. Large bias variations in tropospheric MSU radiances, resulting from on-board calibration errors associated with orbital drift, were accurately identified and corrected in the reanalysis. The method also accounts for systematic errors in the fast radiative transfer models used to interpret the data, which can be substantial. For example, the bias adjustments generated for HIRS data subsequent to the Pinatubo eruption properly compensated for the lingering effects of volcanic aerosols, thereby avoiding some of the problems encountered in ERA-40.

These improvements have led to a much better intercalibration of the various components of the observing system, and consequently to a more consistent representation of climate signals in the reanalysis. For globally averaged temperatures near the surface, in the lower troposphere, and in the lower stratosphere, monthly variability and trends in ERA-Interim compare favourably with estimates from traditional, observation-only climate datasets (Willett *et al.*, 2010). Upper-tropospheric temperatures have been slightly overestimated in recent years due to the assimilation in ERA-Interim of increasing numbers of warm-biased temperature reports from aircraft (Dee and Uppala, 2009). In a detailed study of low-frequency variability in near-surface temperature and relative humidity, Simmons *et al.* (2010) demonstrated very close agreement with independently derived analyses of monthly station temperature data (from CRUTEM3) and synoptic humidity observations

(from HadCRUH) at the locations where such analyses are available. Variability and trends in precipitation over land, as already discussed in section 5.2.1 and in more detail by Simmons *et al.* (2010), are remarkably accurate in ERA-Interim.

Detrimental effects of model biases on trends in ERA-Interim are clearly evident in precipitation estimates over ocean, as explained in section 5.2.1. Problems of a similar nature occur in the upper stratosphere, where mean temperatures are visibly affected by the transition in 1998 from SSU to AMSU-A as the dominant stratospheric observing system. The highest peaking channels of these instruments constrain slightly different vertical integrals of model error (Kobayashi *et al.*, 2009), leading to shifts in time series of global mean temperature at 5 hPa and higher (Figure 20 in Dee and Uppala, 2008). Assimilation of GPSRO data from December 2006 has provided new information about small-scale vertical structures in upper-tropospheric and stratospheric temperature profiles (Healy and Thépaut, 2006), which are not well resolved by the forecast model, nor by other satellite observations. This has led to increased variability and systematic warming of up to 0.1–0.2 K at the tropopause and in the lower stratosphere (Poli *et al.*, 2010).

ERA-Interim used a succession of different SST and SIC data products, as described in section 3.4. A shift to lower SST values, averaging globally to approximately 0.15 K, occurred in 2001 when NCEP operational products began to be used. Simmons *et al.* (2010) discuss implications of SST uncertainties associated with this change in their analysis of trends in temperature, humidity, and precipitation. Hersbach and Drusch (2009) have investigated the impact of the switch from the NCEP RTG product to OSTIA, which was implemented in ERA-Interim in February 2009. Comparison of sea-surface temperatures in the two products did not expose clear differences in variability at the resolution used for ERA-Interim. Systematic differences between the two datasets are negligible in most areas around the globe. However, in the polar regions, OSTIA shows more sea-ice cover and higher SSTs, reaching as much as 10 K in summer at a few locations off the Siberian coast. Since the presence of sea ice can have a large impact on near-surface wind speed in a model forecast, ERA-Interim wind analyses in sparsely observed locations at high latitudes may well be affected by these differences.

5.3. *Other quality aspects*

5.3.1. *Ozone concentrations*

A thorough evaluation of the ERA-Interim ozone estimates has been presented by Dragani (2010a,b). To highlight improvements relative to ERA-40, here we briefly discuss a comparison against sonde measurements from the WOUDC archive (http://www.woudc.org). These high-quality but sparsely distributed data were not assimilated in ERA-Interim. However, a large variety of satellite observations were used, as discussed earlier and shown in Figure 15. A major change in the ozone observing system occurred with the introduction of data from the GOME instrument on ERS-2 in 1996. The impact of ozone profiles retrieved from these data on the quality of the ERA-Interim ozone assimilation is summarised in Figure 30.

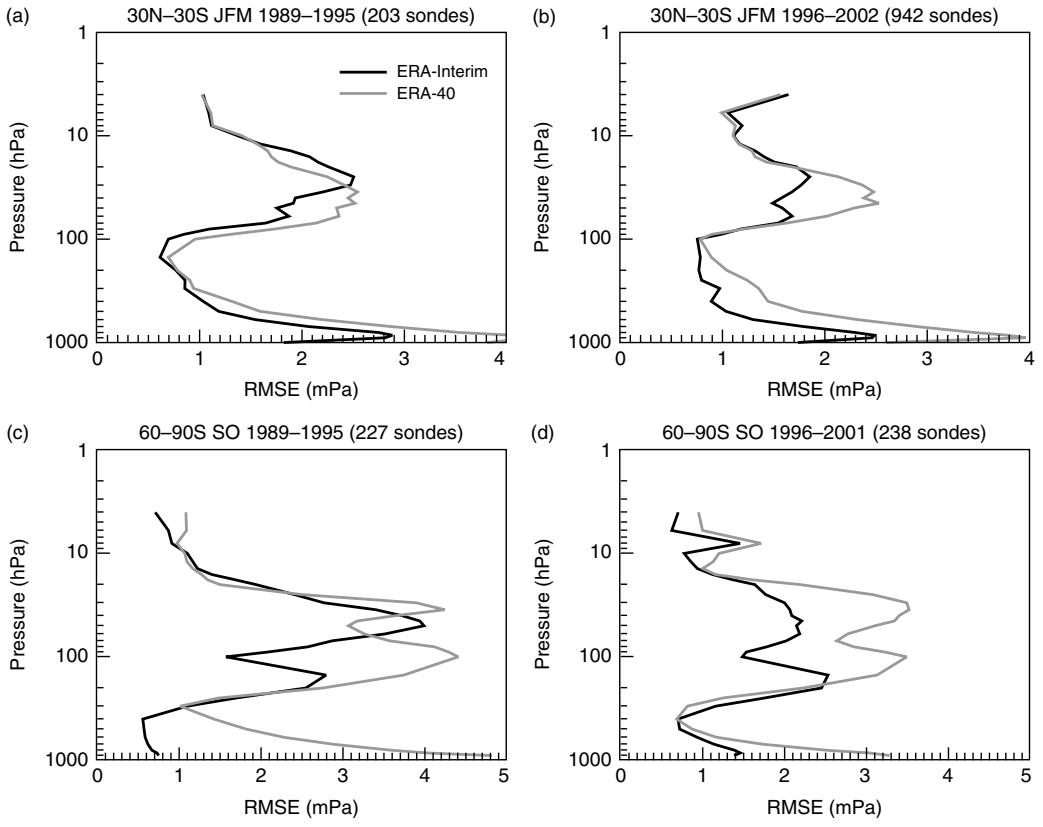

Figure 30. RMS errors (mPa) in ozone estimates from ERA-Interim (black) and ERA-40 (grey), relative to independent ozone sonde measurements, for (a, c) the pre-GOME period 1989–1995, and for (b, d) the period 1996–2002 when GOME data were actively assimilated. (a, b) show results for latitudes between 30°N and 30°S, during January, February, and March. (c, d) are for the southern high latitudes beyond 60°S during September and October. The number of sonde profiles used in each case is indicated at the top of each panel.

In general, the vertical distribution of tropical ozone concentrations compares well with the sonde measurements, although peak values tend to be underestimated. The assimilation of GOME profiles has improved the quality of the ERA-Interim tropical ozone reanalyses relative to ERA-40, particularly in the lower and middle parts of the stratosphere. Similar results are obtained for other seasonal averages.

Figure 30(c, d) show the ERA-Interim and ERA-40 fit to ozone sondes at high-latitude stations in the Southern Hemisphere, for the months September and October during the pre-GOME and GOME periods. Modelling of stratospheric ozone transport and depletion at these latitudes is particularly difficult during winter and spring. Problems are exacerbated by the scarcity of ozone observations retrieved from UV sensors, which do not perform during the polar night. The reanalyses tend to place the ozone maximum too high, with deeper depletion in the middle stratosphere just below the peak. ERA-Interim shows better agreement with ozone sondes than ERA-40 at most levels, and particularly in the troposphere. The assimilation of GOME data has further improved the ERA-Interim ozone analyses, in the upper troposphere and lower stratosphere.

The sonde comparisons combined with further results reported by Dragani (2010a,b) clearly demonstrate improvements in ERA-Interim ozone estimates relative to ERA-40. In some cases the improvements can be related directly to changes in data usage. The accuracy of the ozone estimates also benefit from a better characterisation of the stratospheric circulation in ERA-Interim, via the dependence of prognostic ozone on temperature in the Cariolle and Dèquè (1986) scheme (section 3.1.5). The consistency in time of the ozone reanalysis, however, is still far from adequate, due to biases in observations used, as well as shortcomings of the forecast model. Work is currently under way to improve the homogeneity of the ozone assimilation by applying variational bias adjustments to the various sources of ozone data used in the assimilation (Dragani and Dee, 2008). Development of more realistic error characterisations for the background estimates of 3D ozone fields produced by the forecast model is also required to further improve the ozone estimates in future reanalyses.

5.3.2. *Ocean-wave parameters*

The wave spectra produced in ERA-Interim are constrained by wave-height observations from space-borne radar altimeters on ERS-1, ERS-2, ENVISAT, and Jason-1 as described in section 4.6. Figure 31(a, c, e, g) show time series of the normalised standard deviations and Figure 31(b, d, f, h) of the biases for significant wave height, peak period, zero-crossing mean wave period, and 10 m wind speed. Verification data consist of quality-controlled *in situ* observations of sea-surface wave parameters and 10 m wind speed obtained from buoys, platforms, and weather ships (Bidlot *et al.*, 2002). Since the number of wave observation locations has steadily increased since the late eighties, and buoy locations vary from year to year, any apparent trends in the evolution of the statistics should be interpreted with caution.

