TECHNICAL MEMORANDUM

687

Toward a consistent reanalysis of the climate system

Dick Dee, Magdalena Balmaseda,
Gianpaolo Balsamo, Richard Engelen,
and Adrian Simmons

Research Department

November 2012

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

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Abstract

This report reviews past and current reanalysis activities at ECMWF and outlines plans for future development of a coupled climate reanalysis. The Centre's global reanalyses of the atmosphere, ocean, land surface, and atmospheric composition are closely tied to the development of the IFS and its extended capabilities. Reanalysis data also serve numerous external users in a wide range of applications, including climate monitoring and research. New developments in reanalysis at the Centre are focused on extending the length of atmospheric reanalyses, and on developing a coupled data assimilation capability in the IFS.

1. Reanalysis for medium-range forecasting

Reanalysis activities at ECMWF have always been closely connected with the development of the operational forecasting system. A reanalysis of observations collected for the First Global Experiment of the Global Atmospheric Research Programme (FGGE) started only months after the first operational forecast was issued in August 1979 (Bengtsson et al. 1982a; 1982b). Production of this first (“main”) FGGE reanalysis was completed by summer 1981. Based on its results and feedback to the data providers, various corrections and additions were made to the original FGGE input dataset. A second (“final”) reanalysis covering the two FGGE Special Observing Periods (January-February and June-July 1979) was produced in 1986, using the improved FGGE data and an updated version of the forecasting system. These pioneering reanalyses provided the first global atmospheric datasets available for scientific research, and they were widely utilized in predictability studies and for diagnostic purposes (ECMWF 1985; WMO 1985).

The Centre’s role in the FGGE project set in motion a strong feedback loop between improvements in the global observing system, advances in data assimilation methodology, and development of better forecast models through reanalysis. Between 1993 and 1996 an early version of the newly developed Integrated Forecast System (IFS) was used to complete a reanalysis of the period 1979–1993 (ERA-15, Gibson et al. 1997). Starting in 1998, and building on the experiences gained with ERA-15, a reanalysis of the period 1957-2002 was produced over a 4-year period (ERA-40; Uppala et al. 2005). ERA-40 used a version of the IFS that was operational in 2001, but at a lower resolution (T159L60) and with a three-dimensional variational (3D-Var) analysis. The Centre’s most recent atmospheric reanalysis is ERA-Interim (Dee et al. 2011), at spatial resolution (T255L60), covering the modern satellite era from 1979 and continuing forward in time. ERA-Interim uses a 2006 version of the ECMWF system, including four-dimensional variational (4D-Var) analysis.

ERA-Interim is a global atmospheric reanalysis produced with a 2006 version of the IFS (Cy31r2), configured for a spatial resolution of approximately 79 km, on 60 model levels with the top level at 0.1 hPa (T255L60). The reanalysis covers dates from 1 January 1979 until present; the dataset is updated monthly and can be downloaded from www.ecmwf.int/research/era. A detailed description of the ERA-Interim modelling and data assimilation system, the observations used, and various aspects of the quality of the reanalysis, has been published as an open-access journal article (Dee et al. 2011) at http://onlinelibrary.wiley.com/doi/10.1002/qj.828/abstract.

The data assimilation methodology in ERA-Interim is based on ECMWF's 4-dimensional variational analysis (4D-Var), extended with a variational bias correction system for satellite radiances. Separate bias correction schemes are included for surface pressure observations and for temperature observations from radiosondes. Observations assimilated in ERA-Interim include the great majority of in-situ and space-borne data used for operational forecasting. Boundary conditions for the forecast model used in ERA-Interim were taken from ERA-40 prior to 2002, and from ECMWF's operational forecast system for later dates.

Since the IFS contains fully interactive model components for the land surface (since 1991) and the sea state (since 1998), the ERA datasets provide estimates of land-surface parameters (e.g. soil temperature, soil moisture, snow) and, beginning with ERA-40, parameters that describe the sea state (e.g. wave spectra, significant wave height). These estimates are consistent with the meteorological parameters, in the sense that they are constrained by a physically meaningful (coupled) model. However, the use of separate analysis methods for the different model components destroys some of the consistency; see Section 3.

	Model	**Completed**	**Time period**	**Resolution**
FGGE (main)	1979	1981	Dec 1978 – Nov 1979	N48L15
FGGE (final)	1985	1986	Jan/Feb and Jun/Jul 1979	T63L19
ERA-15	1994	1996	1979 – 1993	T106L31
ERA-40	2001	2003	1957 – 2002	T159L60
GEMS	2006	2009	2003 – 2009	T159L60
ERA-Interim	2006	2009 then monthly updates	1979 – present	T255L60
MACC	2010	2011	2003 – 2010	T255L60

Table 1: Global atmospheric reanalyses produced at ECMWF

Each subsequent ECMWF reanalysis has generated a more comprehensive global dataset, with gridded estimates for a growing list of parameters, over longer time periods, and at higher spatial resolution (Table 1). The increasing quality of the reanalyses clearly reflects the modeling and data assimilation improvements implemented in the IFS (e.g. see Figure 1). Over the years, the Centre has also made great strides in providing efficient data services to users in the Member States and elsewhere. Judging by the numerous references in the scientific literature, it is fair to say that ERA data have become an indispensable resource for the atmospheric sciences. Reanalysis provides worldwide visibility to ECMWF that is highly favourable to its reputation as the leading institute in global medium-range weather forecasting.

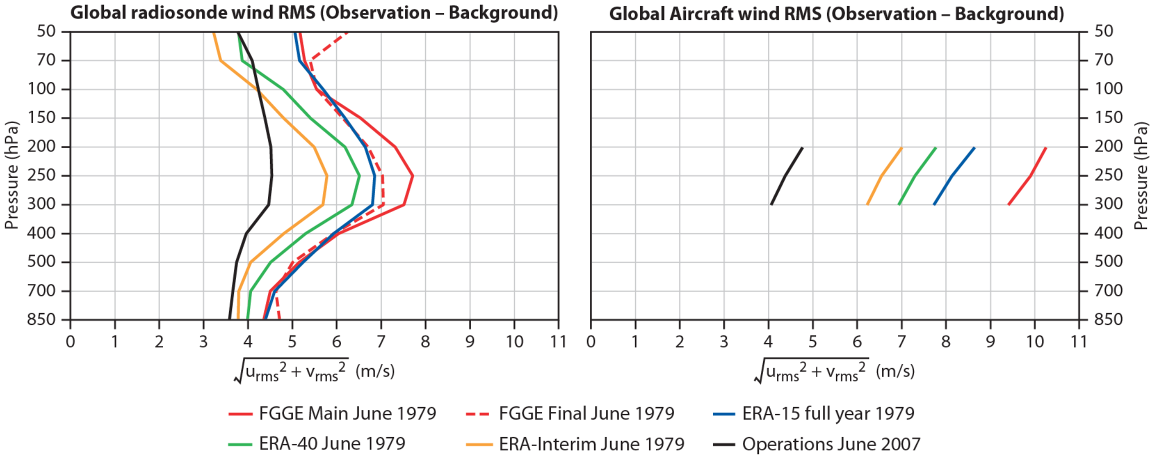


Figure 1: Globally averaged RMS errors in upper-air winds from short-range forecasts produced in successive ECMWF reanalyses, relative to (a) radiosonde observations and (b) aircraft reports. Data are for June 1979. For comparison, background errors in wind estimates from ECMWF operations for June 2007 are also shown. Adapted from Uppala et al. (2008).

2. Reanalysis for monthly and seasonal forecasting

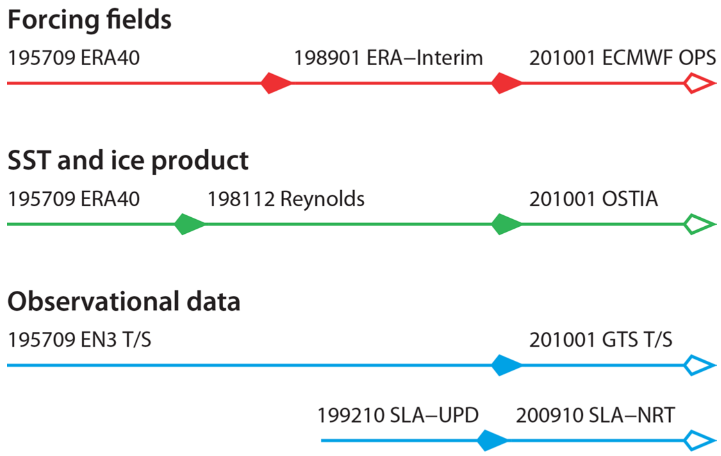
Since 2006, two major reanalyses of the global oceans (ORAS3, Balmaseda et al. 2008 and ORAS4, Balmaseda et al. 2012) have been produced as part of the development of a monthly and seasonal forecasting capability. Prediction of large-scale atmospheric anomalies beyond the medium range depends on the ability to accurately represent slow interactions between the atmosphere and its surface boundaries over land and ocean. The Centre's monthly and seasonal forecast system therefore use a coupled atmosphere-ocean model, based on the IFS extended with the NEMO (Nucleus for European Modelling of the Ocean) ocean model. Developments are currently underway to implement a dynamic sea-ice component in the NEMO model (Tang et al. 2012).

