CMUG Deliverable

Number:	**D3.4**
Due date:	**March 2013**
Submission date:	**15 May 2013**
Version:	**0.3**

Climate Modelling User Group

Deliverable 3.4

Technical note on The role of Reanalysis in the production and quality assessment of CDRs

Centres providing input: ECMWF

Version nr.	Date	Status
0.1	30 Jan 2013	Initial Template RD
0.2	05 Apr 2013	First draft RD
0.3	15 May 2013	Sent to ESA for comment

Deliverable 3.4

Technical note

The role of reanalysis in the production and quality assessment of CDRs

Table of Contents

The role of reanalysis in the production and quality assessment of CDRs

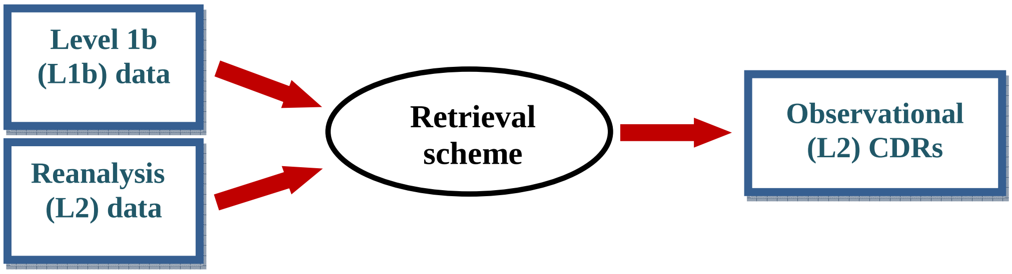
1. Purpose and scope of the Technical note

By combining past observations with a state-of-the-art model, reanalysis provides the most complete, coherent, and comprehensive set of data that could be exploited in the production and quality assessment of Climate Data Records (CDRs). This document discusses the value of reanalysis as a resource for the generation and assessment of data with sufficient quality to be used in climate studies. After briefly introducing the possible linkages between observations and reanalyses (section 2), the remaining of the present paper is structured into two parts. Part 1 (section 3) presents detailed discussions on how a reanalysis is produced, what reanalysis streams are available to date and the consolidated plans for the future, what the key issues and challenges are and how its quality can be monitored and assessed. Part 2 (section 4) discusses how reanalyses can be used in the quality assessment of independent observations with a focus on low-frequency (multi-year) variability.

2. Links between observations and reanalyses

Reanalysis data could be used either as prior / auxiliary information in the retrieval algorithms used to obtain observational CDRs, or as an independent dataset which observations can be compared to. These two cases are schematically illustrated in figure 1 in the case the CDRs, e.g. the Essential Climate Variables (ECVs) defined by the GCOS (see Appendix B for acronyms not defined in the text).

A) Climate Data Record production:



B) Climate Data Record validation:

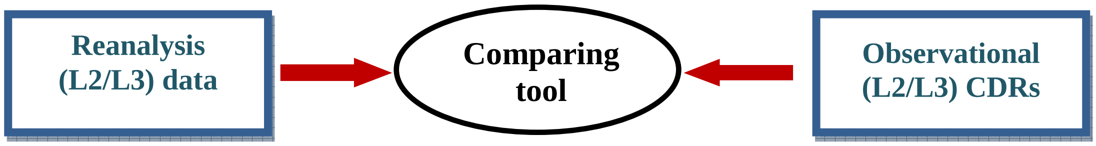
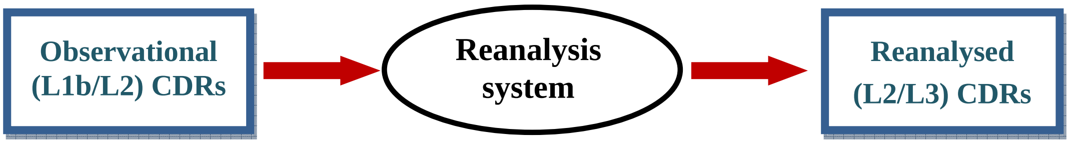


Fig 1: Schematic illustration of how a reanalysis can contribute to the production and quality assessment of CDRs.

The possible ways the CDR validation can be performed are discussed in detail in Part 2 of the present document.

Whether reanalysis data is used directly as auxiliary information in the retrieval schemes (case **A** in figure 1) or indirectly as data which the CDRs can be compared to (case **B** in figure 1), a key requirement is that the reanalysis itself has climate quality and its data products are CDRs. Only recent advances in the model physics and in the characterisation and use of observations have enormously improved the quality of modern reanalyses with consequent growing awareness that one of their possible applications could be the production of CDRs. It is important to recognize that there are a number of key issues and challenges in the production of a reanalysis, so that a climate quality may only be achieved for a subset of variables or during specific periods of time. The key issues and challenges in reanalysis are discussed in detail in section 3.2. As discussed in Part 1 of this document, one condition to ensure a reanalysis product can achieve climate quality is that the observations assimilated during its production possess climate quality, i.e.:



3. Part 1

3.1. *Comprehensive reanalysis of the instrumental record*

A retrospective analysis, or reanalysis, is a scientific method for producing a comprehensive and consistent long-term data record of how weather and climate have been changing over the period under consideration, typically of several decades. This is achieved by integrating observations available from a variety of data sources together within a fixed, state-of-the art model that describe one or more components of the climate system, e.g. the atmosphere.

The use of a fixed model represents a major difference between reanalysis and daily prediction. The daily analyses that aim at enabling the best short-term weather forecasts are conducted with models that are frequently updated, sometimes several times per year. These updates can generate false changes in the analysis records to the point of limiting their value for climate applications. In contrast, reanalysis can be optimized to achieve other objectives such as providing a consistent description of the climate system over an extended time period.

An overview of the past, current and future reanalysis production encompassing the various components of the climate system (including those for coupled reanalyses) is provided below.

Box A: Atmospheric, Land and Ocean Reanalyses

A number of different reanalyses have been realised over the last thirty years, and a new generation (the forth) is currently in production. Most of these reanalyses focussed on the atmospheric domain, thanks to the vast amount of observations available and the advances in atmospheric models driven by weather forecasting. Table 1 lists the established atmospheric reanalyses currently available. References have been provided for each data stream, and the reader is advised to refer to the literature for more in-depth description of the characteristics of each reanalysis production. Further details can also be found at http://www.reanalysis.org

The situation is somewhat different for the oceans. Here routine observations at the same location have been limited to ships or moorings. Satellite data, although important, usually give only information about the ocean surface. Despite the data volume has largely increased over the last ten-fifteen years, it is still small compared to the data volume available for the atmosphere. It is worth mentioning that ocean reanalyses strongly depend on atmospheric reanalyses, and so their quality. A large number of ocean reanalyses have been produced over the last few years. A detailed list of consolidated ocean reanalyses is available at http://icdc.zmaw.de/easy_init_ocean.html?&L=1#c2231 and links therein.