A marked improvement of ERA-Interim relative to ERA-40 is clearly visible for all wave parameters, including the removal of a known problem with the assimilation of ERS-1 data in ERA-40 in 1992–1993 (Sterl and Caires, 2005). The overall quality of the wave parameters in ERA-Interim is at the level of the operational analysis of 2005; this has been confirmed by comparing altimeter wave height data from ENVISAT to background estimates from ERA-Interim and operations (not shown). ERA-Interim wind-speed analyses do not inherit the recent improvements in the operational wind analyses, which are due to resolution increases (to T511 in late 2000 and T799 in early 2006) and other upgrades of the atmospheric forecast model.

5.3.3. *Snow depth*

Several problems with the analysis of snow in ERA-Interim have been uncovered, related to shortcomings in the analysis method as well as errors in data processing. Defects in the Cressman analysis scheme have led to spurious snow-free patches in sparsely observed situations. This has been addressed in a recent upgrade of the operational ECMWF forecasting system, but continues to affect ERA-Interim. In addition, an error in the pre-processing at ECMWF of IMS snow cover data (section 4.6) introduced significant shifts in the locations of the data as well as in the land-sea mask and orography. As a result, some locations in coastal sea and lake areas were falsely identified as ocean locations, which are always defined as snow-free. The geolocation errors have affected the ERA-Interim snow analyses from 1 July 2003, when the IMS product began to be used, until 23 February 2010, when the pre-processing software was corrected.

Some degradation of the accuracy and temporal consistency of reanalysed snow depth, mainly in coastal areas, has resulted from these errors. An assimilation experiment was conducted for the period December 2009–January 2010 specifically to study the impact of the IMS geolocation problem on ERA-Interim products. It was found that 2-day forecasts of 2 m temperature in the affected regions were significantly higher, which is consistent with the fact that erroneous geolocation often led to underestimates of snow coverage. No systematic impact on analysed near-surface temperatures was found, most likely as a result of the observational constraints on the atmospheric analysis during the data-rich ERA-Interim period. Nevertheless, fewer such constraints are available from station observations and radiosondes at high latitudes, where, additionally, background fields may be affected by the poor representation of snow. The accuracy of reanalysed near-surface temperatures is therefore expected to be lower as a result.

5.3.4. *Global conservation properties*

Berrisford *et al.* (2011) have studied the atmospheric mass, energy, and angular momentum budgets in ERA-Interim and ERA-40. Data assimilation methods are not designed to conserve these properties, which therefore provide interesting and useful diagnostics of global quality aspects of the reanalyses. Generally, budget closure is better in ERA-Interim than in ERA-40. Effects of the improving global observing system can be clearly discerned from these diagnostics, particularly in ERA-Interim.

Global dry mass, which is computed from estimates of surface pressure and total column water vapour, is very accurately conserved both in ERA-Interim and ERA-40.

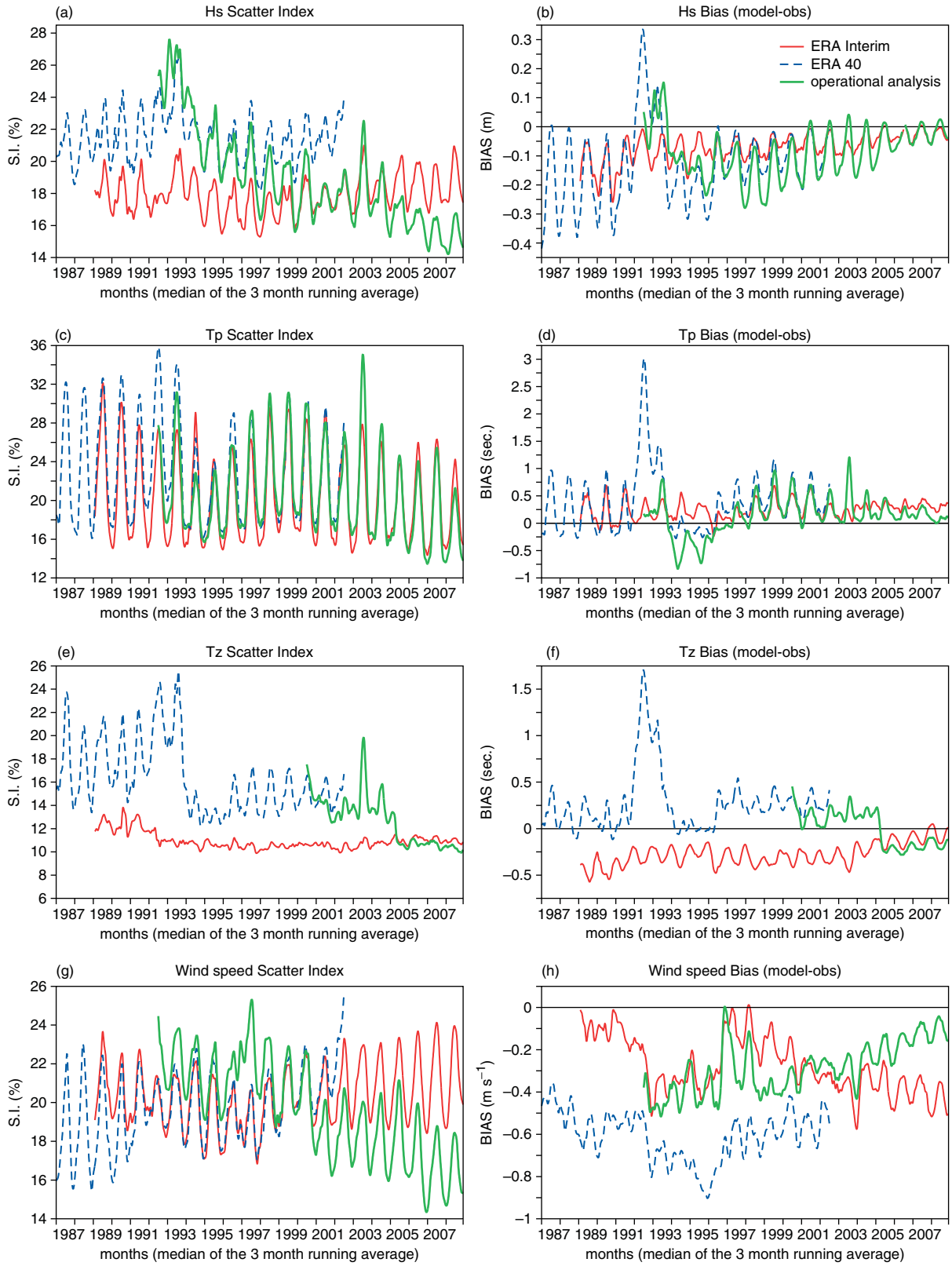

Figure 31. Monthly time series of scatter index (left) and bias (right) relative to *in situ* observations, for ERA-40, ERA-Interim, and operational analyses of (a, b) significant wave height, (c, d) peak period, (e, f) zero-crossing mean wave period and (g, h) 10 m wind speed. The scatter index is defined as the normalised standard deviation of the difference with respect to the *in situ* mean value. Bias is the mean difference between model and observations. Wind observations were adjusted to 10 m using a neutral wind profile. A 3-month running average was used to compute the statistics.

During 1989–2008, the standard deviation of the monthly mean global dry mass from its annual climatology is only 0.0074% for ERA-Interim, compared to 0.0058% for ERA-40 (during 1989–2001). The slightly larger value for ERA-Interim is due to a slow variation in dry mass during the period 1994–2004, also visible in the global mean surface pressure. An explanation for this variation, which is almost certainly spurious, remains to be found. On the other hand, conservation in the mean annual cycle is better in ERA-Interim than in ERA-40.

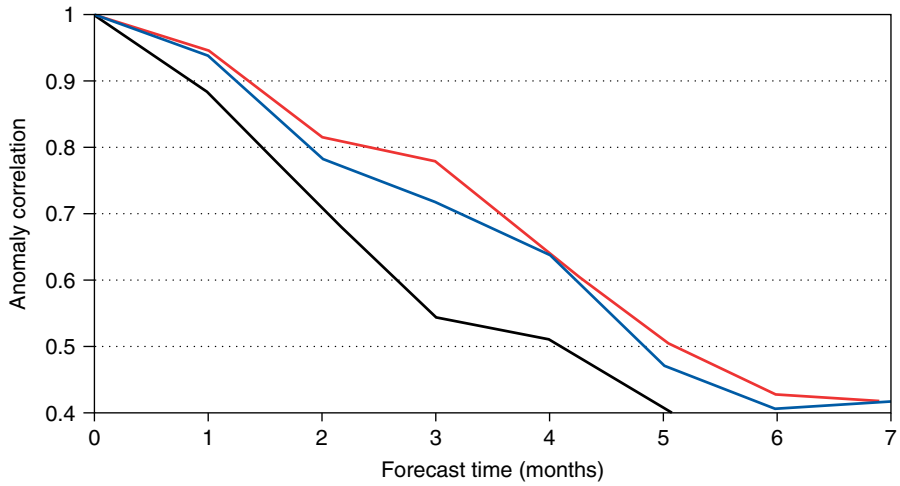

Figure 32. Anomaly correlations for seasonal forecasts of SST in the equatorial Atlantic. Ocean initial conditions for the forecasts were obtained using surface fluxes from ERA-Interim (red) or from a combination of ERA-40 and the ECMWF operational archive (blue). For comparison, the skill of persistence forecasts is also shown (black). Results are based on 71 forecasts issued three months apart during the period April 1989 to October 2006.

Further evaluation of the dry mass budget reveals inconsistencies between cross-equatorial mass fluxes derived from direct estimates of hemispheric mass tendencies and alternative estimates that use divergent tropical winds. The differences are a factor of 2–3 smaller in ERA-Interim than in ERA-40, and they reduce with time as the observing system improves. Given that the global distribution of atmospheric mass is rather accurately constrained by satellite observations in both reanalyses, the improved consistency in ERA-Interim is most likely due to a better representation of tropical divergent winds.