Coupled atmosphere-ocean models tend to develop biases and drifts at the interface (e.g. see Figure 2), which must be corrected a posteriori in order to produce forecasts with any useful skill. In current practice this is accomplished by producing an extensive set of re-forecasts (hindcasts) initialized from a combined ocean-atmosphere reanalysis, then computing the climatology of the model errors relative to the reanalysis, and correcting the model output accordingly. This calibration procedure can be

ORAS4 (Ocean Reanalysis System 4) is the latest ECMWF reanalysis of the global ocean. It consists of an ensemble of 5 members spanning the period 1958 to present. ORAS4 uses the NEMO ocean model, with an approximate resolution of 1° horizontally and 42 vertical levels, and the NEMOVAR data assimilation system in a 3D-Var configuration with a 10-day analysis window. Assimilated observations from the Hadley Centre's EN3 data collection (Ingleby et al. 2007) include temperature and salinity (T/S) profiles, along-track sea level anomalies (SLA) derived from altimeter data, and bathythermograph (XBT) observations with corrections from Wijffels et al. (2009). In addition, gridded maps of observed sea-surface temperatures are used to adjust the heat fluxes from the atmospheric reanalysis via strong relaxation, and global sea-level anomalies are used to constrain fresh-water fluxes.

The ORAS4 data assimilation system includes a model bias correction scheme (Balmaseda et al. 2007) that significantly reduces spurious low-frequency variability associated with changes in the observing system The bias corrections comprise an a-priori component, which is derived from a monthly climatology of model errors estimated during the well-observed Argo period, and an adaptive component, estimated on-line from all available observations. The ORAS4 ensemble is generated by sampling uncertainties arising from errors in wind forcing, limited observation coverage, and the errors in the model representation of the deep ocean.

Quality improvements in ORAS4 relative to the earlier ORAS3 reanalysis are due to the use of atmospheric surface fluxes from the ERA-Interim reanalysis, various improvements in ocean modelling and data assimilation, and the use of improved sub-surface ocean observations. For a detailed description and evaluation of ORAS4 see Balmaseda et al. (2012).



Key input data streams and transition dates for the production of ORAS4. The EN3 database with quality-controlled temperature and salinity (T/S) observations is provided by the Hadley Centre; sea-level anomaly (SLA) data by AVISO; Reynolds SST data by NOAA (OIv2, 1 degree); OSTIA SST by the MetOffice.

expected to work best if the reference dataset is fully consistent with the model to be calibrated, i.e. produced by reanalysis with the same version of the coupled model used for forecasting. Instead, the datasets used to calibrate ECMWF's monthly and seasonal forecast systems have been constructed from separately produced (hence not fully consistent) reanalyses of the atmosphere and ocean.

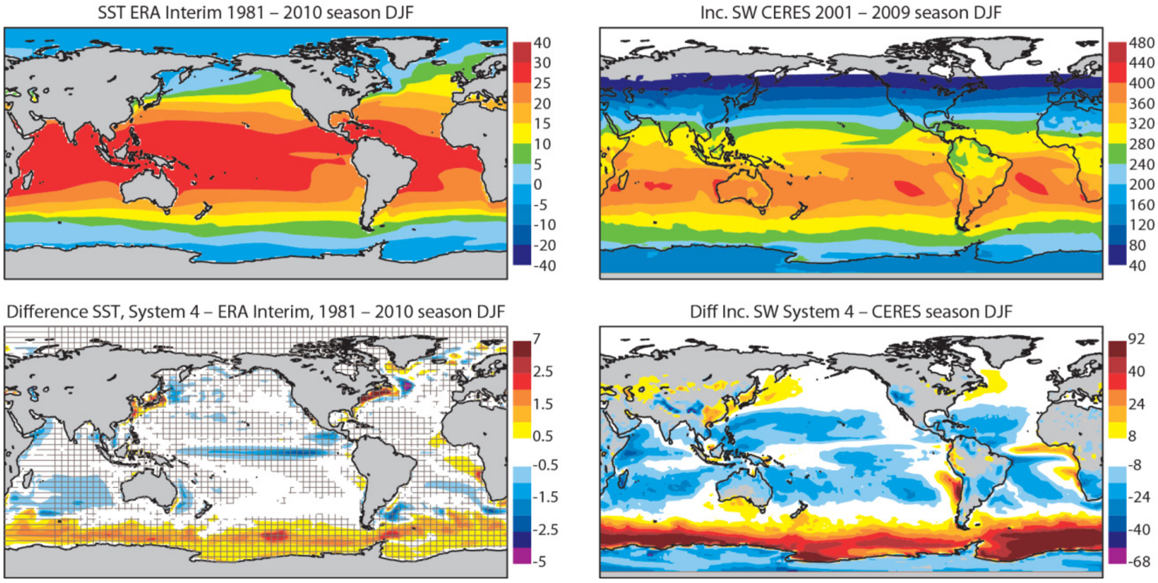


Figure 2: Model drifts in the current ECMWF seasonal forecasting system (System 4). Left panels: Mean forecast errors in SST (bottom) relative to ERA-Interim data (top); averaged over 1981-2010 northern-hemisphere winter seasons. Right: Mean forecast errors in incoming short-wave radiation (bottom) relative to CERES observations (top); averaged over 2001-2009 northern-hemisphere winter seasons.

The ocean reanalyses produced for this purpose make use of the ocean model component of the coupled forecast system, but with prescribed atmospheric forcing (wind, temperature, precipitation). These estimates are provided by the most recent ERA data, possibly adjusted to account for known biases. The lack of feedback between ocean and atmosphere in the reanalyses produces physical inconsistencies in the combined dataset. Nevertheless, the reanalysis of the ocean circulation is quite sensitive to the quality of the atmospheric fluxes, which essentially drive the circulation in the upper layers of the ocean. The impact can propagate to the deeper ocean, which is not well constrained by observations, especially prior to the deployment of the ARGO observing system.

Additional inconsistencies in the combined reference dataset arise as a result of the different release schedules for the various system components. On average, new cycles of the IFS are implemented two or three times yearly. Operational upgrades to the monthly and seasonal forecast systems occur much less frequently; this is mainly due to the complexity of the coupled system and the need for calibration as just described. Production of a new atmospheric reanalysis is even more sporadic, and by the time the dataset is ready for use several new versions of the IFS will have been released.

3. Reanalysis for land-surface monitoring

Discrepancies in model versions also affect the representation of the land surface. The land-surface component of the IFS has undergone significant changes in recent years, e.g. in the area of surface hydrology, the treatment of snow, and the inclusion of carbon. It has been demonstrated (e.g. Koster et al. 2010) that improved initialization of the land surface can enhance the predictability of near-surface temperature by up to 6 weeks. Hence the importance of the representation of the land surface in the monthly and seasonal forecast system, which tends to be more sophisticated than that used for producing a recent atmospheric reanalysis. This creates a further incompatibility in the dataset used to calibrate and initialise the forecasts.

This situation has motivated the development of an off-line land-surface modelling tool at ECMWF (Balsamo et al. 2012) that can be used to upgrade the land-surface state variables associated with an existing atmospheric reanalysis, based on a more recent version of the land-surface model used for reanalysis. The upgraded land-surface is constructed by providing the model with meteorological forcing from the atmospheric reanalysis. Although the off-line land-surface system does not directly assimilate observations, observational constraints are effectively imposed via the meteorological forcing. The system optionally uses independent observational estimates of monthly averaged precipitation from the Global Precipitation Climatology Project (GPCP) to correct biases in reanalysed precipitation. It can produce upgraded land-surface parameters at different spatial resolutions, and is sufficiently efficient to support testing and development of new land-surface model components.

A first upgraded land-surface reanalysis for 1979-2010 (ERA-Interim/Land, see box) has been produced in this fashion and is now available for general use. It is based on a recent version of H-TESSEL (Balsamo et al. 2011) and includes precipitation corrections based on GPCP. Figure 3 illustrates the impact of the new model on the reanalysis of surface hydrology. Balsamo et al. (2012) provide additional information about ERA-Interim/Land with evaluations for various parameters of the upgraded land-surface. A similar off-line reanalysis at higher spatial resolution is in planning.

As currently implemented, the off-line land-surface system may be regarded as a sophisticated tool for downscaling and enhancing the description of the land surface as given by an existing global atmospheric reanalysis. This capability presents interesting opportunities for downstream climate services, as the downscaling to higher resolution offers support for local generation of specialized products tailored to local needs. Additional developments in the off-line system planned at ECMWF will provide the ability to directly assimilate (or re-assimilate) terrestrial observations. Improvements in the land-surface model, especially when combined with enhanced spatial resolution, should allow better use of near-surface observations and can potentially accommodate many observations that are not currently useable in atmospheric reanalyses.

ERA-Interim/Land is a global land-surface reanalysis covering the period 1979-2010, providing 6-hourly estimates of 13 parameters describing soil temperature, soil water content, and snow. These data are available at a spatial resolution of approximately 79 km on a reduced Gaussian grid (N128). ERA-Interim/Land was produced using the Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land (HTESSEL), an extension of the TESSEL scheme (van den Hurk et al. 2000) with improved soil hydrology (Balsamo et al. 2009), a new snow scheme (Dutra et al. 2010), a multi-year satellite-based vegetation climatology (Boussetta et al. 2011), and a revised scheme for bare-soil evaporation (Albergel et al. 2012). Meteorological and radiative forcing needed to drive HTESSEL was derived from the ERA-Interim atmospheric reanalysis. A scale-selective adjustment of the 3-hourly ERA-Interim precipitation estimates was performed in order to improve the match to observed monthly rain accumulations provided by the GPCP v2.1 dataset.