In addition to the reanalyses for atmosphere and ocean, there are a few examples of off-line land-surface simulations associated to given atmospheric reanalysis productions. These simulations are forced by the reanalysis meteorological fields (temperature, surface pressure, humidity and wind). They are useful for land-model development while also offering an affordable way to improve the land-surface component of the original reanalysis. At least two examples have been produced to date. One, called ERA-Interim/Land, is the land off-line simulation associated to the ECMWF ERA-Interim atmospheric reanalysis (Balsamo et al. 2012). The other, MERRA-Land, is the corresponding NASA simulation for MERRA (Reichle et al. 2011).

At the time of writing, a fourth generation of reanalyses is under preparation. ECMWF leads a consortium that aims at producing a reanalysis - ERA-CLIM - that will span the whole 20th century (details can be found at www.era-clim.eu). A similar collaborative reanalysis - The Twentieth Century Reanalysis Project - has been led by NOAA and CIRES in the USA to produce an accurate representation of the large-scale tropospheric circulation by using only surface pressure data (Combo et al, 2011). The main difference between these climate reanalyses and their predecessors is in the use of selected observations of proven quality and specifically produced for climate studies.

It is also important to recognize that there is a growing worldwide effort in producing fully coupled reanalyses. Having fully coupled systems is of paramount importance to produce realistic fluxes at the interface between the atmospheric and oceanic components of the climate system as they can help understanding the forcing and interactions in it. Fully coupled data assimilation systems are not yet available, but first examples of reanalysis production could be realised within the next 3 to 5 years from now. ECMWF, for example, is currently leading the development of the CERA system that will generate its first extended climate reanalyses using a fully coupled data assimilation system for atmosphere and ocean. The project (Dee et al 2012) will cover the whole 20th century and it is expected to be the follow-on project of ERA-CLIM (subject to funding).

Generation	Name	Centre	Model	Resolution	Period	Completed	Products	Resolution Analyses[2]	Reference
Precursor	FGGE	ECMWF	1979	N48 L15	1979	1981	Meteo [1]	6-hourly[3] at 1.875°x1.875°	Bengtsson et al, 1982
I	ERA-15	ECMWF	1994	T106 L31	1979-1993	1996	Meteo	6-hourly at 1.125°x1.125°	Gibson et al 1997
I	NASA/DAO	NASA	1994	2.5^{O}x2^{O} L20	1980-1993	1993	Near surface meteo, surface fluxes	6-hourly at 2.5°x2°	Schubert et al 1993
II	NRA-1	NCEP/NCAR	1995	T62 L28	1948-1988	Completed	Meteo	6-hourly at 2.5°x2.5°	Kalnay et al 1996
II	ERA-40	ECMWF	2001	T159 L60	1957-2002	2003	Meteo, O_3	6-hourly at 2.5°x2.5°	Uppala et al 2005
II	JRA-25	JMA	2004	T106 L40	1979-present	Updated	Meteo	6-hourly at 1.25°x1.25°	Onogi et al 2007
II	NRA-2	NCEP/DOE	1994	T62 L28	1979-2012	2012	Meteo	6-hourly at 2.5°x2.5°	Kanamitsu et al 2002
II	GEMS	GEMS consortium	2006	T159 L60	2003-2009	2010	Composition, aerosols	6-hourly at 2.5°x2.5°	Hollingsworth et al 2008, Benedetti et al 2009
III	ERA-Int	ECMWF	2006	T255 L60	1979-present	Updated monthly	Meteo, O_3	6-hourly at 0.75°x0.75°	Dee et al 2011
III	MERRA	NASA/GMAO	2008	2/3^{O}x0.5^{O} L72	1979-present	Updated	Meteo, O_3	6-hourly at 2/3^{O}x0.5^{O}	Rienecker et al 2011
III	NCEP-CFSR	NCEP	N/A	T382 L64	1979-2008	2010	Atmospheric, oceanic, and land surface	Hourly at 0.5x.5°	Saha et al 2010
III	JRA-55	JMA	2009	T319L60	1958-2012	To be released mid-2013	Meteo	6-hourly	Ebita et al 2011
III	MACC	MACC consortium	2010	T255 L60	2003-2010	2011	Composition, aerosols	6-hourly at 0.75°x0.75°	Inness et al 2013

Table 1: List of consolidated atmospheric reanalyses.[1]Meteo refers to both surface and 3D fields unless otherwise specified. [2]The analysis resolution represents the highest (temporal or spatial) resolution available. [3]6-hourly means at the four main synoptic times (00, 06, 12, and 18 Z).

The result of the data-model integration is a comprehensive, complete, temporally continuous, physically consistent and homogeneous dataset of variables for use in climate research and applications. It is comprehensive because it includes a large set of variables that together provide an adequate description of the component(s) of the climate system it models. It is complete in the sense that they do not simply produce retrospective analyses of meteorological fields. They also provide useful additional information, e.g. on the observation quality control, on the fit of these observations to their model equivalent (background and analysis), and an estimate of the observation biases. It is temporally continuous as its output fields are provided with no spatial or temporal gaps (typically every 6 or 12 hours). It is physically consistent because the model constrains the analyses to be consistent with the fundamental laws of physics. One important implication is that the different state variables depend strongly on one other. This dependence is an important aspect to understand e.g. forcing or physical processes that governs the evolution of the climate system. It provides temporal homogeneous datasets as a reanalysis makes use of fixed data assimilation system and physical model.

In recent years, with the improvements on both the model physics and in the exploitation of observations, the awareness that reanalysis could successfully be used for climate studies has largely increased (subject to the caveats discussed in section 3.2) to the point that the fourth generation of reanalyses is already referred to as <u>climate reanalyses</u>.

What is a climate reanalysis, then? What are the differences with a standard reanalysis? and what role does reanalysis play within a comprehensive climate observing system?

A climate reanalysis is a reanalysis that produces datasets with the same characteristics of CDRs. The latter are defined by the National Research Council as **time series of measurements of <u>sufficient length</u>, <u>consistency</u>, and <u>continuity</u> to enable study and assessment of long-term climate change, with ‘long-term’ meaning year-to-year and decade-to-decade changes.**

A climate reanalysis uses a fixed numerical prediction model and data assimilation method that assimilates quality-controlled observations over an extended period of time, typically several decades, to create a long period climate record. There are two main differences between a climate reanalysis and any of the reanalyses produced so far. The first aspect poses the accent on the “quality-controlled observational data”, i.e. only observations of proved quality are assimilated and used. This might not have been always the case so far, when a weather-forecasting type of approach (of discarding observations only when they were proved to be of poor quality) could have been used. The second aspect relates to the temporal coverage of climate reanalysis productions. These are expected to stretch over about a century long period. Any reanalysis produced over a shorter period of time cannot adequately be used to understand changes in the climate system. Identifying the cause(s) of these changes gives a scientific underpinning for predicting future climate.