The energy balance at the top of the atmosphere in ERA-Interim has improved, with an estimated energy loss of 1.2 W m^{-2} (7.4 W m^{-2} for ERA-40). However, the energy balance at the surface boundary is poor in ERA-Interim, with a global value of 6.9 W m^{-2} (3.8 W m^{-2} for ERA-40). This degradation occurs primarily over oceans and is associated with an increase in net solar radiation there. Over land the surface energy balance actually improves in ERA-Interim, to 0.5 W m^{-2} (1.3 W m^{-2} for ERA-40).

Kållberg (2011) suggests that the model clouds are the major contributor to the imbalance in surface energy, based on a correspondence between spin-up/spin-down of cloudiness and of the net energy fluxes[†]. Cloud cover typically increases by as much as 2.5% during the first 24 h of a model integration, except over subtropical upwelling areas and tropical rain forests where the total cloud amount reduces with time. Surface energy fluxes show a similar behaviour, with a global spin-down but spin-up in the Subtropics, the Atlantic ITCZ, and the high-latitude southern oceans. The reverse occurs at the top of the atmosphere, where the global net energy fluxes in ERA-Interim spin up.

5.3.5. *Surfaces fluxes over oceans*

The achievable skill of seasonal forecasts depends to a large extent on the quality of estimated fluxes of heat, momentum, and water at the atmosphere–ocean interface. Balmaseda and Mogensen (2010) have compared estimates of SST and sea level obtained from several ocean simulations for the period 1989–2001, using atmospheric fluxes from both ERA-40 and ERA-Interim.

Excessive solar radiation from ERA-Interim, already noted in previous paragraphs, causes a net heat flux into the ocean, especially in the convective tropical areas. On the other hand, Balmaseda and Mogensen (2010) show that the seasonal and interannual variability as well as the spatial structure of incoming solar radiation from ERA-Interim have improved significantly compared to ERA-40. They also found that ocean simulations forced by ERA-Interim data produce a more realistic interannual variability of sea-level anomalies, based on a comparison with altimeter-derived estimates.

The quality of ERA-Interim surface fluxes was further evaluated by using them to initialise the ocean component of the ECMWF seasonal forecasting system (S3; Balmaseda *et al.*, 2008). A consistent positive impact on the skill of the seasonal forecast of tropical SST was found, with the greatest improvements occurring in the tropical Atlantic. Aggregated over the tropical North Atlantic, the tropical South Atlantic and the equatorial Atlantic regions, reductions in SST anomaly correlations during the first three months of integration are statistically significant to 5%. To illustrate, Figure 32 shows average anomaly correlations as a function of lead time, for a series of SST forecasts over the equatorial Atlantic region, when using surface fluxes from either ERA-Interim or ERA-40. The positive impact of ERA-Interim fluxes is evident, in experiments with and without ocean data assimilation.

5.3.6. *Madden–Julian Oscillation*

The MJO is the dominant mode of intraseasonal variability in the tropical atmosphere. It has a significant impact on Indian and Australian monsoons (e.g. Murakami, 1976; Hendon and Liebmann, 1990), the onset and development of ENSO events (e.g. Kessler and McPhaden, 1995), tropical cyclone genesis (e.g. Maloney and Hartmann, 2000), and also on extratropical weather (Ferranti *et al.*, 1990; Cassou, 2008). It is therefore important for a monthly and seasonal forecasting system to have skill in predicting the onset and evolution of MJO events.

[†]Section 5.2.1 gives an explanation of model spin-up/spin-down.

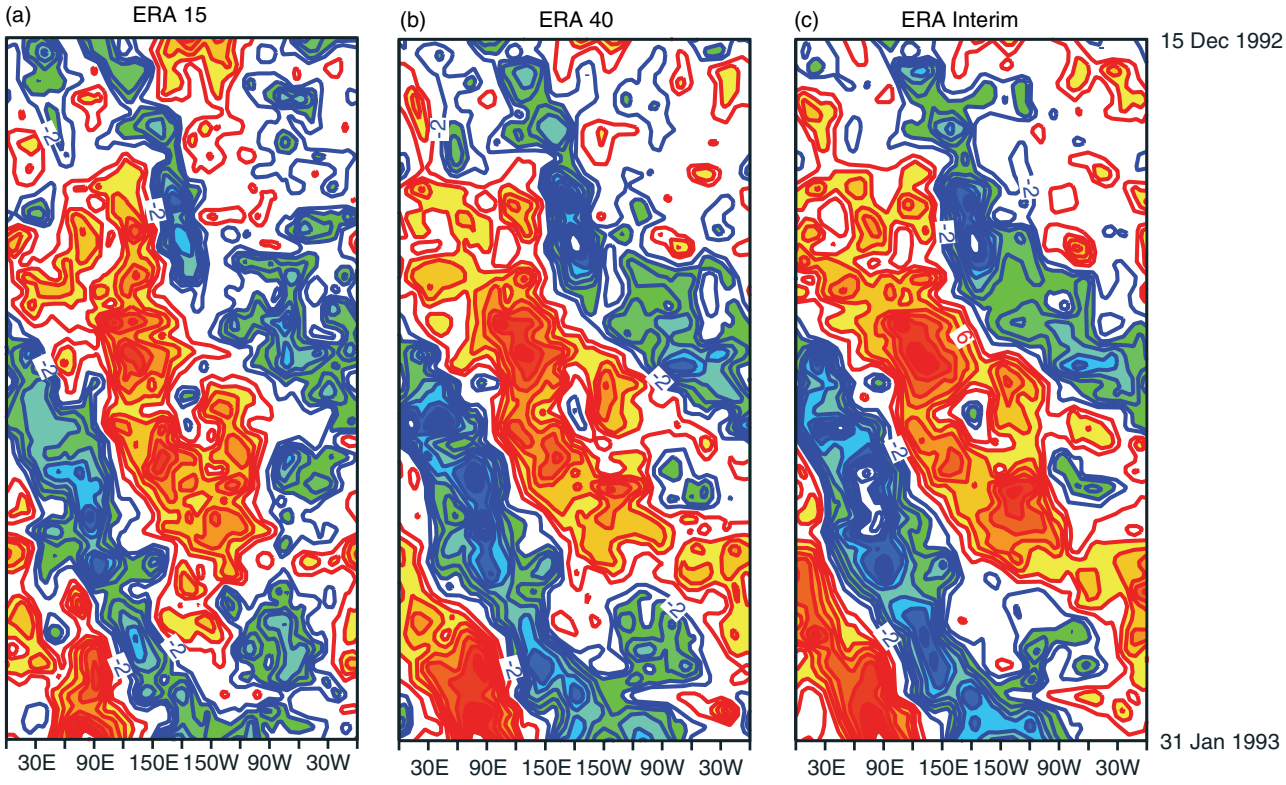

Figure 33. Hovmöller diagram of velocity potential averaged between 10°S and 10°N (contour interval $3 \times 10^6 m^2 s^{-1}$) for the period 15 December 1992 to 31 January 1993, from (a) ERA-15, (b) ERA-40 and (c) ERA-Interim. Negative values are blue/green, positive values are yellow/red.

To evaluate the skill of the ECMWF monthly forecasting system (Vitart, 2004) in predicting MJO events, 32-day coupled ocean–atmosphere integrations using a five-member ensemble have been performed for each day between 15 December 1992 and 31 January 1993 (46 ensemble integrations); this period corresponds to the IOP of TOGA-COARE. The IFS version used in this study is Cy32r2.

Figure 33 presents a Hovmöller diagram of velocity potential at 200 hPa over the period 15 December 1992 to 31 January 1993 from ERA-15, ERA-40 and ERA-Interim, showing increases in amplitude of the MJO in successive reanalyses.

In order to evaluate the impact of those differences on the MJO forecasts, the serial experiment described above has been repeated three times, using atmospheric initial conditions from ERA-15, ERA-40 and ERA-Interim. The MJO is diagnosed in these integrations using a method based on the technique of Wheeler and Hendon (2004). Combined EOFs of OLR, 100 hPa velocity potential and 850 hPa wind averaged over 10°N–10°S are calculated using ECMWF operational data between 2002 and 2004. The EOF analysis is performed on the anomalies relative to the seasonally evolving climatology. The first two EOFs, which represent 18% and 17% of the variance respectively, describe variations associated with the MJO. More details about the experiment settings and the MJO diagnostic can be found in Vitart *et al.* (2007).

Figure 34 shows the correlation of the mean of PC1 and PC2 time series computed from each ensemble mean forecast with their respective analyses (e.g. forecasts using ERA-Interim are verified against ERA-Interim). The 0.6 anomaly correlation is reached around day 13 when the model is initialized with ERA-15. It reaches 0.6 by day 15

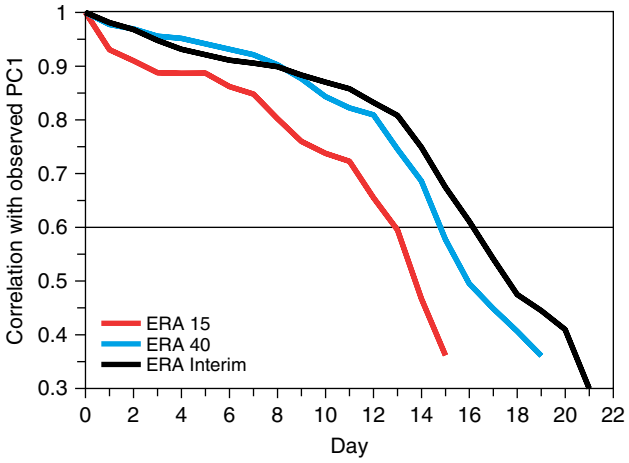

Figure 34. Correlation of the ensemble mean forecast of PC1 and PC2 time series with the time series obtained with the reanalysis used to produce the initial conditions. Skill curves are shown for forecasts initialized with ERA-15 (red), ERA-40 (blue), and ERA-Interim (black).

with ERA-40, and by day 16 with ERA-Interim. Those differences are statistically significant at 5% according to the Wilcoxon-Mann-Whitney test (Wonacott and Wonacott, 1997). Verifying all forecasts against the same reanalysis (ERA-15 or ERA-40 or ERA-Interim) instead of verifying against their respective reanalyses does not change the main conclusion that the skill of the model to predict the MJO improves from ERA-15 to ERA-40 and from ERA-40 to ERA-Interim. The improvement is even larger when all forecasts are verified against ERA-Interim. Those results suggest that the improvements in quality of the ECMWF reanalysis are conducive to improved MJO forecasts.