Balsamo et al. (2012) provide a detailed description and evaluation of the ERA-Interim/Land dataset. Based on independent observations, they demonstrate improvements (relative to the land-surface parameters provided by ERA-Interim) in latent and sensible heat fluxes, soil moisture content, and in various aspects of estimated snow cover. A comparison with river discharge measurements suggests a significantly improved consistency of the water cycle in ERA-Interim/Land.

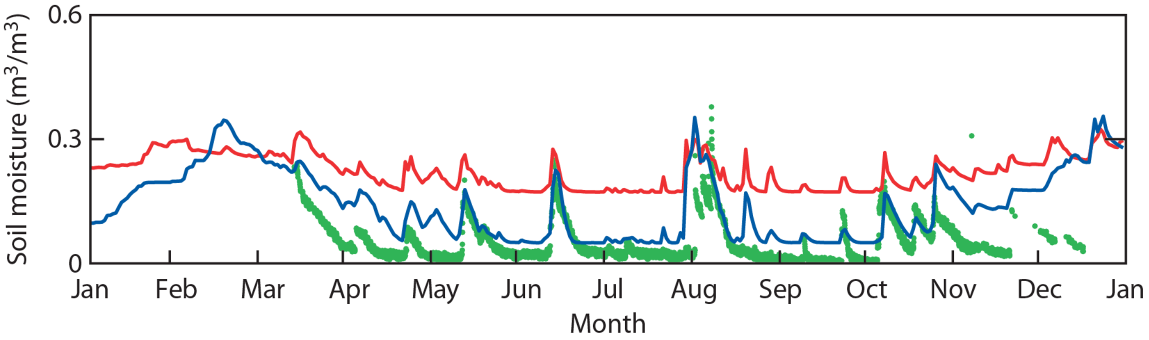


Figure 3: Evolution of volumetric soil moisture at a site in Utah for the year 2010. In-situ observations in green, ERA-Interim estimates in red, and ERA-Interim/Land estimates in blue.

The use of near-surface observations in the IFS, even at the spatial resolutions used for operational forecasts, is still severely limited both by poor model representativity near the surface and by shortcomings in the analysis method used. The 4D-Var analysis of upper-air prognostic variables is performed separately from simpler analyses of screen-level parameters (temperature, humidity) and land-surface parameters (soil moisture, soil temperature, snow depth). In the current setup, incremental 4D-Var updates of the atmospheric state are constricted by fixed (but uncertain) conditions at the surface, even if the available observations indicate that these have changed. The lack of coupling between the land surface and the atmospheric boundary layer in the analysis limits the ability of the IFS to represent fast surface interactions, e.g. associated with precipitation, and this has implications for forecast skill. A great deal of technical work is required to fully integrate the land-surface and atmospheric analyses, including the development of simplified physics needed for improved observation operators.

4. Reanalysis of atmospheric composition

Two reanalyses of global atmospheric composition have been produced within the framework of the collaborative European projects GEMS and MACC, as part of the development of an operational capability for global air-quality monitoring and forecasting at ECMWF. Both reanalyses were produced with an extended version of the IFS that includes chemically reactive gases, aerosols and greenhouse gases. Based on availability of global satellite observations of atmospheric composition, the reanalyses extend back only as far as 2003.

The extensions to the IFS developed in the GEMS and MACC projects allow integrated modelling of meteorological, chemical and aerosol variables, and combined use of observations of trace species and meteorology in the 4D-Var analysis. These are the basic elements needed for a fully coupled data assimilation system, in which observations of atmospheric constituents lead to physically consistent adjustments to the meteorological variables, and conversely, meteorological observations can have an immediate impact on estimates of the constituent concentrations. In principle, such a system can produce coherent global analyses with all estimated variables constrained by a single set of model equations. Coupled data assimilation potentially allows for better use of observations with information about both meteorology and aerosols or chemistry. These are important advantages over uncoupled or weakly coupled systems, in which either the model integration or the analysis of observations (or both) is performed in separate steps.

The actual impact of any single observation in a fully coupled data assimilation system depends on many factors, including the choice of control variables in the analysis, the background error covariances, and details of the forecast model itself. The increased complexity in the system requires a proportional increase in assumptions and choices to be made for its implementation. Fundamentally, a realistic analysis (as in true to nature) can be produced only if the additional degrees of freedom in the modelling system are adequately constrained by accurate observations. This has important implications for climate reanalysis, since the instrumental record available for a reanalysis of atmospheric composition is limited, both in quality and quantity.

Some of the pitfalls associated with coupled data assimilation are evident even in the Centre’s operational forecast system, which contains ozone as a prognostic variable. It was first noticed during the production of ERA-Interim that the 4D-Var analysis of ozone profile data often results in large and unrealistic changes in the upper stratospheric circulation, where the model background is not well constrained by observations (see Figure 4). These upper-level increments provided the most effective way for the 4D-Var analysis to accommodate the observed local changes in ozone concentration further below. It should be possible, in theory, to extract useful information about advection from stratospheric tracer observations in a 4D-Var analysis. In practice this can work well only if both the model background and the observations are sufficiently accurate, which is currently not the case.

As part of the MACC project an eight-year long reanalysis of atmospheric composition data was constructed for the period 2003-2010 (Inness et al 2012). This reanalysis is based on IFS Cy36r1 at resolution T255L60. The period 2003-2010 was chosen based on consideration of the available satellite data on atmospheric composition. The reanalysis is being extended to include 2011 and 2012 as part of MACC-II. Compared to a previous reanalysis performed in the GEMS project, the MACC reanalysis uses a higher model resolution, improved emissions, more sources of satellite data, and variational bias corrections for a subset of the satellite data.

MACC(-II) (Monitoring Atmospheric Composition and Climate) is a research project with the aim of establishing the core global and regional atmospheric environmental services for the European GMES (Global Monitoring for Environment and Security) initiative funded under the Seventh Framework Programme of the European Union. The project combines state-of-the-art atmospheric modelling with Earth observation data to provide information services covering European air quality, global atmospheric composition, climate, and UV and solar energy.

The global model and data assimilation system used in MACC are based on ECMWF's Integrated Forecasting System (IFS), which has been extended to include chemically reactive gases (Flemming et al. 2009; Inness et al. 2009), aerosols (Morcrette et al. 2009; Benedetti et al. 2008) and greenhouse gases (Engelen et al. 2009). Satellite observations of O3, CO, NOx, HCHO, CO2, CH4, and aerosol are used to constrain the system using the ECMWF 4D-Var data assimilation system. Source and sink terms for the reactive gases are supplied by the coupled MOZART-3 chemistry transport model (CTM), which features a full description of stratospheric and tropospheric chemistry (115 species). Tendencies due to chemistry, wet deposition and atmospheric emissions, and tendencies due to surface fluxes (emission, dry deposition) are all included in this coupling. The aerosol model distinguishes five types of tropospheric aerosols (sea salt, dust, organic and black carbon, and sulphate) with sources for each, sedimentation of particles, and wet and dry deposition processes. Stratospheric aerosols are currently obtained from climatology as in the ECMWF operational IFS. The greenhouse gases use prescribed climatological fluxes at the surface representing anthropogenic emissions, ocean fluxes, land biosphere fluxes, biomass burning, and wetlands, plus a climatological OH sink term for CH4 at each model level. On-going MACC-II developments aim for a prognostic representation of stratospheric aerosols, and introduction of full chemistry within the IFS rather than through the coupled CTM.

As a direct result of the discovery of this problem in ERA-Interim, the 4D-Var analysis in the operational forecast system was modified in 2007 to prevent any changes in temperature and wind resulting directly from the analysis of ozone data. This effectively breaks the connection between observations of changes in ozone concentration and information about the atmospheric flow. Recent improvements in the IFS, including implementation of variational bias corrections for ozone observations, have ameliorated the problem, and it is now being investigated whether a fully coupled 4D-Var analysis for ozone can be safely reinstated.

For the same reasons, the MACC coupled assimilation system similarly does not allow any trace-gas observations to modify the flow. Variational bias corrections are being developed for most of the constituents used in the MACC system as a prerequisite to a fully coupled 4D-Var analysis. Additional developments currently taking place in the MACC project will couple the atmospheric data assimilation with the carbon component of the land surface model (C-TESSEL). This then provides the ability to estimate surface fluxes of CO2 for the land biosphere that are fully consistent with the meteorology. Proper ways to constrain this coupling with observations are being investigated.