It is not possible to have accurate projection of the future climate without being able to accurately reproduce the past climate.

Climate reanalysis plays a central role in assessing how and how much the climate system has changed. Observations are at the core of a reanalysis. They can affect its quality and limit/increase its ability in reproducing the evolution of the climate system. However, they cannot on their own answer the fundamental questions of how, how much and eventually why the climate system has changed over the past century or so. This can only be achieved when these observations are used together with a model that describes at best our understanding of the key physical processes and how these are related one another.

Understanding the strengths and limitations of current reanalysis data, including representations of climate changes and trends, increases our confidence in these products, but that inevitably depends on the careful use and exploitation of the assimilated observations. Comparisons of datasets from different reanalyses and observational sources are a valuable mean to provide a measure of the uncertainty in these products and how well they represent past climate.

3.2 Key issues and challenges related to climate quality reanalysis

It was said above that reanalysis uses advanced statistical methods to assimilate observations from multiple sources into a state-of-the-art forecast model. The result is a physically and dynamically coherent global dataset that comprises several ECV estimates over several decades and is consistent with both the observations and the laws of physics.

However, in reality, reanalyses combine an inaccurate and incomplete Global Observing System (GOS) with imperfect models. Its quality is impacted by a number of practical decisions and compromises on the analysis methodology, on the data quality control, on the choice of observations and the description of their error characteristics, on how they are used, and of course on the model. CMUG (2013, section 3) has already presented an overview of some of the challenges and key issues that can affect the quality of a reanalysis production. Those aspects are now reviewed and discussed in a wider context.

The first difficulty one needs to understand is that the GOS used in reanalysis was never designed for climate studies and for long-term climate variability assessments. It was instead designed for weather forecasting. Observations are normally obtained from a variety of different sources (surface, upper air, satellite data) yet they do not provide complete spatial coverage of all relevant components of the climate system. This means that the available GOS often does not provide long-term, comprehensive and consistent observations of the climate system, including observations of the land and ocean. These are critical to understand and predict atmospheric variability over seasonal and longer time periods. In order to adequately detect the climate long-term variability, well characterised observations would need to be continuously available for many decades, observe the whole globe, include all key climate parameters, and be consistent with our best physical understanding. This has also been recognised by GCOS that defined a number of key requirements for a climate observing system that should - to mention some:

- Give high priority for additional observations to data-poor regions, poorly observed parameters, regions sensitive to change, and key measurements with inadequate temporal resolution.

- Have suitable period of overlap between new and old satellite systems adequate to determine inter-satellite biases and maintain the homogeneity and consistency of time-series observations.
- Provide continuity of satellite measurements (i.e. elimination of gaps in the long-term record) through appropriate launch and orbital strategies.
- Sustain operational production of priority climate products.
- Have rigorous pre-launch instrument characterization and calibration.
- Have on-board calibration adequate for climate system observations and monitoring of the associated instrument characteristics.

Despite these directives, the current reanalysis productions are still very much affected by the overall inadequacy of the existing GOS, and a number of challenges and issues need to be addressed when preparing a new reanalysis. One of the key limitations of current and foreseeable observing systems - that reflects in most of the challenges outlined in the box below - is that they do not provide complete coverage in time, space and of all relevant components of the climate system. Many of these limitations determine some of the key challenges in a reanalysis production (see Box B).

Box B: Challenges in Reanalysis

Challenge 1: Data rescue

A reanalysis to be used for climate change assessment must extend over several decades so that climate signals can be appreciated. To this end, a great effort is needed to collect all the observations required. This task is not trivial. Over the timeframe spanned by a climate reanalysis, the GOS has undergone great changes (figure 2). Until the mid-twentieth century, the observing system mainly consisted of surface-based observations, typically limited to land areas and ship reports. An upper-air radiosonde network of observations over land (particularly in the Northern Hemisphere) only became available in the late 1940s. While a global observing system only became available in the 1970s with the advent of satellites.

In many cases, the early surface observations are not even available in a digitised form but only archived in local reports. In these cases, the data need to be imaged, digitised and undergo a preliminary screening before being considered for assimilation.

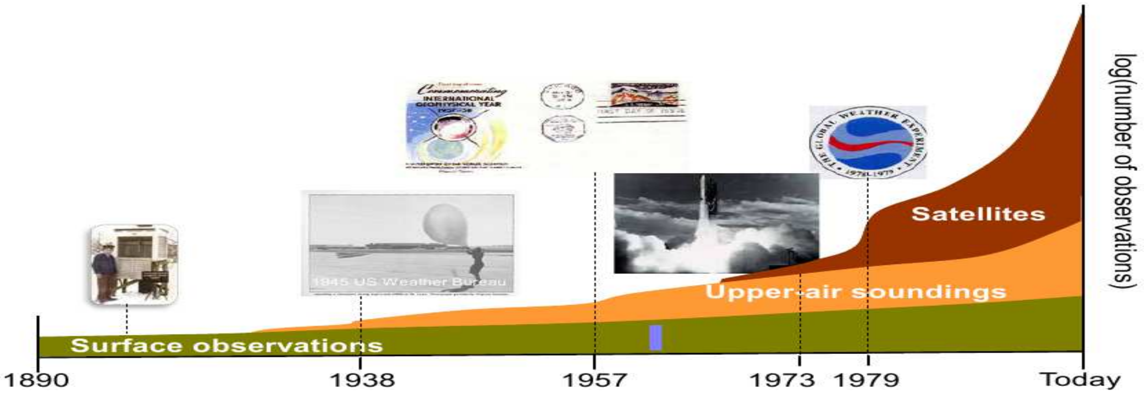


Figure 2: A schematic representation of the changes in observation type and data volume from 1890 to the present days.

Challenge 2: Detection of calibration errors

Observations are affected by a number of different issues (such as calibration problems and more generally systematic errors) that can limit their reliability if not properly accounted for. Figure 3 shows the NOAA-14 MSU channel 2 recorded warm target temperature change due to the satellite orbital drift (bottom panel) as reported by Grody et al. 2004, and the bias correction that was applied during the assimilation in ERA-Interim. The resemblance between the w-shaped feature detected by the ERA-Interim bias correction (Auligné *et al*, 2007) and that in the warm temperature target due to the satellite orbital drift visible from mid-2001 to 2003 is remarkable. If the observations had not been corrected for it, the ERA-Interim temperature reanalyses would have presented a similar (artificial) feature.