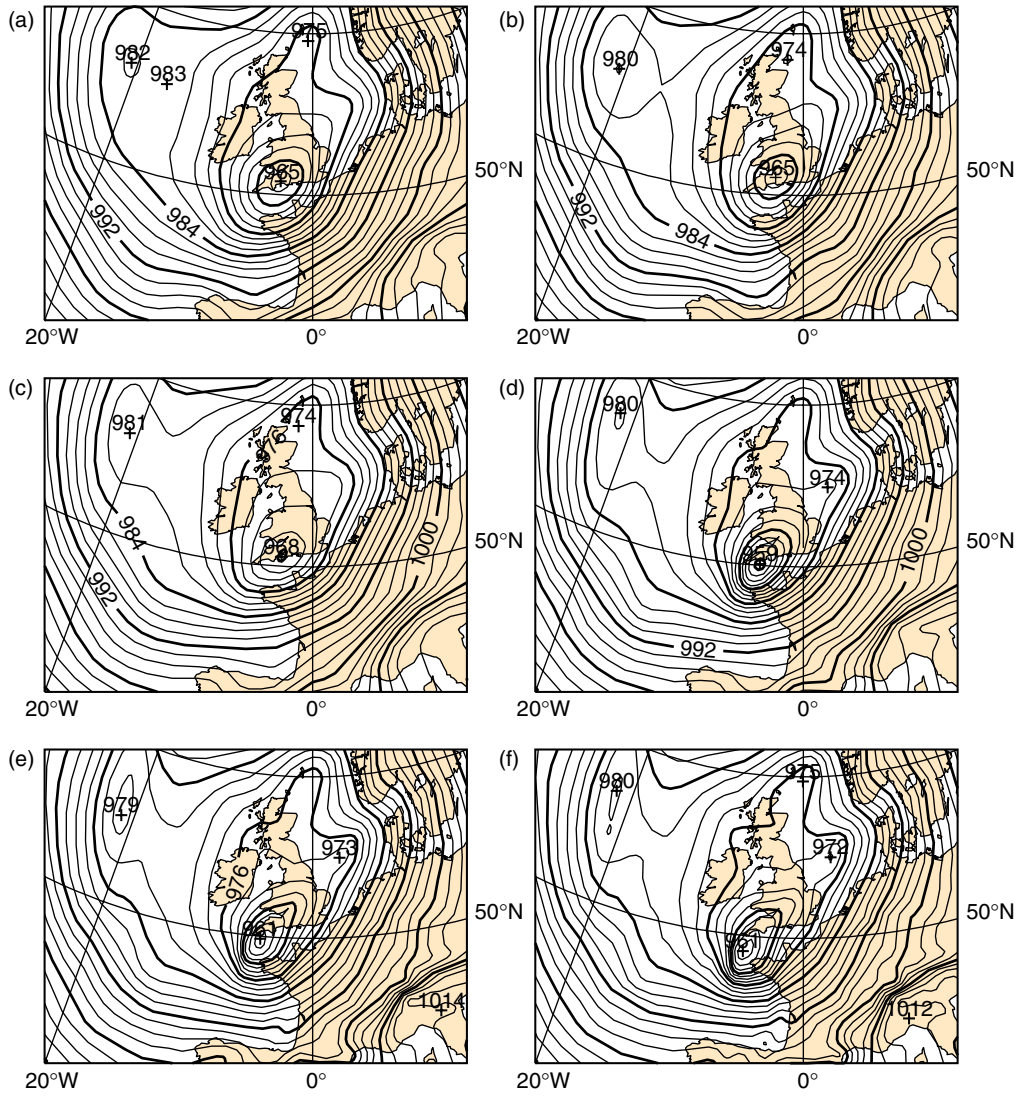

Figure 35. Mean-sea-level pressure analysis at 0000 UTC on 16 October 1987: (a) ECMWF operations at the time, (b) ERA-15, (c) ERA-40, (d) ERA-Interim, (e) ERA-Interim at T511, and (f) ERA-Interim at T799.

5.3.7. An extreme weather event

The Great Storm of 16 October 1987, which caused enormous damage in parts of southern England and France, provides a good test of the ability to represent extreme weather events in a global reanalysis at relatively moderate resolution. Operational forecasts at the time failed to provide adequate short-range guidance in terms of the track, speed and intensity of the storm, even though indications of its development had been present in the medium-range forecasts (Morris and Gadd, 1988; Jarraud *et al.*, 1989). This case presents an interesting challenge for data assimilation, since a rapid intensification on 15 October took place over the Bay of Biscay, which adjoins the data-dense continent but is downstream of an area with only few conventional data. It has been demonstrated (Shutts, 1990) that a mesoscale forecast model can simulate the evolution of the storm when initialised using late-arriving aircraft reports, which indicates the sensitivity to the use of observations in the data assimilation. In order to capture this event in a global reanalysis, the data assimilation system must preserve the baroclinic structure of the developing storm, while making use of abundant TOVS and SSM/I radiance data over the Atlantic, in a manner consistent with the information from conventional observations over land.

Jung *et al.* (2004) describe reanalysis experiments of the Great Storm using the version of the forecast model used for ERA-40, but with enhanced resolution (T511) and a 4D-Var analysis scheme, assimilating conventional observations and AMV wind data. Only minor improvements in forecast quality could be obtained, even when a more recent version of the forecast model was used. To test the ability of the ERA-Interim system to capture the evolution of this event, and to explore the potential performance of higher-resolution reanalyses, the month of October 1987 has been reanalysed using the ERA-Interim configuration with increasing spatial resolutions (T255, T511 and T799). Radiance data together with all other available observations were included in these assimilation experiments.

Six analyses of mean sea-level pressure valid for 16 October 1987 at 0000 UTC over Europe are shown in Figure 35, produced by ECMWF operations at the time (OPER), and reconstructed from ERA-15, ERA-40, ERA-Interim (T255), ERA-Interim (T511) and ERA-Interim (T799). According to a hand-drawn analysis, the centre

of the storm was located just north of Brittany, with an estimated central pressure of 953 hPa. Both the location and intensity of the storm are incorrect in the OPER, ERA-15 and ERA-40 analyses, all of which are at relatively low resolution and based on static methods (OI for OPER and ERA-15, and 3D-Var for ERA-40). The three ERA-Interim variants, each using 4D-Var, are able to intensify the storm, and with increasing model resolution improve its location to a nearly perfect match at T799.

Any analysis can be forced to draw more closely to observations by manipulating error statistics. A more meaningful indication of the quality of the data assimilation is the accuracy of the background forecasts. For OPER, ERA-15 and ERA-40, the background forecasts (not shown) are not sufficiently intense and have large location errors; the storm moves too fast. In ERA-Interim (at T255) the storm intensity is improved but it still moves too fast. Both aspects are improved in the T511 and T799 ERA-Interim assimilations. The impact of increasing the resolution of the ERA-Interim system is more clearly visible in the 36-hour forecasts of the storm (not shown). Results could perhaps be further improved by using a better representation of flow-dependent background errors.

6. Conclusions and future outlook

The ERA-Interim project was conceived to serve as a bridge between the highly successful ERA-40 atmospheric reanalysis, completed in 2002, and future generations of reanalyses to be produced at ECMWF. A key objective was to address several difficult data assimilation problems encountered in ERA-40, mostly related to the use of satellite data. Good progress was achieved in this regard, resulting in an improved representation of the hydrological cycle, a more realistic stratospheric circulation, and better temporal consistency on a range of time-scales. In many respects, the quality of ERA-Interim products has exceeded expectations; the project was conducted with limited resources, and only modest efforts were made to use the best input data available. The success of the project speaks volumes about the remarkable achievements in modelling and data assimilation realised at ECMWF in recent years.

Further progress is possible in several key areas. For example, unphysical changes in global mean precipitation, while still present in ERA-Interim, are now well understood and will likely be reduced in future reanalyses. A great deal has been learned about assimilation of cloud- and rain-affected satellite radiances for numerical weather prediction since ERA-Interim began. Representation of moist physical processes in forecast models has improved, and the ability to usefully assimilate *in situ* observations of accumulated rainfall is now well within reach. Advances in the treatment of soil hydrology and snow in land-surface models can lead to improved fluxes of heat and moisture in the atmospheric boundary layer. This, in turn, will allow better use of satellite observations over land, and lead to more accurate representations of clouds and precipitation. Similar improvements in the exchange of fluxes between atmosphere and ocean can be achieved by including the ocean mixed layer in the forecast model. These, and other developments e.g. in modelling atmospheric composition, are necessary to further improve the consistency of global budgets of heat, water, and momentum, which are often used to identify remaining shortcomings in the climate quality of reanalysis products (Trenberth *et al.*, 2009).

ERA-Interim also provided an opportunity to improve the technical infrastructure for reanalysis production at ECMWF. This includes the handling of input observations, procedures for quality control and bias correction of the observations, and tools for monitoring the data assimilation system and its overall performance. Each of these aspects has affected the quality of the reanalysis. In particular, the variational bias adjustments for radiance data used in ERA-Interim have effectively anchored the various components of the satellite observing system to the *in situ* observations assimilated in the reanalysis. Naturally this presumes a certain level of confidence in the accuracy and stability of the observational record from radiosondes, aircraft, and surface stations. Further improvement of these data is essential to the success of future reanalyses. Monitoring of the bias adjustments and analysis departures produced by ERA-Interim has exposed remaining issues with the observations (e.g. warm biases in some aircraft reports) that can now be addressed. In the main, the bias adjustments generated in ERA-Interim have led to a more accurate representation of low-frequency variability than had been obtained in previous reanalyses. Confidence in trend estimates for temperature, both near the surface and at higher levels in the atmosphere, is such that ERA-Interim data are now used along with estimates from traditional, observation-only climate datasets to monitor climate change (Willett *et al.*, 2010).