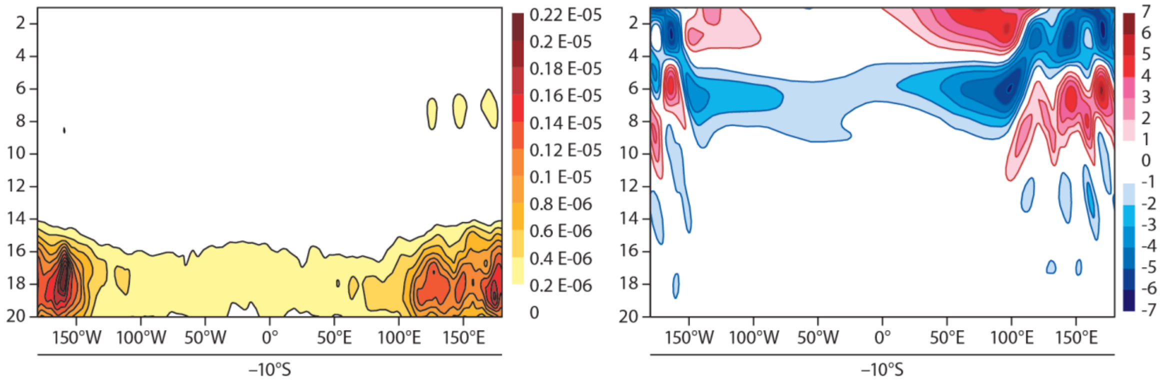


Figure 4: Impact of GOME ozone profile observations only, in a single 12h 4D-Var cycle (4 July 1995, 0 UTC), along the latitude circle 10S for the top 20 model levels (of a 60-level model, i.e., from 40hPa up to 0.1hPa). Ozone increments (left panel) with maximum values of about 2 g/kg are concentrated in locations where the satellite track crosses 10S. They are everywhere positive in this vertical plane, because the model ozone concentrations are biased low. Unrealistic temperature increments (right panel) ranging from -6.6K to +6.3K occur at much higher levels.

5. Reanalysis for climate services

Climate services include a wide range of activities that deal with generating and processing information about past, present and future climate and its effect on society and the environment. The development of the WMO Global Framework for Climate Services (GFCS; WMO 2011) foresees a prominent role for reanalysis as a key element of the observations and monitoring component of the GFCS. The need for a sustainable climate reanalysis capability in Europe is clearly recognized in the emerging plans for a GMES Climate Service component (Uppala et. al., 2011). Within GMES an important climate services component is already being developed in the MACC project, as its name indicates. It is only natural that the technical expertise and resources available at ECMWF, which are the result of many years of European investment, are increasingly enlisted in the effort to address the most pressing scientific problem of our times.

The use of atmospheric reanalysis data for climate change assessment has been, and still is, somewhat controversial (e.g. Thorne and Vose 2010). This is due to well-known difficulties with the representation of low-frequency variability in reanalysis data. Early generations of reanalyses, as well as some recent ones, show spurious shifts and other artifacts that can be identified with changes in the observing system, improper use of observations, transitions between multiple production streams, or various mistakes that can occur in a complex reanalysis production. Many of the issues are strictly technical, but clearly there are fundamental limitations as well to what can be achieved with incomplete observations and imperfect models.

Since the production of ERA-40 considerable progress has been made at ECMWF in addressing many of the technical issues just mentioned, resulting in a much better representation of climate variability and trends in ERA-Interim (e.g. Simmons et al. 2010). The progress is partly due to improved data assimilation, including the introduction of 4D-Var and the use of variational bias corrections for satellite observations (Dee and Uppala 2009). However, the importance of ongoing developments in technical tools and computing services at the Centre, including improved facilities for observation handling, monitoring and diagnostics, cannot be overstated. It is now relatively straightforward to maintain a reanalysis production in near-real time and to provide regular monthly updates of the dataset to a large number of users. A significant technical breakthrough was achieved with the recent 10-year backward extension of ERA-Interim in 2010, which demonstrated that it is indeed feasible to produce multidecadal atmospheric reanalyses in separate segments without introducing artificial discontinuities (see Figure 5). This has important practical implications for the planning of future reanalysis productions that extend even further back in time.

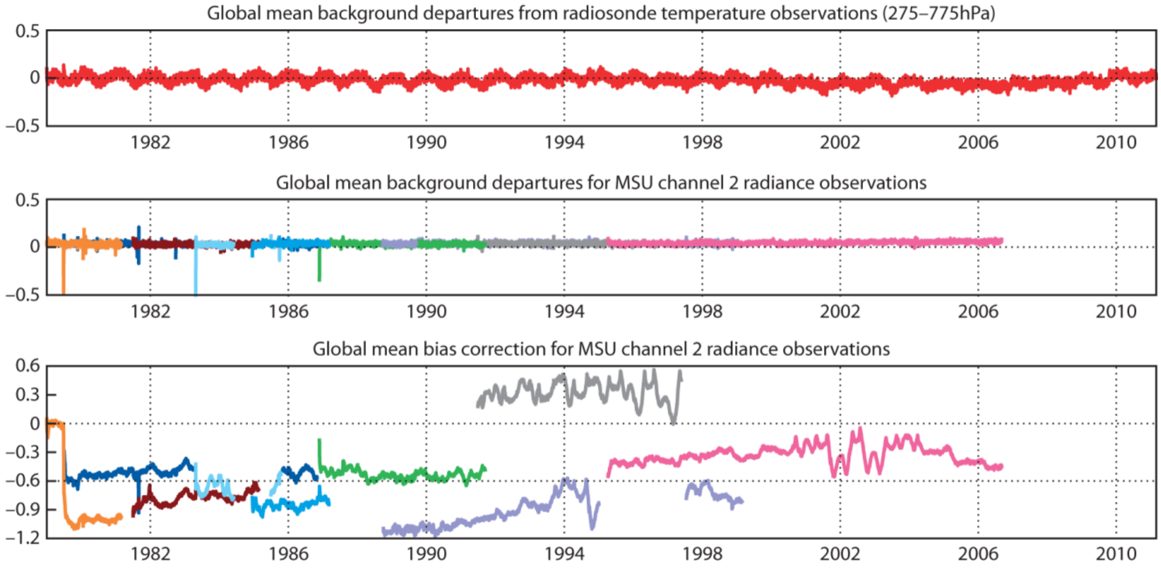


Figure 5: The three panels illustrate the stability and temporal consistency of the extended ERA-Interim reanalysis, and the nearly seamless transition between the two production streams on 1 January 1989. Reanalysed temperatures in the mid-troposphere are largely consistent with radiosonde observations (top panel) and with bias-corrected radiance measurements from Microwave Sounding Units flown on successive NOAA satellites (centre panel; colours indicate different satellites). The bias corrections for the MSU data, produced by the variational analysis in ERA-Interim, account for calibration differences, orbital drifts and various other instrument errors (lower panel).

Model-based reanalysis arguably offers the best approach for extracting maximum information about the recent climate from the existing instrumental record (Dee et al. 2010). Today's models encapsulate a great deal of knowledge about the climate system that can be used to relate and combine information from otherwise disparate observations in a physically meaningful way. This provides reanalysis with a distinct advantage over traditional observation-based methods, which rely primarily on spatial interpolation. Even though global observations from satellites have now been available for more than 3 decades, the current benchmark datasets for monitoring global temperature change, e.g. from the

Met Office Hadley Centre, NOAA/NCDC, and NASA/GISS, still have major gaps in regions that are critical for climate, such as the high latitudes and some parts of the tropics. Reanalysis methodology is now sufficiently mature to be able to match these observational estimates where they exist (Simmons et al. 2004, 2010; and see Figure 6). As a result, the usefulness of reanalysis data for monitoring the climate is increasingly recognized by the scientific community; data from ERA and MACC are now routinely included in the annual State of the Climate special issue of the Bulletin of the American Meteorological Society (SOC 2010; 2011; 2012).

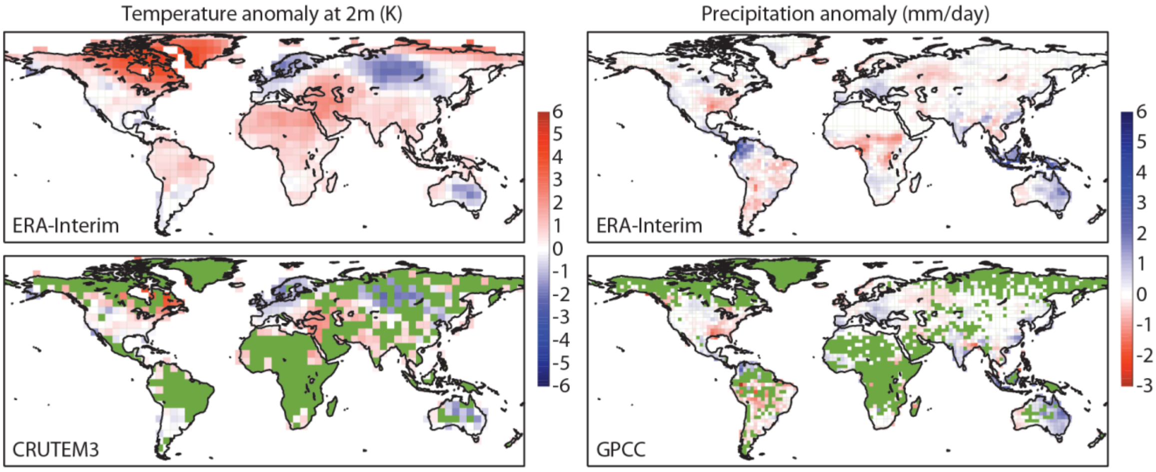


Figure 6: Anomalies for 2010 relative to 1979-2010 in surface air temperature (K; left) from ERA-Interim (upper) and CRUTEM3 (lower), and in precipitation (mm/day; right) from ERA-Interim (upper) and GPCC (lower). Values are plotted over land for grid-squares with a complete monthly data record for 2010 and no more than 12 missing months from 1979 to 2009. For GPCC, it is also required that there be at least one station per grid-box. Green colours indicate no data.