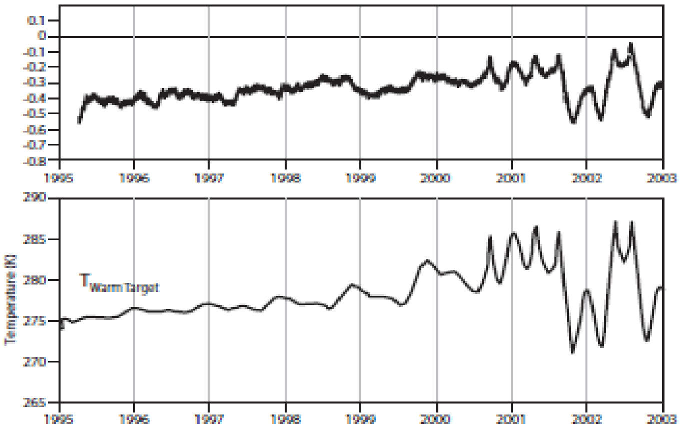


Figure 3: Bottom panel: recorded variations of the warm-target calibration temperature on board NOAA-14 from Grody et al. (2004). Top panel: global mean bias estimates from ERA-Interim for NOAA-14 MSU channel 2.

Challenge 3: Homogenisation of observations with different temporal coverage

A data assimilation system uses a statistical method to ensure that, in the absence of bias with respect to the true state of the climate system, the observations and model first guess are combined in an optimal way to minimize their errors (under the hypothesis that these errors follow a normal distribution). Therefore, ideally a data assimilation system would be presented with observations corrected for any systematic error. This is not yet a standard practice, so that any systematic error in the data records has to be removed at the time of the assimilation using bias correction schemes (Dee, 2005). Advances in the development of these bias correction schemes have been such that a number of inconsistencies in the GOS can nowadays be successfully detected and corrected for, reducing the occurrence of artificial features and jumps in the final reanalysis products.

One of the biggest challenges is represented by the homogenisation of the large number of different instruments available during the timeframe spanned by a climate reanalysis.

For example, surface observations have been available continuously over the 1900s, but the manufacture of the instruments utilised to make these measurements evolved over time and even at one given time they could have been different from one station to another. Information about these changes and how the measuring systems evolved over time is crucial to determine the potential bias affecting the corresponding measurements, but not always (well) documented.

In the case of satellite observations, difficulties arise when merging together observations with the same nominal characteristics but obtained from different sources. If these observations are not harmonised prior or during the assimilation, inter-instrumental biases may affect the resulting analyses. Figure 4 (adapted from figure 10 of CMUG, 2013) shows the bias corrections in brightness temperature (middle panel) that were applied to the channel 2 observations measured by various MSU sensors on the NOAA platforms (each indicated by a different colour), so that the resulting observation minus temperature background departures (top panel) were mostly unbiased. Comparisons between radiosonde temperature observations and the same temperature background (bottom panel) show a good agreement in the mid troposphere, reassuring that the applied corrections were appropriate.

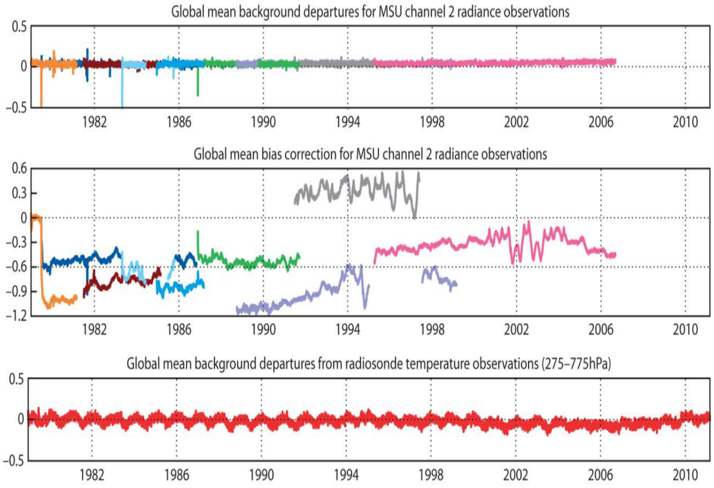


Figure 4: The top panel shows the bias-corrected observation minus background departures from MSU channel 2 on successive NOAA satellites (colours indicate different satellites). The applied global mean bias corrections for the MSU are shown in the middle panel. The comparison between the ERA-Interim temperature and radiosonde observations is plotted in the bottom panel.

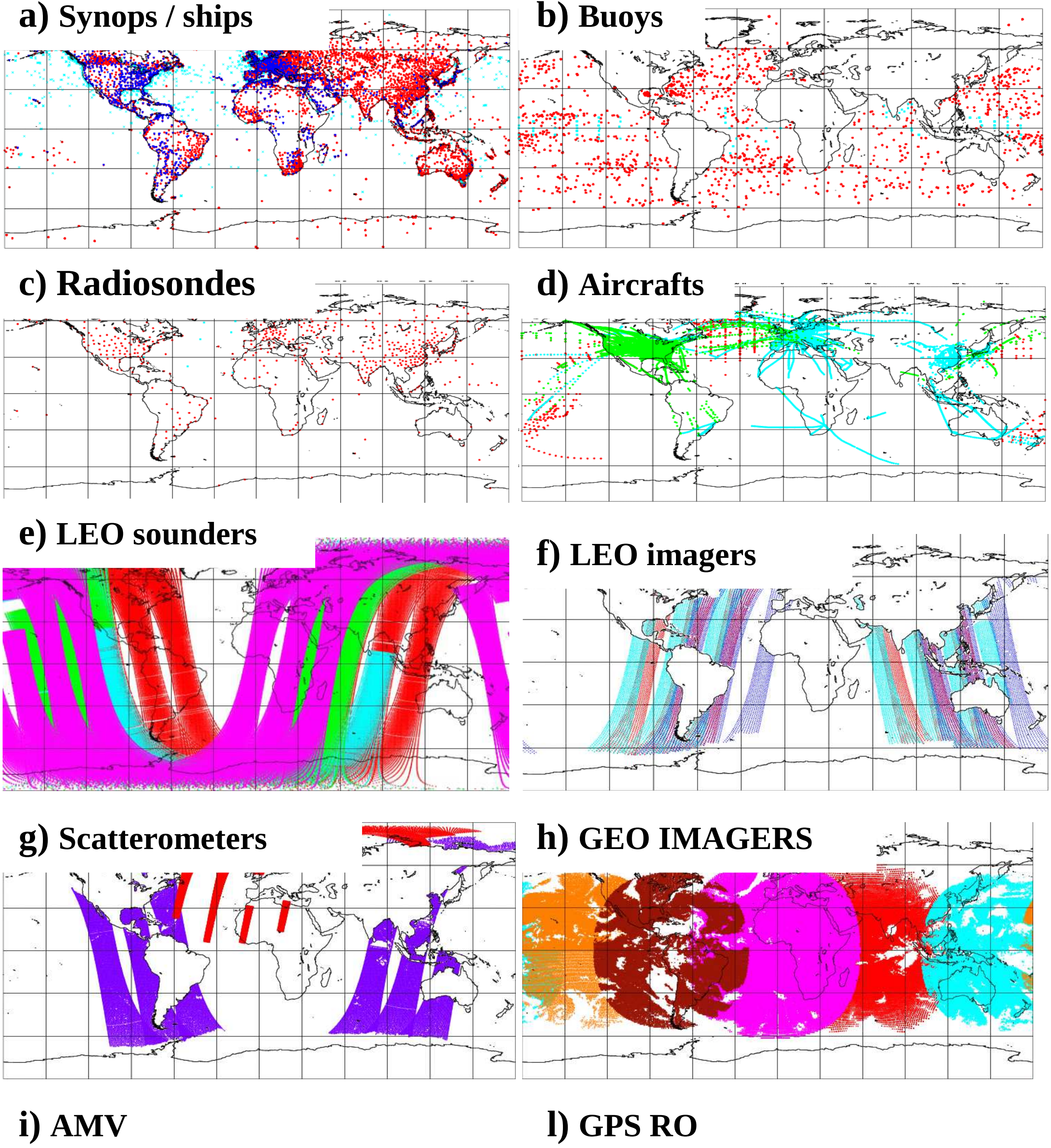
A more complex situation occurs when an instrument type is dismissed and replaced with one presenting slightly different characteristics. In these situations, the bias correction scheme may be less successful in detecting the most appropriate corrections to apply so that the resulting reanalyses could present artificial changes. CMUG (2013) discussed one case (see their figure 11) of a discontinuity in the ERA-Interim stratospheric temperature reanalyses that occurred in summer 1998 when the assimilation of AMSU-A replaced that of SSU.