An advantage of using reanalysis for climate change assessment is that the data provide a global view that encompasses many essential climate variables in a physically consistent framework, with only a short time delay. To illustrate, Figure 36 shows estimates of anomalies in near-surface temperature, humidity, and precipitation for the year 2010 obtained from ERA-Interim. Temperature anomalies reflect exceptional warming over most land areas, especially at high latitudes in the Northern Hemisphere. In most places higher (lower) temperatures are associated with more (less) moisture; exceptions are Australia, Brazil, southern Africa, and Russia, which experienced a well-publicised dry heat wave during the summer of 2010. Very large positive anomalies in humidity and precipitation over most of Australia, consistent with a strong La Niña, have caused extreme flooding in large parts of Queensland. The magnitudes of precipitation anomalies over the tropical oceans are somewhat overestimated, as can be inferred from Figure 24.

The quality of trends derived from reanalysis data needs to be verified for each geophysical quantity, e.g. by comparison with independent observations as has been done by Simmons *et al.* (2010) for surface temperature and humidity. Unfortunately, high-quality observations from conventional sources are often not available in sufficient numbers, especially in locations critical for climate change such as the Tropics and the polar regions.

The only fundamental limitation on our ability to describe the evolution of the atmosphere with increasing accuracy lies in the quality and availability of observations. Reanalysis of the past few decades, while technically complex, has been greatly facilitated by the global reach of the current observing system and the variety of information it provides. Satellite agencies and other data providers are now increasingly engaged in activities to

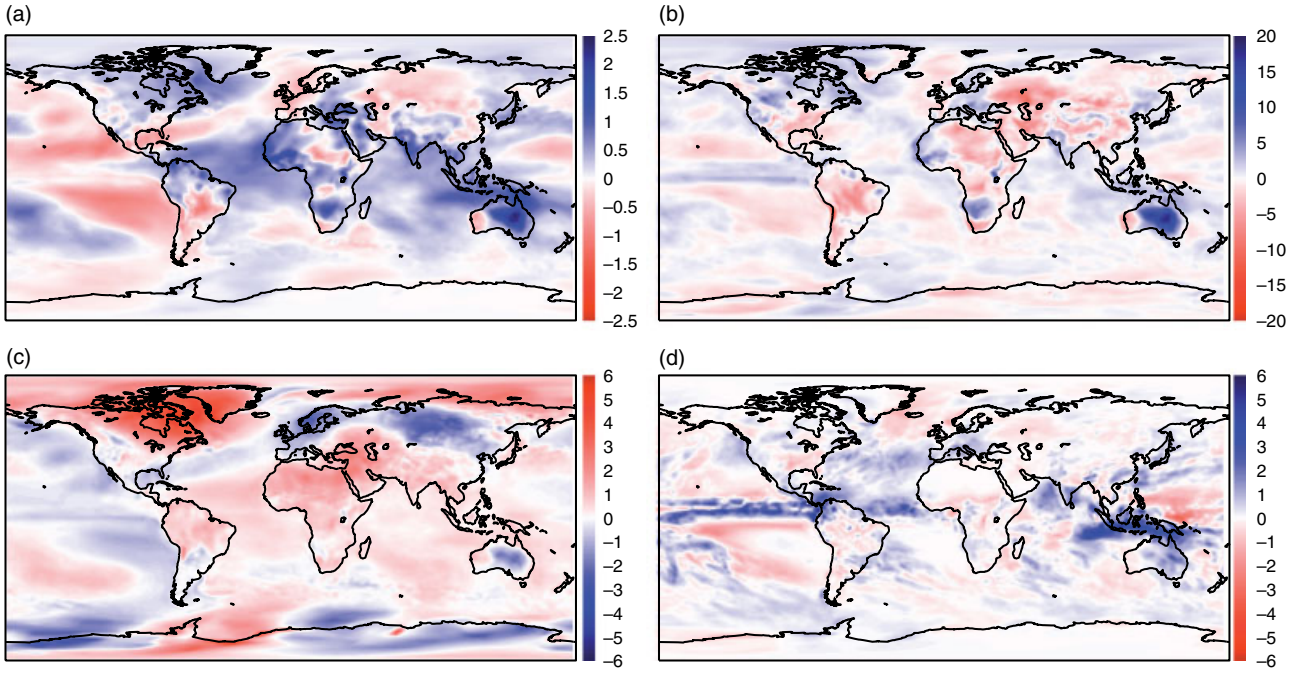

Figure 36. Anomalies for the year 2010 relative to 1989–2009 means from ERA-Interim. (a) Specific humidity analysis ($g\,kg^{-1}$) at a height of 2 m. (b) Relative humidity analysis (%) at a height of 2 m. (c) Temperature analysis (K) at a height of 2 m. (d) Precipitation ($mm\,day^{-1}$) accumulated in the 12–24 h range from twice-daily forecasts initiated at 0000 and 1200 UTC.

reprocess and recalibrate their data holdings, offering further potential for improving and refining estimates of essential climate variables, either by reanalysis or otherwise. Nevertheless, limitations remain, e.g. in observational data on humidity, wind, and the distribution of aerosols and greenhouse gases in large parts of the atmosphere. Data assimilation can fill the gaps by adding physically meaningful information from forecast models, but not without uncertainty.

After completion of the extension of ERA-Interim back to 1979, reanalysis activities at ECMWF will shift focus to the collection and preparation of input observations, boundary conditions, and atmospheric forcing data for a comprehensive atmospheric reanalysis spanning the entire twentieth century. Data assimilation for the earlier (pre-satellite) segment of the instrumental record presents many new challenges. The NOAA-CIRES Twentieth Century Reanalysis from 1871 to 2008 (Compo *et al.*, 2011) has demonstrated that it is possible to extract meaningful information about the global atmospheric circulation, in this case from surface observations only. This pioneering project has been closely linked with data recovery and digitisation efforts carried out by numerous groups around the world, under the umbrella of the ACRE initiative (http://www.met-acre.org). It illustrates well the truly global nature of the reanalysis effort, bringing together data providers, reanalysis producers, and the user community.

Acknowledgements

ECMWF reanalysis is the product of many years of outstanding scientific research and development by a large number of individuals. Human as well as technical resources for computing, data services, and user support at ECMWF are excellent and greatly appreciated. Interest and feedback from users of reanalysis data are instrumental to improve the quality of our products. International collaborations with data providers and reanalysis producers have been invaluable to our work. Preparation and production of ERA-Interim has been supported by staff secondments from the NCAS, the JMA, and the KMA. Work at the University of Leeds on the evaluation of the stratospheric circulation produced by ERA-Interim was partly funded by the NERC and by the European Commission via the GEOMON project.

Appendix. Acronyms

ACRE	Atmospheric Circulation Reconstructions over the Earth
AIREP	Aircraft meteorological report
AIRS	Advanced Infrared Sounder
AMSR-E	Advanced Microwave Scanning Radiometer for EOS
AMSU-A	Advanced Microwave Sounding Unit-A
AMSU-B	Advanced Microwave Sounding Unit-B
AMV	Atmospheric Motion Vector
ASCAT	Advanced Scatterometer
ATOVS	Advanced TIROS Operational Vertical Sounder
AVHRR	Advanced Very High Resolution Radiometer
CFC	Chlorofluorocarbon
CHAMP	Challenging Mini-Satellite Payload
CIRES	Cooperative Institute for Research in Environmental Science (USA)
COARE	Coupled Ocean–Atmosphere Response Experiment
COSMIC	Constellation Observing System for Meteorology, Ionosphere and Climate
CRUTEM3	Climatic Research Unit and Hadley Centre (UK) land surface temperature dataset

CSR	Clear-sky radiance
DMSP	Defense Meteorological Satellite Program
DRIBU	Report from drifting buoy
ECMWF	European Centre for Medium-Range Weather Forecasts
ENSO	El Niño/La Niña–Southern Oscillation
ENVISAT	Environmental Satellite
EOF	Empirical Orthogonal Function
EOS	Earth Observing System
EPIC	East Pacific Investigation of Climate
ERA	ECMWF Reanalysis
ERA-15	A 15-year ERA starting from 1979
ERA-40	A 45-year ERA from September 1957 to August 2002
ERA-Interim	An ERA from January 1989 onward (to be extended back to January 1979)
ERS	European Remote Sensing Satellite
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FGAT	First Guess at the Appropriate Time
FGGE	First GARP Global Experiment
GARP	Global Atmospheric Research Programme
GEOMON	Global Earth Observation and Monitoring of the Atmosphere
GLCC	Global Land Cover Characteristics
GMF	Geophysical Model Function
GMS	Geostationary Meteorological Satellite
GOES	Geostationary Operational Environmental Satellite
GOME	Global Ozone Monitoring Experiment
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GPS	Global Positioning System
GPSRO	GPS Radio Occultation
GRAS	Global Navigation Satellite Systems Receiver for Atmospheric Sounding
GTOPO30	Global 30 Arc-Second Elevation Data Set
GTS	Global Telecommunication System
HadCRUH	Hadley Centre and Climate Research Unit (UK) global surface humidity dataset
HALO	Harmonized coordination of Atmosphere, Land and Ocean projects
HIRS	High-Resolution Infrared Sounder
IASI	Infrared Atmospheric Sounding Interferometer
IFS	Integrated Forecast System
IMS	Interactive Multisensor Snow and Ice Mapping System
IOP	Intensive Observing Period
IPCC	Intergovernmental Panel on Climate Change
ISCCP	International Satellite Cloud Climatology Project
ITCZ	Intertropical Convergence Zone
JMA	Japan Meteorological Agency
JRA-25	Japanese 25-year Reanalysis Project
KMA	Korea Meteorological Administration
METAR	Meteorological Aviation Report
MHS	Microwave Humidity Sounder
MJO	Madden–Julian Oscillation
MLS	Microwave Limb Sounder
MODIS	Moderate Resolution Imaging Spectroradiometer
MSU	Microwave Sounding Unit
MTSAT	Multifunctional Transport Satellite
NCAS	National Centre for Atmospheric Science (USA)
NCEP	National Centers for Environmental Prediction (USA)
NERC	Natural Environment Research Council (UK)
NESDIS	National Environmental Satellite, Data, and Information Service (USA)
NOAA	National Oceanic and Atmospheric Administration (USA)
OI	Optimal interpolation
OLR	Outgoing Long-wave Radiation
OMI	Ozone Monitoring Instrument
OSTIA	Operational Sea Surface Temperature and Sea Ice Analysis
PILOT	Wind report from pilot balloon
PRISM	Parameter-elevation Regressions on Independent Slopes Model
PROFILER	Report from wind profiler
QBO	Quasi-Biennial Oscillation
QuikSCAT	Quick Scatterometer
RAOBCORE	Radiosonde Observation Correction using Reanalyses
RMS	Root mean square
RSS	Remote Sensing Systems
RTG	Real-time gridded
RTTOV	Radiative Transfer for TOVS
SBUV	Solar Backscattered UltraViolet
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Cartography
SHIP	Report from ship
SIC	Sea-ice concentration
SSM/I	Special Sensor Microwave/Imager
SSMI/S	Special Sensor Microwave Imager/Sounder
SST	Sea-surface temperature
SSU	Stratospheric Sounding Unit
SYNOP	Land surface synoptic report
TCWV	Total column water vapour
TEMP	Report from radiosounding
TESSEL	Tiled ECMWF Scheme for Surface Exchanges over Land
TIROS	Television Infrared Observation Satellite
TMI	TRMM Microwave Imager
TOGA	Tropical Ocean–Global Atmosphere
TOFD	Turbulent Orographic Form Drag
TOMCAT	Toulouse Off-line Model of Chemistry And Transport
TOMS	Total Ozone Mapping Spectrometer
TOVS	TIROS Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission
UCAR	University Corporation for Atmospheric Research (USA)
VarBC	Variational bias correction
WAM	Wave modelling
WOUDC	World Ozone and Ultraviolet Radiation Data Centre
*x*D-Var	*x*-dimensional variational analysis