The use of a skilful model in reanalysis permits estimation of a large set of climate variables, even for variables that are not well observed, e.g. stratospheric winds, radiative fluxes, root-zone soil moisture, etc. (see Figure 7). These estimates are useful because they are indirectly constrained by the observations used to initialise the model. In the absence of direct observations, however, it is difficult to quantify the uncertainties in estimates of model-generated variables, as they depend on errors in the model as well as on the strength of the (indirect) observational constraint. Some insight into the uncertainties can be obtained by using ensemble techniques, with the important caveat that it is not practical to sample more than a few selected sources of uncertainty in a reanalysis.

Nevertheless, the complete description of a physically plausible atmosphere consistent with observations provided by reanalysis makes it possible to do many things that simply cannot be achieved otherwise. It permits, for example, detailed diagnostics of the global energy budget and the hydrological cycle (Trenberth et al. 2011). Such diagnostics are especially useful if they involve known time-invariant properties of the climate system. These are (usually) conserved by the assimilating model in a reanalysis, but tend to be destroyed by the assimilation increments, depending on the nature of the observational constraints and on the method of assimilation. Budget diagnostics can be used to demonstrate shortcomings as well as progress in climate reanalysis (Berrisford et al. 2011). Ironically, inconsistencies in the mass and energy budgets are often used to question the usefulness of reanalysis data for climate applications, even though it is clearly not possible to estimate most of the quantities involved from observations alone.

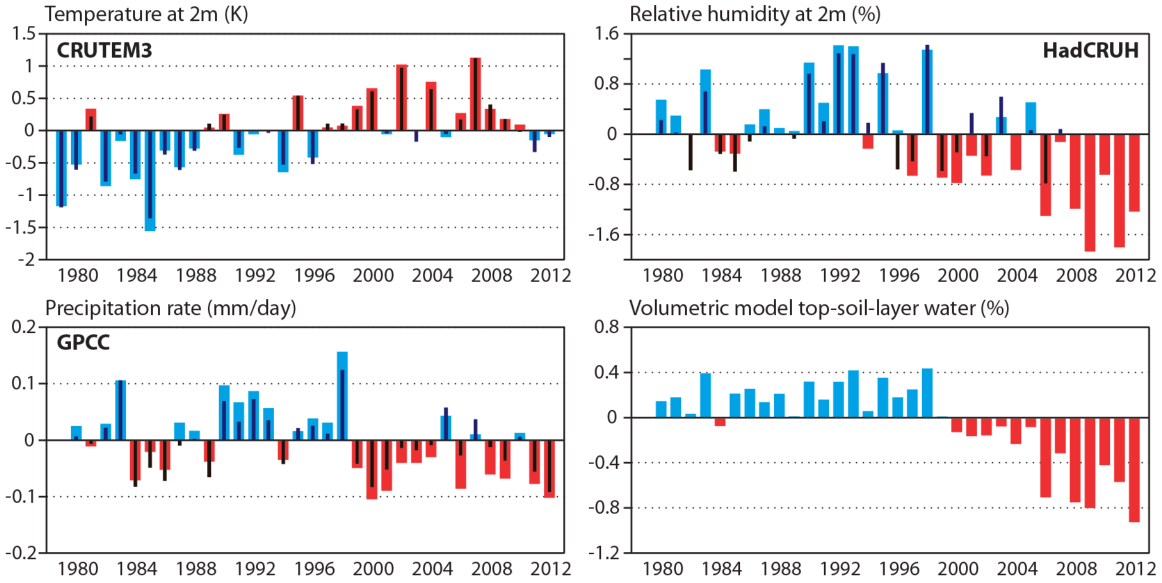


Figure 7: Dec-Mar anomalies averaged over northern-hemisphere land locations computed from ERA-Interim data and independently from in-situ observations. Light-coloured bars show ERA-Interim estimates; vertical darker lines show corresponding estimates from CRUTEM3 (for temperature; upper left), HadCRUH (for relative humidity; upper right), and GPCC (for precipitation, lower left).

There is still a great deal of room for improvement in reanalysis; various issues with the quality of ERA-Interim have been identified that remain to be addressed in subsequent reanalysis projects (Dee et al. 2011). These concern, for example, the consistency in time of mean precipitation over the tropical oceans, which, although much improved over ERA-40, is affected by incorrect use of rain-affected satellite radiances in the ERA-Interim system. Various aspects of the land surface representation in ERA-Interim, in particular snow cover, have been affected by problems with the input data as well as shortcomings in the surface analysis scheme. The energy balance at the surface boundary in ERA-Interim is poor, especially over tropical oceans due to a net increase in solar radiation.

As always, many of the improvements needed both in the forecast model and in data assimilation methodology are being addressed in current and upcoming IFS releases. However, several directions for development are needed specifically to address climate requirements: (1) reanalyses need to extend further back in time to provide a longer record for climate studies and climate model validation; (2) interactions and feedbacks between the atmosphere and other components of the climate system need to be better represented, and (3) users need to be provided with useful information about uncertainties relevant for the estimation of low-frequency variability and trends.

ERA-CLIM (European Reanalysis of Global Climate Observations) is a collaborative research project involving 9 partners, funded by the European Union for a three-year period through 2013. The project is coordinated by ECMWF. The goal is to prepare input data sets and assimilation systems for a new global atmospheric reanalysis of the 20th century to be undertaken at ECMWF.

Activities by the project partners include recovery, digitization, and quality control of early meteorological observations, reprocessing and recalibration of radiance measurements from satellites, and the preparation of climate-quality atmospheric forcing data and boundary conditions for the IFS. ECMWF will use these data sets to produce several new reanalyses, including an exploratory atmospheric reanalysis of the 20th century based on surface observations only; a corresponding high-resolution reanalysis of land-surface parameters, and a new atmospheric reanalysis of the satellite era from 1979 to present.

Using ERA-CLIM funding, ECMWF has developed a new set of technical tools and infrastructure for storing and retrieving feedback data using MARS. This Observation Feedback Archive will provide users with a powerful interface to observations used in ECMWF reanalyses, together with valuable data quality ('feedback') information such as background departures, bias estimates, and quality control decisions.

See www.era-clim.eu for additional information about ERA-CLIM.

ERA-20CM	Ensemble of model integrations 1900-2010 HadISST2 with CMIP5 forcing	125 km 10 members	Available end 2012
ERA-20C	Ensemble of atmospheric reanalyses 1900-2010 Surface observations only	125 km 10 members	Available mid 2013
ERA-20CL	Global land-surface reanalysis 1900-2010 Consistent with ERA-20C	40 km 10 members	Available mid 2013
ERA-SAT	Atmospheric reanalysis 1979-present To replace ERA-Interim	60 km 1 member	Available end 2014

6. Extended climate reanalysis

The production of a model-based reanalysis extending as far back as the instrumental record allows was first pursued in the 20th-Century Reanalysis Project (Compo et al. 2006) at NOAA's Earth Systems Research Laboratory. The project was based on the idea that a reanalysis of surface pressure observations only is relatively straightforward to compute and avoids many of the problems associated with observing system changes[1]. Furthermore, historic surface weather observations provide reasonable global coverage throughout the 20th century, and modern data assimilation methods are capable of reconstructing an accurate representation of the large-scale tropospheric circulation from surface pressure observations alone (Whitaker et al. 2009). Many of the earlier observations needed for this ground-breaking project were obtained from analogue sources, then digitized and collected in a global database (the ISPD, or International Surface Pressure Databank), which is now available for general use. A reanalysis of the 140-year period 1871 - 2010 was subsequently completed using NCEP's Global Forecast System (GFS) at resolution T62L28 and an ensemble Kalman filter especially developed for the purpose (Compo et al. 2011). Future versions have been proposed that will extend even further back in time.

At ECMWF a similar reanalysis for the period 1900 - 2010 is being prepared within the framework of the EU-funded ERA-CLIM project (see box). The ERA-20C reanalysis will be produced with the IFS as an ensemble of data assimilations (EDA) at resolution T159L91, using surface pressure and marine wind observations from the ISPD and ICOADS. The IFS forecast model has been modified to ingest climate data sets related to model radiation and land-surface parametrizations obtained from the Coupled Model Intercomparison Project Phase 5 (CMIP5). An ensemble of consistent global estimates of sea-surface temperature (SST) and sea-ice concentrations (SIC) for the entire period (HadISST2) has been developed for the project by the Met Office Hadley Centre, taking into account the uncertainties in the available observational sources (Figure 8). The ERA-20C reanalysis will comprise 10 members, each based on a different evolution of SST/SIC drawn from the HadISST2 ensemble, in order to provide users with information about at least one key source of uncertainty in the reanalysis. Using the tools described in Section 3, a set of global land-surface products consistent with the ERA-20C reanalyses but at higher spatial resolution will be produced as well.