Challenge 4: Homogenisation of data with different spatial coverage

Even if the observations were accurate, the sampling of the instruments across space changes over time and even at one given time, particularly in the satellite era, there is an overwhelming variety of observation types. Considerable resources have been invested in obtaining observations of the different components of the climate system using both satellite and ground-based networks, with plans to further improve and expand these observations in coming years.

Figure 5 shows an example of how different the data coverage offered by different instruments can be within a typical 12-hour window in modern times. A climate analysis then plays an essential role by combining these diverse observations together to enable improved descriptions

variations and ch



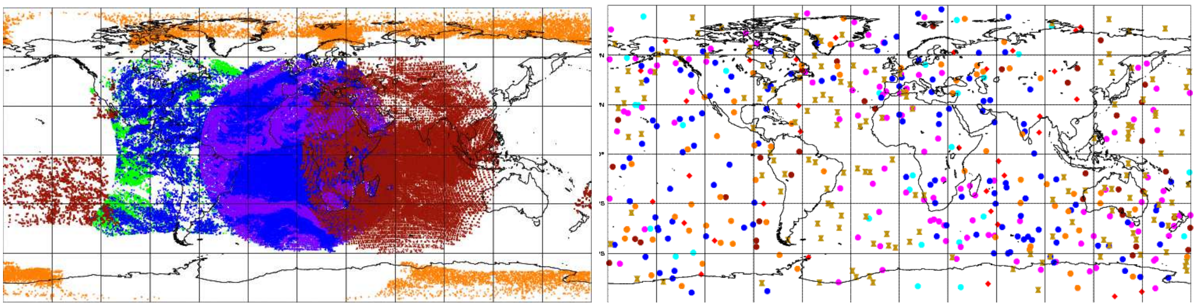


Figure 5: Examples of data coverage from a variety of in-situ (panels a) to d)) and remotely sensed data (panels e) to l)) that are available within a typical 12-hour assimilation window during the satellite era.

What is evident from figure 5 is the presence of data dense areas, where observations from multiple sources can potentially be in conflict, as opposed to areas with sparse observations, where the model is required to fill the gap and transfer information from better constrained regions. Not all variables are equally observed, e.g. at all times and places. Some of them (e.g. temperature) are normally better constrained by observations than others. In the latter case, the quality of the reanalyses may depend on how well the physical processes affecting these poorly observed variables are represented or parameterized and how efficiently the model can transfer information in space and time.

Challenge 5: Differences in the observational method

An additional difficulty associated with the diverse GOS - particularly during the satellite era - is that different sources of information while using different observational methods can provide observations affecting the same model variable. Blending together the information from these different sources is not always straightforward. An example is discussed by Dragani and McNally (2013) who described the steps that were taken to successfully merge the ozone information provided by ozone-sensitive infrared radiances from three infrared sounders (AIRS, IASI, and HIRS) with that provided by ozone products retrieved from a number of UV sensors.

Challenge 6: Interaction between different variables

An important aspect of complex systems, such as a reanalysis, is that not only are different variables related through the equations that describe / parameterise the physical processes represented in the forecast model, they are also related through the data assimilation system. It is not unusual then that the assimilation of observations that are meant to constrain one variable could also generate increments in another one. This increased complexity in the system requires a proportional increase in assumptions and choices to be made resulting in additional degrees of freedom in the modelling system. Problems can arise when (and at locations where) these additional degrees of freedom in the model are not properly constrained by the observations. An example was discussed by CMUG (2013) in their section 3. That example referred to (unrealistically) large increments in the upper-level temperature generated in an early version of ERA-Interim by the data assimilation system to accommodate observed local changes in ozone concentration caused by the assimilation of some (inaccurate) ozone products.

Challenge 7: Generation of not well observed variables

It is worth mentioning that the use of a model also enables estimates of quantities and physical processes that are difficult to observe directly, such as vertical motions, surface heat fluxes, latent heating, and many of the other physical processes that determine how the climate system evolves over time. In general, the estimated quantities are model dependent and careful interpretation is required. Any incorrect representation of physical processes (called parameterizations) will be reflected in the reanalysis to some extent. Only recently models have improved enough to be used with some confidence in individual physical processes. Previously, most studies using assimilated data have indirectly estimated physical processes by computing them as a residual of a budget that involves only variables that are well observed. Thus, it is important to understand which quantities are strongly constrained by the observations, and which are indirectly constrained and depend on model parameterizations.

The key issues and challenges discussed above can only be addressed by periodically updating the reanalysis production and the data reprocessing. There are several reasons for these updates: (1) to include important or additional observations missed in previous productions; (2) to correct observational data errors identified through subsequent quality-control efforts; (3) to assimilate observations reprocessed with state-of-the-art algorithms; and (4) to take advantage of scientific advances in models and data assimilation techniques, including bias correction schemes (Dee, 2005).

3.3 Tools for monitoring reanalysis production and data quality

As the quality of the reanalysis products has improved with time, their popularity has also grown. This is confirmed by the increasing number of users and applications where these datasets are routinely used and testified by the number of citations to the numerous papers that have been produced. For example the Uppala et al. (2005) paper describing the ERA-40 reanalysis became in 2008 the most cited paper in geosciences. It is fair to say, however, that the quality of reanalysis products within a given production varies with location, altitude, time period, and variable of interest. It is paramount then to monitor and assess the quality of every component of a reanalysis production. There are several tools that can be used to that end.