References

ECMWF publications are available on the internet at http://www.ecmwf.int/publications

Abdalla S, Hersbach H. 2004. 'The technical support for global validation of ERS wind and wave products at ECMWF'. *Final Report for ESA contract 15988/02/I-LG*, ECMWF: Reading, UK.

Abdalla S, Hersbach H. 2007. 'The technical support for global validation of ERS wind and wave products at ECMWF (April 2004 – June 2007)'. *Final report for ESA contract 18212/04/I-OL*, ECMWF: Reading, UK.

Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P. 2003. The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979 – present). *J. Hydrometeorol.* **4**: 1147–1167.

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

Andersson E, Bauer P, Beljaars A, Chevallier F, Hólm EV, Janisková M, Kållberg P, Kelly G, Lopez P, McNally AP, Moreau E, Simmons AJ, Thépaut J-N, Tompkins AM. 2005. Assimilation and modeling of the atmospheric hydrological cycle in the ECMWF forecasting system. *Bull. Amer. Meteorol. Soc.* **86**: 387–402.

Andrae U, Sokka N, Onogi K. 2004. *'The radiosonde temperature bias correction in ERA-40'*. *ERA-40 Project Report Series*, No. 15. ECMWF: Reading, UK.

Andrews AE, Boering KA, Daube BC, Wofsy SC, Loewenstein M, Jost H. 2001. Mean ages of stratospheric air derived from *in situ* observations of CO_2, CH_4 and N_2O. *J. Geophys. Res.* **106**: 32295–32314.

Balmaseda MA, Mogensen K. 2010. *'Evaluation of ERA-Interim forcing fluxes from an ocean perspective'*. *ERA Report Series*, No. 6. ECMWF: Reading, UK.

Balmaseda MA, Vidard A, Anderson D. 2008. The ECMWF ORA-S3 ocean analysis system. *Mon. Weather Rev.* **136**: 3018–3034.

Balsamo G, Boussetta S, Lopez P, Ferranti L. 2010. *'Evaluation of ERA-Interim and ERA-Interim-GPCP-rescaled precipitation over the USA'*. *ERA Report Series*, No. 5. ECMWF: Reading, UK.

Bauer P, Kelly G, Andersson E. 2002. 'SSM/I radiance assimilation at ECMWF'. In *Proceedings of the ECMWF/GEWEX workshop on Humidity Analysis*, Reading, 8–11 July 2002, 167–175.

Bauer P, Lopez P, Benedetti A, Salmond D, Moreau E. 2006a. Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. I: 1D-Var. *Q. J. R. Meteorol. Soc.* **132**: 2277–2306.

Bauer P, Lopez P, Salmond D, Benedetti A, Saarinen S, Moreau E. 2006b. Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var. *Q. J. R. Meteorol. Soc.* **132**: 2307–2332.

Bechtold P, Chaboureau JP, Beljaars ACM, Betts AK, Köhler M, Miller M, Redelsperger J-L. 2004. The simulation of the diurnal cycle of convective precipitation over land in a global model. *Q. J. R. Meteorol. Soc.* **130**: 3119–3137.

Beljaars ACM, Brown AR, Wood N. 2004. A new parametrization of turbulent orographic form drag. *Q. J. R. Meteorol. Soc.* **130**: 1327–1347.

Beljaars ACM, Bechtold P, Köhler M, Orr A, Tompkins AM. 2006. 'Developments in model physics after ERA-40'. In *Proceedings of the ECMWF/GEO Workshop on atmospheric reanalysis,* Reading, UK, 19–22 June, 81–90.

Beljaars ACM, Viterbo P. 1994. The sensitivity of winter evaporation to the formulation of aerodynamic resistance in the ECMWF model. *Boundary-Layer Meteorol.* **71**: 135–149.

Berrisford P, Dee DP, Fielding K, Fuentes M, Kållberg P, Kobayashi S, Uppala SM. 2009. *'The ERA-Interim Archive'*. *ERA Report Series*, No. 1. ECMWF: Reading, UK.

Berrisford P, Kållberg P, Kobayashi S, Dee DP, Uppala SM, Simmons AJ, Poli P. 2011. *'Atmospheric conservation properties in ERA-Interim'*. *ERA Report Series*, ECMWF: Reading, UK. In press.

Bidlot J-R, Holmes DJ, Wittmann PA, Lalbeharry R, Chen HS. 2002. Intercomparison of the performance of operational ocean wave forecasting systems with buoy data. *Weather Forecasting* **17**: 287–310.

Bidlot J-R, Janssen P, Abdalla S. 2007. *'Impact of the revised formulation for ocean wave dissipation on the ECMWF operational wave model'*. Tech. Memo. 509, ECMWF: Reading, UK.

Bormann N, Thépaut J-N. 2004. Impact of MODIS polar winds in ECMWF's 4D-VAR data assimilation system. *Mon. Weather Rev.* **132**: 929–940.

Bromwich DH, Nicolas JP, Monaghan AJ. 2011. An assessment of precipitation changes over Antarctica and the Southern Ocean since 1989 in contemporary global reanalyses. *J. Climate*, DOI: 10.1175/2011JCLI4074.1.

Cariolle D, Déqué M. 1986. Southern hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model. *J. Geophys. Res.* **91**: 10825–10846.

Cariolle D, Teyssèdre H. 2007. A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations. *Atmos. Chem. Phys.* **7**: 2183–2196.

Cassou C. 2008. Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature* **455**: 523–527.

Chipperfield MP. 2006. New version of the TOMCAT/SLIMCAT off-line chemical transport model. *Q. J. R. Meteorol. Soc.* **132**: 1179–1203.

Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason Jr BE, Vose RS, Rutledge G, Bessemoulin P, Brönnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY, Jones PD, Kruk MC, Kruger AC, Marshall GJ, Maugeri M, Mok HY, Nordli Ø, Ross TF, Trigo RM, Wang XL, Woodruff SD, Worley SJ. 2011. The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol. Soc.* **137**: 1–28.

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

Courtier P, Andersson E, Heckley W, Pailleux J, Vasiljevic D, Hamrud M, Hollingsworth A, Rabier F, Fisher M. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Q. J. R. Meteorol. Soc.* **124**: 1783–1807.

Dee DP. 2005. Bias and data assimilation. *Q. J. R. Meteorol. Soc.* **131**: 3323–3343.

Dee DP. 2008. 'Importance of satellites for stratospheric data assimilation'. In *Proceedings of the ECMWF Seminar on Recent developments in the use of satellite observations in numerical weather prediction*, 3–7 September 2007. ECMWF: Reading, UK.

Dee DP, Uppala SM. 2008. *'Variational bias correction in ERA-Interim'*. *Tech. Memo. No.* 575. ECMWF: Reading, UK.

Dee DP, Uppala SM. 2009. Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Q. J. R. Meteorol. Soc.* **135**: 1830–1841.

Dee DP, Källén E, Simmons AJ, Haimberger L. 2011. Comments on 'Reanalyses suitable for characterizing long-term trends'. *Bull. Amer. Meteorol. Soc.* **92**: 65–70.

Delsol C, Dee DP, Uppala SM, Bormann N, Thépaut J-N, Bauer P. 2008. 'Use of reprocessed AMVs in the ECMWF Interim reanalysis'. In *Proceedings of the 9th International Winds Workshop, Annapolis, USA*, 14–18 April 2008.

Derber J, Bouttier F. 1999. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus* **51A**: 195–221.

Dethof A, Hólm EV. 2004. Ozone assimilation in the ERA-40 reanalysis project. *Q. J. R. Meteorol. Soc.* **130**: 2851–2872.

Douville H, Royer J-F, Mahfouf J-F. 1995. A new snow parameterization for the Météo-France climate model. Part I: Validation in stand-alone experiments. *Climate Dyn.* **12**: 21–35.

Douville H, Mahfouf J-F, Saarinen S, Viterbo P. 1998. *'The ECMWF surface analysis: Diagnostics and prospects'*. *Tech. Memo.* No. 258. ECMWF: Reading, UK.

Douville H, Viterbo P, Mahfouf J-F, Beljaars ACM. 2000. Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data. *Mon. Weather Rev.* **128**: 1733–1756.