The ERA-20C reanalysis can be regarded as an *extended climate reanalysis*, which differs from all earlier ECMWF reanalyses in that it makes use of a restricted set of observations and other input datasets specifically prepared for climate applications. In contrast, previous ERA reanalyses have assimilated the majority of observations used in operational forecasting, in the attempt to produce the best possible estimate of the atmospheric state at any given time. Clearly both types of reanalysis have an important role to play in climate studies and climate services. Extended reanalyses serve to provide the longest possible record of low-frequency variability and change consistent with observations, which is needed to put current large-scale anomalies in perspective. The spatial and temporal resolution that can be achieved in such a reanalysis will be limited mainly by the information available from observations. A shorter *reanalysis of the satellite era*, such as ERA-Interim, can provide a more detailed and complete view of recent changes taking place in the climate system, which can be continuously updated by making use of observations used for operational forecasting.

---

[1] It should be noted that any such reanalysis requires model boundary conditions, e.g. for sea-surface temperature, which can also be affected by changes in the observing system.

The two kinds of reanalyses have different requirements in terms of data usage. A traditional NWP-inspired reanalysis of the satellite era tends to assimilate all available observations unless they are known to be unusable or of poor quality; this is the familiar blacklisting approach. In contrast, an extended climate reanalysis ideally follows a whitelisting approach to data selection, where observations are used only if they are known to be of high quality and suitable for climate applications. In practice the distinction may not be quite as strict, since judgments on data quality are often difficult to make, but the contrasting objectives provide useful guidance. It implies, for example, that a climate reanalysis requires extra effort on selection and preparation of input data prior to assimilation, with preference given to observations obtained from compiled data collections that have been subject to some form of quality control to ensure temporal consistency.

The second major direction for development needed to better address climate requirements is coupled data assimilation, with an eye toward producing a more consistent description of mass and energy transport and a better representation of climate feedback mechanisms. The use of fixed model boundary conditions in an atmosphere-only reanalysis generates unrealistic surface fluxes and inconsistencies in the mass and energy balances. As reanalyses extend further back in time, uncertainties in the boundary conditions increase; e.g. see Figure 8 for the case of sea-surface temperature. A coupled reanalysis has the potential to extract information from near-surface observations to reduce these uncertainties. For example, the conventional in-situ observations used to construct gridded SST fields in the pre-satellite era, such as HadISST2, provide global information on a monthly timescale at best. In an atmosphere-only reanalysis the daily boundary conditions needed for the model are constructed simple by interpolating the monthly estimates; this yields physically unrealistic variability. In a coupled reanalysis, however, it should be possible to extract information about daily SST variability from marine winds and other near-surface weather observations. Even in the absence of observations, the reanalysis would generate global daily estimates of SST that are physically consistent in the sense of the model. The challenge, of course, is to constrain the drift in these estimates that can arise from systematic errors in the coupled model (e.g. see Figure 2).

The recently completed NCEP Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) is the first coupled global reanalysis of the satellite era (1979-present). CFSR uses a coupled model of the atmosphere, land surface, ocean, and sea-ice, and was specifically designed to support the development, calibration, and initialization of NCEP's seasonal forecast system. The data assimilation method used in CFSR is best described as weakly coupled, in the sense that it uses the coupled model only to generate background estimates for the analysis. The analysis itself is performed in separate and independent steps for each of the four model components. Biases in sea-surface temperature, sea-ice, precipitation and snow are constrained by means of strong relaxation to prescribed external datasets.

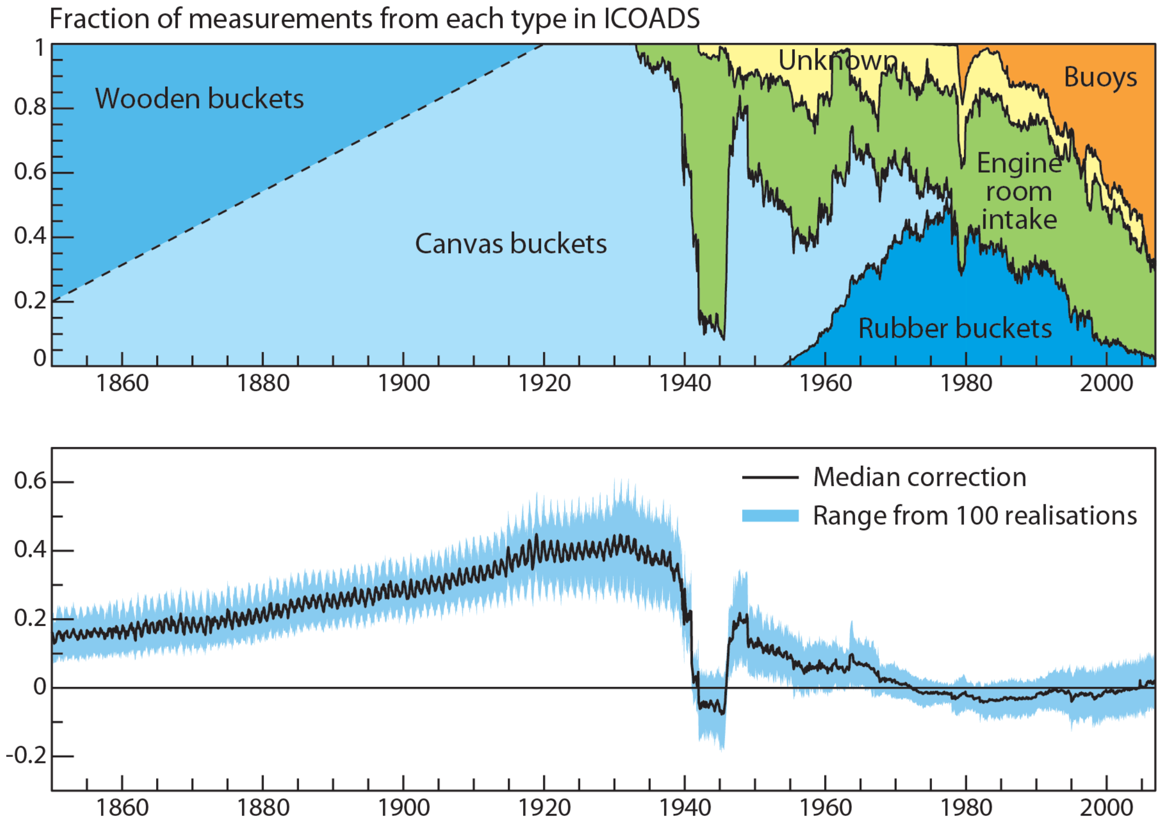


Figure 8: Representation of observational uncertainty in the HadISST2. The top panel displays the evolution over time of different measurement methods used for in-situ observations of sea-surface temperature, as a fraction of total. The exact proportion between wooden and canvas buckets used in the early part of the record is not known. The lower panel illustrates uncertainties in bias corrections applied in the construction of the HadISST2 ensemble of global SST fields. See also Kennedy et al. (2011a,b).

7. The CERA framework

Finally we describe a roadmap for planned IFS developments during the next 3-5 years that will permit the production of an extended climate reanalysis using coupled data assimilation for the atmosphere and ocean. These developments are specifically targeted for reanalysis but will also provide the IFS with the ability to initialise coupled forecasts, offering prospects for forecast skill improvements in the medium as well as the monthly to seasonal range. The work to be performed includes many preliminary technical steps required to incorporate and consolidate the tools and systems currently used for ocean reanalysis within the IFS-based ERA environment. Important details to be addressed include the use of different methods for handling input data streams, dealing with incompatible grids and output formats used in NEMO and IFS, possible MARS developments needed to accommodate the combined output, development of monitoring tools and new diagnostics for coupled systems, etc. These technical aspects are not discussed here, but they will require substantial effort and coordination. It is conceivable – and in many ways desirable – that the IFS developments described should move to the OOPS environment, but it is currently unclear whether and when this can be feasible.

We refer to the enhanced technical infrastructure for coupled climate reanalysis as the CERA (for Coupled ERA) framework. Application of this framework to produce a first ECMWF coupled climate reanalysis spanning the 20th century is the subject of a new proposal to the European Commission, currently in preparation by the Reanalysis Section. The proposed project, tentatively named ERA-CLIM2, will be presented as a natural continuation of ERA-CLIM that will further strengthen European capabilities for Earth-system reanalysis and climate services. If successful, it will provide ECMWF with substantial resources for research and development in coupled data assimilation, and will position the Centre for a pivotal role in future GMES climate services.

The following milestones summarize the incremental development of the CERA framework, which is schematically represented in Figure 9:

- Develop and implement a sequential model bias correction scheme to constrain the drift of the coupled IFS/NEMO model, using external estimates of monthly averaged SST. This provides the ability to produce an ensemble of coupled climate model integrations constrained by observed SST, e.g. as given by HadISST2.

- Extend the existing ERA framework by using the coupled IFS/NEMO model to compute the 4D-Var nonlinear trajectories. This provides the ability to assimilate atmospheric observations in the coupled model, while constraining the model drift using HadISST2.

  At this stage, re-linearizations in the 4D-Var outer loop are performed relative to a coupled-model trajectory. Observations used in the atmospheric analysis can therefore directly affect the ocean state. A potential benefit for an extended climate reanalysis is that the use of a coupled model can enhance SST variability in a physically plausible way – in contrast to e.g. simple interpolation as currently used in atmospheric reanalyses.