It is important to recognise that the products of a reanalysis are not merely the reanalyses themselves. The data product archive normally includes comprehensive information about how the reanalysis was produced, e.g. it includes all observations that were assimilated, together with any additional information about their quality, for example bias corrections, quality flags, and observation uncertainties. It also contains additional variable computed during the assimilation such as the observation minus background and observation minus analysis departures, and the analysis increments. These extra fields can also provide insight on the quality of the reanalyses themselves. It is now recognised that in order to have fully traceable products - important to reassure users about the data (climate) quality - one should be able to access this information at any time. For example, within the ERA-CLIM reanalysis project, ECMWF is developing a fully supported **Observation Feedback Archive** (OFA) specifically designed to give users quick and open access to all input data that were assimilated in the reanalysis.

All the products available in the database are routinely monitored to ensure that any problem can timely be detected. A number of tools have been developed over the years to that end. These tools manipulate large data volume and calculate statistics that can be plotted for easy assessment. As part of the OFA, various new tools for processing vast numbers of observations from various instruments and sources will also be made available to users.

One of the most useful tools is the **Observation Monitoring Facility (OMF)**. This OMF routinely provides statistics of the observations and of their residuals from their model equivalent (the so-called background and analysis departures), but also the observation uncertainties, the bias corrections applied and the data amount. An example, which shows some of these statistics for the SCIAMACHY total column ozone assimilated in ERA-Interim, is plotted in figure 6. The use of an OMF allows one to immediately identify issues with a given data set. Time series of the background and analysis departures represent a robust tool to identify systematic differences between observations and their modelled equivalent, useful to assess the impact of changes in the GOS and monitor the reanalysis quality over time.

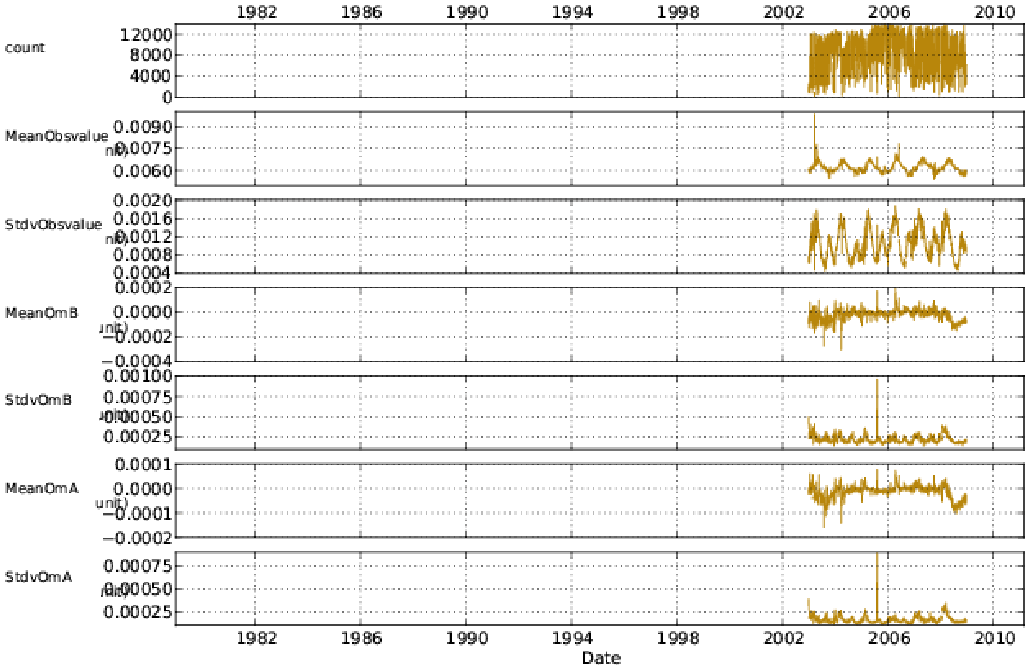


Figure 6: Statistics for the assimilated SCIAMACHY total column ozone in the ERA-Interim reanalysis. Data are in kg/m^2.

Monitoring the **Analysis Increments** (i.e. analysis minus short range forecasts) permits to identify the impact of the observations assimilated during the assimilation window (typically 12 hours) they refer to. Time-pressure cross-sections of these differences are valuable tools to assess the changes produced by the observing system. Figure 7 shows as an example the timeseries of the stratospheric temperature analysis increments over the North Pacific obtained from an early experiment of ERA-Interim. Large temperature increments of several degrees were produced in a deep layer around the stratopause (around model level 4) from 4 July 1995 onwards. An investigation of the possible causes of such large changes in the upper-air temperature showed that they were caused by the assimilation of ozone profiles from ERS-2 GOME that was turned on exactly on 4 July 1995 (CMUG 2013, section 3).

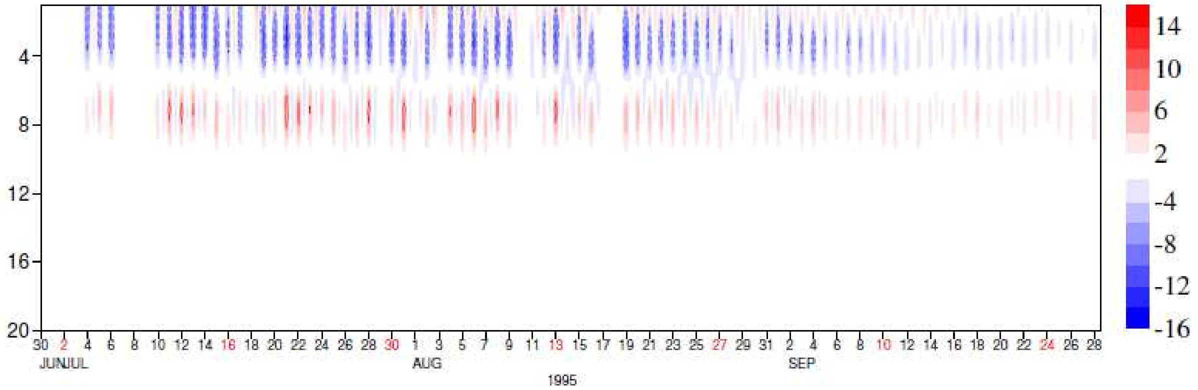


Figure 7: Mean stratospheric temperature analysis increments for an early ERA-Interim experiment averaged over the North Pacific region. Level 20 corresponds to about 40hPa; the top level is at 0.1hPa.

In addition to the daily monitoring, **Monthly Mean Products** are also routinely monitored. These include standard surface and pressure level analysed fields (temperature, winds, total column water vapour, etc…) averaged over a selected domain (global or regional), but also a number of derived variables (e.g. monthly mean time series of various global budgets, circulation indices, and single-level accumulated forecast parameters). Figure 8 shows a few examples (see captions for details) of the long term variation of some of these additional variables from ERA-Interim (red), and ERA-40 field (black). The panels c) and d) also show the comparison with the Japanese reanalyses JRA-25 (cyan) and the NCEP reanalysis NRA-2 (violet).