Dragani R. 2010a. *'On the quality of the ERA-Interim ozone reanalyses. Part I: Comparisons with in situ measurements'*. *ERA Report Series*, No. 2, ECMWF: Reading, UK.

Dragani R. 2010b. *'On the quality of the ERA-Interim ozone reanalyses. Part II: Comparisons with satellite data'*. *ERA Report Series*, No. 3, ECMWF: Reading, UK.

Dragani R, Dee DP. 2008. 'Progress in ozone monitoring and assimilation'. *ECMWF Newsletter* **116**: 35–42.

Drusch M, Vasiljevic D, Viterbo P. 2004. ECMWF's global snow analysis: Assessment and revision based on satellite observations. *J. Appl. Meteorol.* **43**: 1282–1294.

Ferranti L, Palmer TN, Molteni F, Klinker E. 1990. Tropical–extratropical interaction associated with the 30–60 day oscillation

and its impact on medium and extended range prediction. *J. Atmos. Sci.* **47**: 2177–2199.

Fiorino M. 2004. '*A multi-decadal daily sea surface temperature and sea ice concentration data set for the ERA-40 reanalysis*'. *ERA-40 Project Report Series* No. 12, ECMWF: Reading, UK.

Fisher M. 2003. 'Background-error covariance modelling'. In *Proceedings of seminar on recent developments in data assimilation for atmosphere and ocean*, 8–12 September 2003, ECMWF: Reading, UK.

Fisher M. 2004. 'Generalized frames on the sphere, with application to the background-error covariance modelling'. In *Proceedings of seminar on recent developments in numerical methods for atmospheric and ocean modelling*, 6–10 September 2004, ECMWF: Reading, UK.

Fisher M. 2005. 'Wavelet J_b – A new way to model the statistics of background errors'. *ECMWF Newsletter* **106**: 23–28.

Fisher M, Courtier P. 2005. '*Estimating the covariance matrices of analysis and forecast error in variational data assimilation*'. *Tech. Memo.* No. 220, ECMWF: Reading, UK.

Fortuin JPF, Langematz U. 1994. An update on the global ozone climatology and on concurrent ozone and temperature trends. *Atmos. Sensing Modeling* **2311**: 207–216.

Gérard E, Saunders R. 1999. Four-dimensional variational assimilation of Special Sensor Microwave/Imager total column water vapour in the ECMWF model. *Q. J. R. Meteorol. Soc.* **125**: 3077–3102.

Geer AJ, Bauer P, Lopez P. 2008. Lessons learnt from the operational 1D+4D-Var assimilation of rain- and cloud-affected SSM/I observations at ECMWF. *Q. J. R. Meteorol. Soc.* **134**: 1513–1525.

Gibson JK, Kållberg P, Uppala SM, Nomura A, Hernandez A, Serrano E. 1997. '*ERA description*'. *ERA-15 Report Series*, No. 1, ECMWF: Reading, UK.

Haimberger L, Tavolato C, Sperka S. 2008. Toward elimination of the warm bias in historic radiosonde temperature records: Some new results from a comprehensive intercomparison of upper-air data. *J. Climate* **21**: 4587–4606.

Healy SB, Thépaut J-N. 2006. Assimilation experiments with CHAMP GPS radio occultation measurements. *Q. J. R. Meteorol. Soc.* **132**: 605–623.

Hendon HH, Liebmann B. 1990. A composite study of onset of the Australian summer monsoon. *J. Atmos. Sci.* **47**: 2227–2240.

Hersbach H, Drusch M. 2009. 'Usage of sea-surface temperature and sea-ice cover at ECMWF'. In *Proceedings from Group for High-Resolution Sea Surface Temperature (GHRSST) 2009 International Users Symposium*, 28–29 May 2009, Santa Rosa, CA, USA.

Hess M, Koepke P, Schult I. 1998. Optical properties of aerosols and clouds: The software package OPAC. *Bull. Amer. Meteorol. Soc.* **79**: 831–844.

Hólm EV. 2003. 'Revision of the ECMWF humidity analysis: Construction of a Gaussian control variable'. In *Proceedings of the ECMWF/GEWEX Workshop on Humidity Analysis*, 8–11 July 2002, ECMWF: Reading, UK.

Holtslag AAM. 1998. Modelling of atmospheric boundary layers. Pp. 85–110 in: *Clear and cloudy boundary layers,* Holtslag AAM, Duynkerke P. (eds.) Royal Netherlands Academy of Arts and Sciences: Amsterdam.

IPCC. 1996. *Climate Change 1995. Second assessment report of the Intergovernmental Panel on Climate Change.* Houghton JT, Meira Filho LG, Callander BA, Harris N, Kattenberg A, Maskell K. (eds.) Cambridge University Press: Cambridge, UK.

Isaksen L, Janssen PAEM. 2004. The benefit of ERS scatterometer winds in ECMWF's variational assimilation system. *Q. J. R. Meteorol. Soc.* **130**: 1793–1814.

Janssen PAEM. 2004. *The Interaction of Ocean Waves and Wind.* Cambridge University Press: Cambridge, UK.

Janssen PAEM. 2008. Progress in ocean wave forecasting. *J. Comput. Phys.* **227**: 572–3594.

Janssen PAEM, Bidlot J-R, Abdalla S, Hersbach H. 2005. '*Progress in ocean wave forecasting at ECMWF*'. *Tech. Memo.* 478, ECMWF: Reading, UK.

Jarraud M, Goas J, Deyts C. 1989. Prediction of an exceptional storm over France and Southern England (15–16 October 1987). *Weather Forecasting* **4**: 517–536.

Jung T, Klinker K, Uppala SM. 2004. Reanalysis and reforecast of three major European storms of the twentieth century using the ECMWF forecasting system. Part I: Analyses and deterministic forecasts. *Meteorol. Appl.* **11**: 343–361.

Kållberg P. 2011. '*Forecast drift in ERA-Interim*'. *ERA Report Series*, No. 10, ECMWF: Reading, UK.

Kalnay E, Kanamitsu M, Kirtler 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 KC, Ropelewski C, Wang J, Leetma A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteorol. Soc.* **77**: 437–471.

Kessler KS, McPhaden M. 1995. Oceanic equatorial waves and the 1991–93 El Niño. *J. Climate* **8**: 1757–1774.

Klein SA, Hartmann DL. 1993. The seasonal cycle of low stratiform cloud. *J. Climate* **6**: 1587–1606.

Kobayashi S, Matricardi M, Dee DP, Uppala SM. 2009. Toward a consistent reanalysis of the upper stratosphere based on radiance measurements from SSU and AMSU-A. *Q. J. R. Meteorol. Soc.* **135**: 2086–2099.

Köhler M. 2005. 'Improved prediction of boundary layer clouds'. *ECMWF Newsletter* **104**:.

Köhler M, Ahlgrimm M, Beljaars ACM. 2011. Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model. *Q. J. R. Meteorol. Soc.* **137**: 43–57.

Komen GJ, Cavaleri L, Donelan M, Hasselmann K, Hasselmann S, Janssen PAEM. 1994. *Dynamics and Modelling of Ocean Waves.* Cambridge University Press: Cambridge, UK.

Lionello P, Günther H, Janssen PAEM. 1992. Assimilation of altimeter data in a global third-generation wave model. *J. Geophys. Res.* **97**: 14453–14474.

Lock AP. 1998. The parameterization of entrainment in cloudy boundary layers. *Q. J. R. Meteorol. Soc.* **124**: 2729–2753.

Loveland TR, Reed BC, Brown JF, Ohlen DO, Zhu Z, Youing L, Merchant JW. 2000. Development of a global land cover characteristics database and IGB6 DISCover from the 1 km AVHRR data. *Int. J. Remote Sensing* **21**: 1303–1330.

Mahfouf J-F, Viterbo P, Douville H, Beljaars ACM, Saarinen S. 2000. 'A revised land-surface analysis scheme in the integrated forecasting system'. *ECMWF Newsletter* **88**: 8–13.

Maloney ED, Hartmann DL. 2000. Modulation of eastern North Pacific hurricanes by the Madden–Julian oscillation. *J. Climate* **13**: 1451–1460.

Marécal V, Mahfouf J-F. 2000. Variational retrieval of temperature and humidity profiles from TRMM precipitation data. *Mon. Weather Rev.* **128**: 3853–3866.

Marécal V, Mahfouf J-F. 2002. Four-dimensional variational assimilation of total column water vapour in rainy areas. *Mon. Weather Rev.* **130**: 43–58.

Marécal V, Mahfouf J-F. 2004. Experiments on 4D-Var assimilation of rainfall data using an incremental formulation. *Q. J. R. Meteorol. Soc.* **129**: 3137–3160.

Matricardi M, Saunders R. 1999. Fast radiative transfer model for simulations of infrared atmospheric sounding interferometer radiances. *Appl. Opt.* **38**: 5679–5691.

Matricardi M, Chevallier F, Kelly G, Thépaut J-N. 2004. An improved general fast radiative transfer model for the assimilation of radiance observations. *Q. J. R. Meteorol. Soc.* **130**: 153–173.

McNally AP, Watts PD. 2003. A cloud detection algorithm for high-spectral-resolution infrared sounders. *Q. J. R. Meteorol. Soc.* **129**: 2411–2323.

McNally AP, Watts PD, Smith JA, Engelen R, Kelly GA, Thépaut J-N, Matricardi M. 2006. The assimilation of AIRS radiance data at ECMWF. *Q. J. R. Meteorol. Soc.* **132**: 935–957.

Monge-Sanz BM, Chipperfield MP, Simmons AJ, Uppala SM. 2007. Mean age of air and transport in a CTM: Comparison of different ECMWF analyses. *Geophys. Res. Lett.* **34**: L04801.

Morris RM, Gadd AJ. 1988. Forecasting the Great Storm. *Weather* **43**: 70–89.

Murakami T. 1976. Cloudiness fluctuations during the summer monsoon. *J. Meteorol. Soc. Japan* **54**: 802–817.

Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hatsushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K, Kadokura S,

Wada K, Kato K, Oyama R, Ose T, Mannoji N, Taira R. 2007. The JRA-25 Reanalysis. *J. Meteor. Soc. Japan* **85**: 369–432.

Paltridge GW, Platt CMR. 1976. *Radiative Processes in Meteorology and Climatology*. Elsevier: New York.

Peuch A, Thépaut J-N, Pailleux J. 2000. Dynamical impact of total ozone observations in a four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.* **126**: 1641–1659.

Phalippou L. 1996. Variational retrieval of humidity prole, wind speed and cloud liquid-water path with the SSM/I: Potential for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **122**: 327–355.

Poli P, Healy SB, Dee DP. 2010. Assimilation of Global Positioning System radio occultation data in the ECMWF ERA-Interim reanalysis. *Q. J. R. Meteorol. Soc.* **136**: 1972–1990.

Rabier F, Thépaut J-N, Courtier P. 1998. Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Q. J. R. Meteorol. Soc.* **124**: 1861–1887.

Rabier F, Järvinen H, Klinker E, Mahfouf J-F, Simmons AJ. 2000. The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **126**: 1143–1170.

Randel W, Udelhofen P, Fleming E, Geller M, Gelman M, Hamilton K, Karoly D, Ortland D, Pawson S, Swinbank R, Wu F, Baldwin M, Chanin M-L, Keckhut P, Labitzke K, Remsberg E, Simmons AJ, Wu D. 2004. The SPARC intercomparison of middle atmosphere climatologies. *J. Climate* **17**: 986–1003.

Ray EA, Moore FL, Elkins JW, Dutton GS, Fahey DW, Voemel H, Oltmans SJ, Rosenlof KH. 1999. Transport into the Northern Hemisphere lowermost stratosphere revealed by *in situ* tracer measurements. *J. Geophys. Res.* **104**: 26565–26580.

Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim G-K, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W, Joiner J, Koster RD, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redder CR, Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woollen J. 2011. MERRA – NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, DOI: 10.1175/JCLI-D-11-00015.1.

Rodwell MJ. 2005. 'The local and global impact of the recent change in aerosol climatology'. *ECMWF Newsletter* **105**: 17–23.

Rodwell MJ, Jung T. 2008. Understanding the local and global impacts of model physics changes: An aerosol example. *Q. J. R. Meteorol. Soc.* **134**: 1479–1497.

Saha S, Moorthi S, Pan H-L, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu H, Stokes D, Grumbine R, Gayno G, Hou Y-T, Chuang H-Y, Juang H-MH, Sela J, Iredell M, Treadon R, Kleist D, van Delst P, Keyser D, Derber J, Ek M, Meng J, Wei H, Yang R, Lord S, van den Dool H, Kumar A, Wang W, Long C, Chelliah M, Xue Y, Huang B, Schemm J-K, Ebisuzaki W, Lin R, Xie P, Chen M, Zhou S, Higgins W, Zou C-Z, Liu Q, Chen Y, Han Y, Cucurull L, Reynolds RW, Rutledge G, Goldberg M. 2010. The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteorol. Soc.* **91**: 1015–1057.

Saunders R, Matricardi M, Brunel P. 1999. '*A fast radiative transfer model for assimilation of satellite radiance observations: RTTOV-5*'. *Research Department Tech. Memo.* 282, ECMWF: Reading, UK.

Schneider U, Fuchs T, Meyer-Christoffer A, Rudolf B. 2008. 'Global precipitation analysis products of the GPCC'. Deutscher Wetterdienst: Offenbach, Germany. Available at http://gpcc.dwd.de.

Schoeberl MR, Douglass AR, Stolarski RS, Pawson S, Strahan SE, Read W. 2008. Comparison of lower stratospheric tropical mean vertical velocities. *J. Geophys. Res.* **113**: D24109.

Schubert SD, Rood R, Pfaendtner J. 1993. An assimilated dataset for earth science applications. *Bull. Am. Meteorol. Soc.* **74**: 2331–2342.

Shutts GJ. 1990. Dynamical aspects of the October storm, 1987: A study of a successful fine-mesh simulation. *Q. J. R. Meteorol. Soc.* **116**: 1315–1347.

Siddans R, Reburn WJ, Kerridge BJ, Munro R. 2002. 'Height-resolved ozone information in the troposphere and lower stratosphere from GOME'. Rutherford Appleton Laboratory: Didcot, UK. Available at http://badc.nerc.ac.uk/data/gome/ers_html.

Siebesma P, Teixeira J. 2000. 'An advection-diffusion scheme for the convective boundary layer: Description and 1D-results'. In *Proceedings of the 14th symposium on boundary layers and turbulence*, Amer. Meteorol. Soc: Boston.

Simmons AJ, Willett KM, Jones PD, Thorne PW, Dee DP. 2010. Low-frequency variations in surface atmospheric humidity, temperature and precipitation: Inferences from reanalyses and monthly gridded observational datasets. *J. Geophys. Res.* **115**: D01110.

Stark JD, Donlon CJ, Martin MJ, McCulloch ME. 2007. 'OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system'. *Proceedings of Oceans '07 IEEE Conference, 'Marine Challenges: Coastline to Deep Sea'*, 18–21 June 2007, Aberdeen, UK.

Sterl A, Caires S. 2005. Climatology, variability and extrema of ocean waves – The web-based KNMI/ERA-40 wave atlas. *Int. J. Climatol.* **25**: 963–977.

Stoffelen ACM, Anderson DLT. 1997. Scatterometer data interpretation: Derivation of the transfer function CMOD4. *J. Geophys. Res.* **102**(C3): 5767–5780.

Tanré D, Geleyn J-F, Slingo JM. 1984. First results of the introduction of an advanced aerosol-radiation interaction in the ECMWF low resolution global model. Pp. 133–177 in *Aerosols and their Climatic Effects*, Gerber HE, Deepak A. (eds.) A. Deepak Publishing: Hampton, Va, USA.

Tavolato C, Isaksen L. 2010. '*Data usage and quality control for ERA-40, ERA-Interim, and the operational ECMWF data assimilation system*'. *ERA Report Series*, No. 7, ECMWF: Reading, UK.

Tegen I, Hoorig P, Chin M, Fung I, Jacob D, Penner J. 1997. Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. *J. Geophys. Res.* **102**: 23895–23915.

Thépaut J-N. 2006. 'Assimilating only surface pressure observations in 3D- and 4D-Var'. Pp. 107–112 in *Proceedings of the ECMWF workshop on atmospheric reanalysis*, 19–22 June 2006, ECMWF: Reading, UK.

Thépaut J-N, Courtier P, Belaud G, Lemaître G. 1996. Dynamical structure functions in a four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.* **122**: 535–561.

Thorne PW, Vose RS. 2010. Reanalyses suitable for characterizing long-term trends. *Bull. Amer. Meteorol. Soc.* **91**: 353–361.

Tompkins AM, Cardinali C, Morcrette J-J, Rodwell M. 2005. Influence of aerosol climatology on forecasts of the African Easterly Jet. *Geophys. Res. Lett.* **32**: L10801.

Tompkins AM, Gierens K, Rädel G. 2007. Ice supersaturation in the ECMWF integrated forecast system. *Q. J. R. Meteorol. Soc.* **133**: 622.

Trémolet Y. 2004. Diagnostics of linear and incremental approximations in 4D-Var. *Q. J. R. Meteorol. Soc.* **130**: 2233–2251.

Trenberth KE, Fasullo JT, Kiehl J. 2009. Earth's global energy budget. *Bull. Amer. Meteorol. Soc.* **90**: 311–324.

Troen I, Mahrt L. 1986. A simple model of the atmospheric boundary layer; Sensitivity to surface evaporation. *Boundary-Layer Meteorol.* **37**: 129–148.

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Van De Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette J-J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* **131**: 2961–3012.

Uppala SM, Dee DP, Kobayashi S, Simmons AJ. 2008. 'Evolution of reanalysis at ECMWF'. In *Proceedings of Third WCRP International Conference on Reanalysis*, 28 January–1 February 2008, Tokyo, Japan.

Vasiljevic D. 2006. 'Handling biases in surface pressure (Ps) observations in data assimilation'. Pp. 107–117 in *Proceedings of ECMWF/EUMETSAT NWP-SAF Workshop on bias estimation and correction in data assimilation*, 8–11 November 2005. ECMWF: Reading, UK.

Vasiljevic D, Andersson E, Isaksen L, Garcia-Mendez A. 2006. 'Surface pressure bias correction in data assimilation'. *ECMWF Newsletter* **108**: 20–27.

Veersé F, Thépaut J-N. 1998. Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.* **124**: 1889–1908.

Vitart F. 2004. Monthly forecasting at ECMWF. *Mon. Weather Rev.* **132**: 2761–2779.

Vitart F, Woolnough S, Balmaseda MA, Tompkins AM. 2007. Monthly forecast of the Madden–Julian Oscillation using a coupled GCM. *Mon. Weather Rev.* **135**: 2700–2715.

Viterbo P, Beljaars ACM. 1995. 'An improved land surface parametrization scheme in the ECMWF model and its validation'. Technical Report 75, ECMWF: Reading, UK.

Viterbo P, Betts AK. 1999. Impact on ECMWF forecasts of changes to the albedo of the boreal forests in the presence of snow. *J. Geophys. Res.* **104**: 27803–27810.

Viterbo P, Beljaars ACM, Mahouf J-F, Teixeira J. 1999. The representation of soil moisture freezing and its impact on the stable boundary layer. *Q. J. R. Meteorol. Soc.* **125**: 2401–2426.

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

Wheeler MC, Hendon HH. 2004. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Weather Rev.* **132**: 1917–1932.

Whitaker JS, Compo GP, Thépaut J-N. 2009. A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Mon. Weather Rev.* **137**: 1991–1999.

Willett KM, Alexander LV, Thorne PW. (eds.) 2010. Global climate (in 'State of the Climate in 2009'). *Bull. Amer. Meteorol. Soc.* **91**: S19–S52.

Wonnacott TH, Wonnacott R. 1997. *Introductory Statistics*. John Wiley & Sons: Chichester, UK.