- Implement the linearized NEMOVAR minimization step in parallel with the IFS 4D-Var minimization. This provides the ability to assimilate both atmospheric and ocean observations in a coupled model.

  At this point, linear incremental updates for the ocean and atmosphere components are performed separately in the inner loops, and exchange of observational information takes place during the coupled model integration in the outer loop. Spin-up effects may occur within the analysis window if the separate analyses produce inconsistent information.

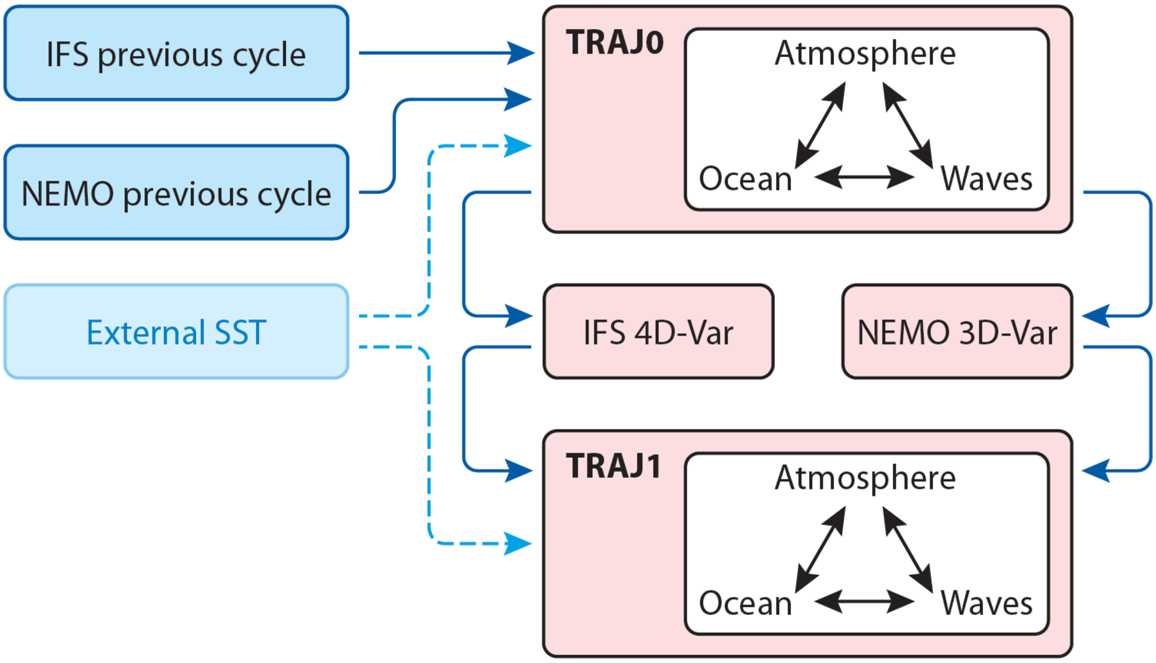


Figure 9: Schematic of the CERA framework for coupled data assimilation.

When using repeated outer loops, the CERA framework moves a step beyond ‘weakly coupled’ data assimilation as described in Section 6, in which the coupled model is used to produce a background estimate, but the analysis is performed in separate and independent steps.

The ERA-CLIM2 project, expected to start in 2014, will make use of the CERA framework to produce CERA-20C, a first coupled atmosphere-ocean reanalysis spanning the 20th century. Similar in concept to ERA-20C, it will consist of an ensemble of data assimilations at relatively low resolution, constrained by external SST and sea-ice products, and using a restricted set of observations from climate data collections. The off-line land-surface system described in Section 3 will be used to create a consistent but upgraded reanalysis of the land surface. In a similar manner, consistent supplementary reanalyses of carbon fluxes and stocks for the 20th century will be produced in collaboration with LSCE and Mercator Ocean, using their specialized versions of the ORCHIDEE land-surface and PISCES ocean biogeochemical models.

The ERA-CLIM2 proposal includes a substantial research and development component for coupled data assimilation targeted for inclusion in the CERA framework at ECMWF. This work will be done in collaboration with several external partners, including the Met Office, Reading University, Mercator Ocean, CERFACS, INRIA, and CMCC, who will together contribute approximately 13 person-years to the effort. The work envisioned includes incorporation of the LIM2 dynamic sea-ice model in the CERA framework, implementation of 4D-Var in NEMOVAR, use of ensembles to improve background error covariances for the coupled system, and development of the ability to directly assimilate SST observations in the system.

The CERA framework lends itself to many additional developments and refinements of the coupled data assimilation capability of the IFS. With a fully coupled nonlinear model implemented in the outer loop of 4D-Var, one can consider moving additional components of the analysis (e.g. for screen-level parameters) to the inner loops. Doing so requires work on simplified physics for the models involved, and development of tangent linear and adjoint codes. Once implemented, all inner-loop components can be computed in parallel, with information exchange taking place in the nonlinear outer loop only. A final step to a fully coupled version of incremental 4D-Var, where all inner-loop components share the same extended state vector, may not be worthwhile. Development would be rather difficult and the resulting system less flexible. Any additional benefit would depend on the ability to formulate accurate error covariance models for the coupled system.

8. Conclusion

Throughout the Centre’s history, reanalysis has been essential for the development of forecasting capability and skill. Reanalysis projects have led to numerous improvements in data assimilation, and the datasets produced are indispensable for model verification and development. Conversely, the evolution of the Centre’s operational products has gone hand in hand with the growth in diversity and sophistication of reanalysis products. As a result of recent advances in data assimilation, including an improved ability to handle biases and changes in the observing system, we can now foresee a central role for ECMWF reanalysis data in future climate services. This presents the Centre with a major opportunity to apply its unique technical capabilities to help address what is arguably the most pressing scientific problem of our times, namely adaptation and mitigation of climate change.

To best respond to society’s demand for climate information, while continuing to meet the Centre’s requirements for forecast system research and development, two different types of reanalysis products are needed. These are extended climate reanalyses, which reach back in time as far as the instrumental record reasonably allows, and reanalyses of the recent satellite-dominated observing system. The purpose of both types is to make the best possible use of available observations; however, the vastly changing characteristics of the observing system imply different approaches to data selection and usage in each case. The available observational constraints also have important implications for the types of models that can be used, and how they can best be constrained by means of data assimilation. In all cases there are, of course, fundamental limitations on the ability to represent what is essentially unobserved.

We have outlined plans for developing a coupled data assimilation capability in the IFS, targeted for the production of a consistent climate reanalysis. Resources needed to implement these plans are, for now, entirely dependent on available external funding. Many other aspects of data assimilation development, even though highly relevant for climate reanalysis, have not been discussed here. These include the application of weak-constraints 4D-Var to control the effects of model biases during the assimilation; the use of long analysis windows to take advantage of observations ahead of as well as behind the analysis time; extensions of the variational bias correction scheme for conventional observations; implementation of a hybrid data assimilation scheme to allow adaptive estimation of background error covariances. Reanalysis activities contribute to, and, in some cases, drive progress in these important areas.

This paper pays scant attention to observations, even though these in fact provide the essential element of value, and require by far the most effort in any reanalysis project. We have also omitted any discussion of improving access to reanalysis data, e.g. by providing better web services, visualization tools, and user guidance and support. Clearly these are all important elements needing full attention in a future climate-service oriented reanalysis activity at ECMWF.

9. References

Albergel, C., G. Balsamo, P. de Rosnay, J. Muñoz-Sabater, and S. Boussetta, 2012: A bare ground evaporation revision in the ECMWF land-surface scheme: evaluation of its impact using ground soil moisture and satellite microwave data, *Hydrol. and Earth System Sciences* (submitted).

Balmaseda, M., D. Dee, A. Vidard, D. L. T. Anderson, 2007: A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. *Q. J. R. Meteorol. Soc.* **133**: 167–179.

Balmaseda, M., K. Mogensen, A. T. Weaver, 2012: Evaluation of the ECMWF Ocean Reanalysis ORAS4. Submitted to *Q. J. Roy. Meteorol. Soc.*

Balmaseda, M., A. Vidard, D. L. T. Anderson, 2008: The ECMWF Ocean Analysis System: ORAS3. *Mon. Wea. Rev*. **136**, 3018-3034.

Balsamo, G., C. Albergel, A. Beljaars, S. Boussetta, E. Brun, H. Cloke, D. Dee, E. Dutra, F. Pappenberger, P. de Rosnay, J. Muñoz Sabater, T. Stockdale, F. Vitart, 2012: ERA-Interim/Land: A global land-surface reanalysis based on ERA-Interim meteorological forcing. *ERA Report Series* No. 13

Balsamo, G., S. Boussetta, E. Dutra, A. Beljaars, P. Viterbo, B. Van den Hurk, 2011: Evolution of land surface processes in the IFS, *ECMWF Newsletter*, 127, 17-22.

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

Balsamo, G., P. Viterbo, A. Beljaars, B. van den Hurk, M. Hirschi, A.K. Betts & K. Scipal, 2009: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *J. Hydrometeor*., **10**, 623–643.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., Suttie, M., and the GEMS-AER team, 2008: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Data Assimilation. *J. Geophys. Res*., D13205, **114**, doi:10.1020/2008JD011115.