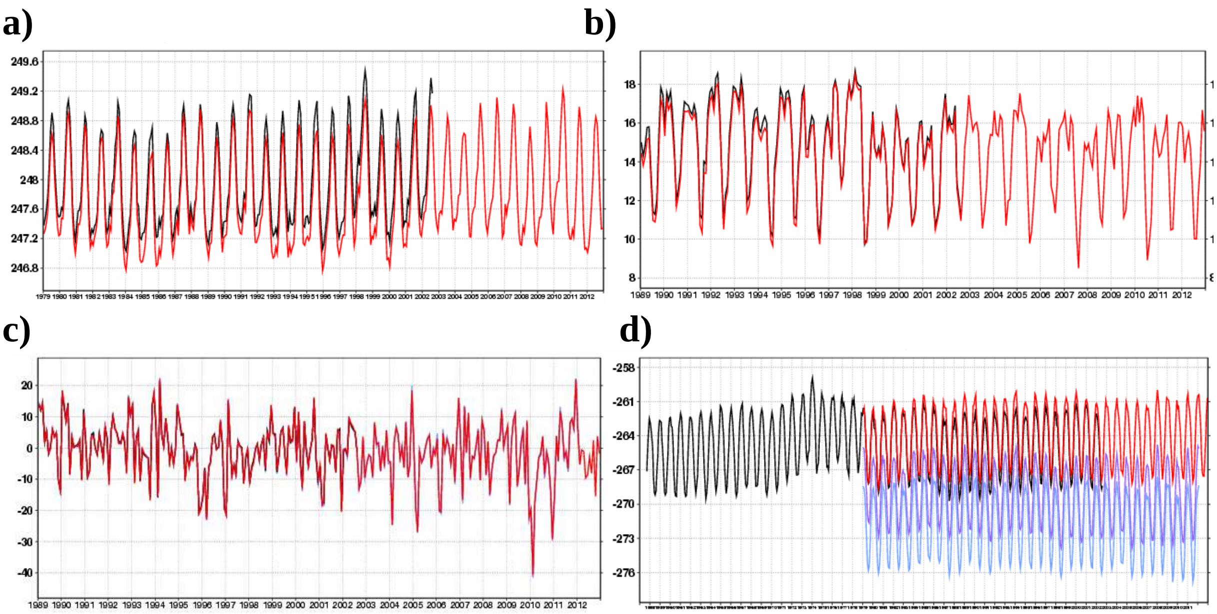


Figure 8: Comparisons of derived parameters from ERA-Interim (red) and ERA-40 (black). When available, the NCEP NRA-2(violet) and the Japanese JRA-2 (cyan) reanalyses are also shown. ***Panel a):*** *The +00h global mean middle tropospheric temperature analyses from ERA-40 (black) and ERA-Interim (red).* ***Panel b):*** *Global mean Wind angular momentum (10^{25} kg m^2/s).* ***Panel c):*** *Global mean North Atlantic Oscillation Index.* ***Panel d):*** *Global mean top-of-atmosphere thermal radiation in W/m^2. Other examples can be found on-line at [http://www.ecmwf.int/products/forecasts/d/inspect/catalog/research/eraclim/timon/.](http://www.ecmwf.int/products/forecasts/d/inspect/catalog/research/eraclim/timon/)*

Additionally, all fields that are required to fulfil **conservation laws** (e.g. global mass, energy, and imbalances) are also continuously monitored to ensure that they do not drift over time. It is important to recognize that the conservation laws are not normally verified within the analysis cycle (e.g. Berrisford et al. 2011). This is because while the conservation laws are enforced by the model, the model variables are modified during the analysis to be closer to the assimilated observations. Yet, these estimates are important because they are indirectly constrained by the observations used to initialise the model (see, for example, the diagnostics the global energy budget and the hydrological cycle discussed by Trenberth et al. 2011) and represent a valuable tool to assess the quality of the reanalysis over time, and to check the consistency among inter-related fields.

Low-frequency variability and long-term trends are also evaluated as additional quality control for some of the main fields. It is fair to say that the ability of reanalyses to accurately detect long-term trends is still controversial (e.g. see Thorne and Vose 2010, and comments by Dee et al 2011a). The factors constraining the quality of the reanalyses for trend detection are several, as discussed in the previous section (e.g. changes in observing systems over time; deficiencies in observational data quality and spatial coverage; model limitations in representing interactions across interfaces, etc…). However, considerable progress has been achieved in this area in recent years, mainly due to advances in data assimilation related to the treatment of biases in satellite observations (Dee and Uppala 2009). It has been demonstrated that near-surface temperature and humidity anomalies estimated from reanalysis data closely match those obtained independently from station observations (Simmons et al. 2004, 2010), and reanalysis data have begun to be routinely used to assess global climate change, e.g. in the annual State of the Climate special issues of the Bulletin of the American Meteorological Society (SOC 2010; 2011; 2012).

4. Part 2

4.1 Reanalysis as a framework for CDR quality assessment

Part 1 of this document presented a detailed overview of 1) the available reanalyses and the consolidated plans in this field for the near future, 2) the key issues and challenges related to climate quality reanalyses and 3) the tools that can routinely be used to monitor a reanalysis production and its quality.

The quality assurance represents a constant preoccupation in the generation of any data record. This is particularly important if these products are used for climate studies - as any inaccuracy could lead to the detection of wrong trends and long term variability - or when they become a reference to assess the value of other data records, e.g. the CCI ECVs.

There are different ways in which an observational data record can be compared with its reanalysis equivalent. These have been summarised in Table 2. Comparisons can be performed using any level of data from level 1b (radiances) to level 3 (monthly mean data sets) and it can be performed within the data assimilation system (labelled as “on-line” in table 2) or *a posteriori* (i.e. “off-line”). For completeness, the cases in which the observations (either as L1b or L2) are assimilated (i.e. their status is defined “active”) have also been included (cases B and D in red). In these two cases, however, the comparisons cannot provide an independent validation of the observations.

Cases A and C refer to situations where the observations flow through the data assimilation system but without being assimilated and thus without making any impact on the final reanalysis products, which are therefore a completely independent dataset. Although the observations do not make an impact on the reanalyses, this type of comparison is particularly useful as the data assimilation system computes and stores in the reanalysis database a number of additional fields based on the original data (e.g. short range forecast and analysis departures from the observations, information on the quality of the data, including the bias correction the observations would have been given if they were assimilated, etc…). These additional fields can be used a posteriori to calculate statistics that can provide an in-depth understanding of the quality of the original dataset in different areas of the globe and at different levels.