Bengtsson, L., M. Kanamitsu, P. Kållberg, S. Uppala, 1982a: FGGE research activities at ECMWF. *Bull. Amer. Meteorol. Soc*., **63**: 277-303.

Bengtsson, L., M. Kanamitsu, P. Kållberg, S. Uppala, 1982b: FGGE 4 dimensional data assimilation at ECMWF. *Bull. Amer. Meteorol. Soc.*, **63**: 29-43.

Berrisford, P., P. Kållberg, S. Kobayashi, D. Dee, S. Uppala, A. J. Simmons, P. Poli, and H. Sato, 2011: Atmospheric Conservation Properties in ERA-Interim. *Q. J. Roy. Meteorol. Soc.,* **137** (659) (July 7): 1381–1399. doi:10.1002/qj.864.

Boussetta, S., G. Balsamo, A. Beljaars & J. Jarlan, 2011: Impact of a satellite-derived Leaf Area Index monthly climatology in a global Numerical Weather Prediction model. *ECMWF Tech. Memo*. No. 640.

Compo, G. P., J. S. Whitaker, and P. D. Sardeshmukh, 2006: Feasibility of a 100-Year Reanalysis Using Only Surface Pressure Data. *Bull. Amer. Meteorol. Soc.*, **87** (2): 175–190. doi:10.1175/BAMS-87-2-175.

Compo, G. P., J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin, B. E. Gleason, et al. 2011: The Twentieth Century Reanalysis Project. *Q. J. Roy. Meteorol. Soc.,* **137** (654): 1–28. doi:10.1002/qj.776.

ECMWF, 1985: *Proceedings of 1984 seminar/workshop on data assimilation systems and observing system experiments with particular emphasis on FGGE*. ECMWF, Reading, UK.

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

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

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae et al. 2011: The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System. *Q. J. R. Meteorol. Soc*., **137** (656): 553–597. doi:10.1002/qj.828.

Dutra, E., G. Balsamo, P. Viterbo, P. Miranda, A. Beljaars, C. Schär, K. Elder, 2010: An improved snow scheme for the ECMWF land surface model: description and offline validation. *J. Hydrometeor*., **11**, 899–916

Engelen R. J., Serrar, S., Chevallier, F., 2009: Four-dimensional data assimilation of atmospheric CO2 using AIRS observations, *J. Geophys. Res*., **114**, D03303, doi:10.1029/2008JD010739.

Flemming, J., A. Inness, H. Flentje, V. Huijnen, P. Moinat, M.G. Schultz and O. Stein, 2009: Coupling global chemistry transport models to ECMWF's integrated forecast system. *Geosci. Model. Dev*., **2**, 253-265.

Gibson, J. K., P. Kållberg, S.M. Uppala, A. Nomura, A. Hernandez, E. Serrano, 1997: ERA description. *ERA-15 Report Series* No. 1, ECMWF, Reading, UK.

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, 2009: Improving the global precipitation record: GPCP Version 2.1, *Geophys. Res. Lett*., **36**, L17808, doi:10.1029/2009GL040000

Ingleby, B., and M. Huddleston, 2007: Quality control of ocean temperature and salinity profiles - historical and real-time data. *J. of Marine Systems*, **65**, 158-175 10.1016/j.jmarsys.2005.11.019

Inness, A., F. Baier, A. Benedetti, I. Bouarar, S. Chabrillat, H. Clark, C. Clerbaux, P. Coheur, R. J. Engelen, Q. Errera, J. Flemming, M. George, C. Granier, J. Hadji-Lazaro, V. Huijnen, D. Hurtmans, L. Jones, J. W. Kaiser, J. Kapsomenakis, K. Lefever, J. Leitão, M. Razinger, A. Richter, M. G. Schultz, A. J. Simmons, M. Suttie, O. Stein, V. Thouret, M. Vrekoussis, C. Zerefos and the MACC team, 2012: The MACC reanalysis: An 8-year data set of atmospheric composition. *ECMWF Tech. Memo*. *No* 671.

Inness, A., Flemming, J., Suttie, M. and Jones, L., 2009: GEMS data assimilation system for chemically reactive gases. *ECMWF Tech. Memo*. No.587.

Kennedy, J. J., N. A. Rayner, R. O. Smith, D. E. Parker, and M. Saunby, 2011a: Reassessing Biases and Other Uncertainties in Sea-Surface Temperature Observations Measured in Situ Since 1850, Part 1: Measurement and Sampling Uncertainties. *J. Geophys. Res*., **116**: D14103. doi:10.1029/ 2010JD015218.

Kennedy, J. J., N. A. Rayner, R. O. Smith, D. E. Parker, and M. Saunby, 2011b: Reassessing Biases and Other Uncertainties in Sea-Surface Temperature Observations Measured in Situ Since 1850, Part 2: Biases and Homogenisation. *J. Geophys. Res*., **116**: D14104. doi:10.1029/2010JD015220.

Koster, R. D., S. P. P. Mahanama, T. J. Yamada, G. Balsamo, A. A. Berg, M. Boisserie, P. A. Dirmeyer, F. J. Doblas-Reyes, G. Drewitt, C. T. Gordon, Z. Guo, J.-H. Jeong, W.-S. Lee, Z. Li, L. Luo, S. Malyshev, W. J. Merryfield, S. I. Seneviratne, T. Stanelle, B. J. J. M. van den Hurk, F. Vitart, and E. F. Wood, 2010: The Second Phase of the Global Land-Atmosphere Coupling Experiment: Soil Moisture Contributions to Subseasonal Forecast Skill, *J. Hydrometeor*., doi: 10.1175/2011JHM1365.1

Madec, G., 2008: NEMO reference manual, ocean dynamics component : NEMO-OPA. Preliminary version. *Note du Pôle de modélisation* 27, Institut Pierre-Simon Laplace (IPSL), France.

Morcrette, J. J. et al., 2009: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling, *J. Geophys. Res*., **114**, D06206, doi:10.1029/2008JD011235

Saha, S., S. Moorthi, H. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp et al., 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteor. Soc*., 1010-1057. doi:10.1175/2010BAMS3001.1.

Simmons, A. J., K. M. Willett, P. D. Jones, P. W. Thorne, and D. P. Dee, 2010: Low-frequency variations in surface atmospheric humidity, temperature and precipitation: Inferences from reanalyses and monthly gridded observational datasets. *J. Geophys. Res*., **115**, D01110, doi:10.1029/2009JD012442.

Simmons, A. J., P. D. Jones, V. da Costa Bechtold, A. C. M. Beljaars, P. W. Kållberg, S. Saarinen, S. M. Uppala, P. Viterbo, and N. Wedi, 2004: Comparison of trends and low-frequency variability in CRU, ERA-40 and NCEP/NCAR analyses of surface air temperature, *J. Geophys. Res*., **109**, D24115, doi:10.1029/2004JD005306.

SOC 2010: State of the Climate in 2009. *Bull. Amer. Meteor. Soc*., **91** (7), 2010.

SOC 2011: State of the Climate in 2010. *Bull. Amer. Meteor. Soc*., **92** (6), 2011.

SOC 2012: State of the Climate in 2011. *Bull. Amer. Meteor. Soc.*, **93** (7), 2012

Tang, Y. M., M. A. Balmaseda, and K. Morgensen, 2012: Sea-ice modelling in a coupled global ocean-ice model. Manuscript in preparation.

Thorne, P. W., R. S. Vose, 2010: Reanalyses suitable for characterizing long-term trends. *Bull. Amer. Meteorol. Soc*. **91**: 353–361.

Trenberth, K. E., J. T. Fasullo, and J. Mackaro, 2011: Atmospheric Moisture Transports From Ocean to Land and Global Energy Flows in Reanalyses. *Journal of Climate* 4907-4924 (24): doi:10.1175/2011JCLI4171.1.

Uppala, S. M., D. P. Dee, S. Kobayashi, A. J. Simmons, 2008: Evolution of reanalysis at ECMWF. In *Proceedings of Third WCRP International Conference on Reanalysis*, 28 January–1 February 2008, Tokyo, Japan.

Uppala, S. M., Kållberg P., A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson, et al. 2005: The ERA-40 Re-Analysis. *Q. J. Roy. Meteorol. Soc.,* **131** (612): 2961–3012. doi:10.1256/qj.04.176.

Uppala, S. et al, 2011: GMES climate service: towards a european knowledge base in support of mitigation and adaptation. *Draft Report European Commission* DG ENTR/H.4/CC (2011).

van den Hurk, B., P. Viterbo, A. Beljaars, and A.K. Betts, 2000: Offline validation of the ERA-40 surface scheme. *ECMWF Tech. Memo*. No. 295.

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

Wijffels, S., J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K. Ridgway, J. A. Church, 2009: Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. *J. Climate*. **21**: 5657–5672.

WMO, 1985: *GARP Special Reports* Nos. 42, 43, and 45. World Meteorological Organization, Geneva, Switzerland.

WMO, 2011: Climate Knowledge for Action: A Global Framework for Climate Services. WMO-No.1065, World Meteorological Organization, Geneva, Switzerland.