Case	Observation level	Reanalysis level	Mode	Status	Type of comparison
A	L1b	L1	On-line	Passive	Independent
B	**L1b**	**L1**	**On-line**	**Active**	**Dependent**
C	L2	L2	On-line	Passive	Independent
D	**L2**	**L2**	**On-line**	**Active**	**Dependent**
E	L2	L2	Off-line	N/A	Independent
F	L3	L3	Off-line	N/A	Independent

Table 2: Summary of the possible ways observations and reanalyses can be compared. The cases with shaded background refer to a posteriori comparisons. Cases A to D refer to comparisons performed within the data assimilation system.

Cases E and F refer to situations where the comparison is performed outside the data assimilation system (i.e. off-line). This could happen, for example, if the observations were not yet available at the time the reanalysis production was run. In the off-line cases, comparisons can be performed for the Level 2 data (i.e. retrievals along the satellite orbits), or for Level 3 (i.e. monthly mean area averaged fields). The first of these two cases (case E) is most useful when assessing the high frequency variability or rapid changes in a given field. The second case (case F) is indicated and suitable for the characterization of the long-term variability, for which the low frequency signal is more important, e.g. for climate assessments.

In the framework of a long term assessment of the CCI ECVs, ECMWF has developed the prototype of an interactive interface to perform comparisons of CDRs with several reanalysis products. The design of this tool, named the Climate Monitoring Facility or simply CMF, was described in detail by CMUG (2013a). Section 4.2 provides an overview of the kind of assessment that the CMF permits to perform.

4.2 Illustration of ECV product assessments

Funded through the ESA CCI, ECMWF has developed the prototype of an interactive interface for assessing low-frequency (multi-year) variability of statistical averages (typically monthly/regional means). Thus, the focus of this tool (CMF) is to perform time-series analysis in order to evaluate long-term homogeneity and consistency of CDRs.

As explained in CMUG (2013a), the CMF tool is made up of three main components as follows:

- **Dataset Pre-processing & Ingestion Interfaces** used to create the Climate Database starting from monthly mean fields (observations and model outputs, such as reanalyses).
- **Climate Monitoring Database** that holds a wide range of data (in the form of area averaged monthly mean fields)
- **Post-processing & Extraction Interfaces** used to extract, manipulate and plot the data time series.

The Database currently includes a substantial data volume that counts several reanalysis streams (e.g. ERA-40, ERA-Interim, NRA-2, JRA-2), a total of about eighty different variables (e.g. Temperature, Ozone, total column water vapour, SST, etc…) averaged over thirty eight different regions (e.g. global, tropics, midlatitudes, Africa, Antarctica, etc…). For three dimensional fields, a set of seventeen different pressure levels spanning the atmosphere from surface up to 1 hPa are available. Besides displaying area averaged monthly mean fields, a number of additional statistics (e.g. anomaly, standard deviation, RMS, etc…) can also be selected and displayed. Appendix A provides a number of tables of all the currently available options for each of these selection criteria and those being processed at the time of writing.

The CMF is an efficient tool for immediate visualization and data quality assessment. As an illustration, a number of examples are shown below using reanalysis products.

1. A tool to assess the long term consistency among datasets

The CMF permits the assessment of the long term quality and consistency of a CDR by comparing it with validating datasets available in the Climate Monitoring Database. Figure 9 shows for example the SST anomaly time series obtained from three reanalysis streams (ERA-40, ERA-Interim, and JRA-25). It is evident that for example the Japanese JRA-25 SST was rather different from the two ERA products between 1985 and 1990, while the level of agreement, especially with ERA-Interim, is much higher in other periods, particularly after 1990.

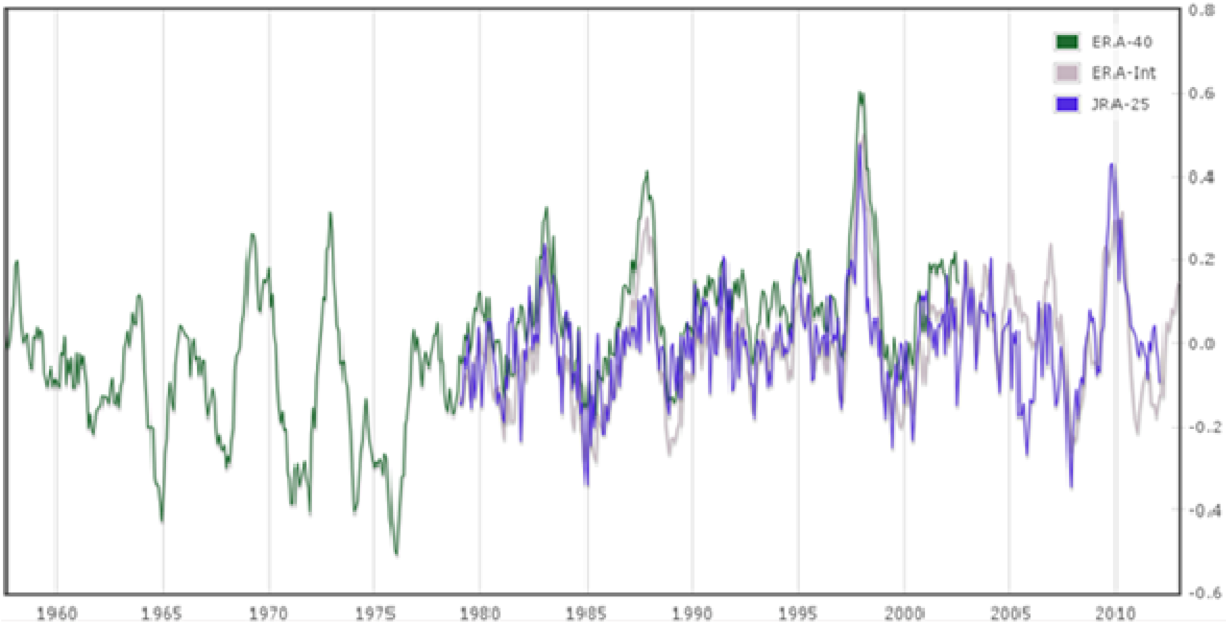


Figure 9: Timeseries of the SST anomaly from three reanalysis streams: ERA-40 (green), ERA-Interim (pink), and JRA-25 (purple).

Investigating the reasons of sudden changes in the time series is crucial to correctly detect trends and long term variability. In particular, it is crucial to understand if a sudden change is related to a problem with a particular dataset or rather it represents a real change in the environment. Figure 10 shows the global mean time series of the temperature anomaly at 100 hPa from different streams (ERA-Interim and four ensemble members of the forthcoming ERA-20C reanalysis). All the time series show sudden changes, e.g. the three most recent occurred around 1963, then 1982, and 1993. The agreement among the datasets gives confidence that these changes were not artefacts in the CDRs, but more likely real events. It is possible that they represent temperature changes occurred near the tropopause after some major volcanic eruptions, namely the eruptions of Mount Agung (1963), of Mount El Chichon (1982), and of Mount Pinatubo (1993).

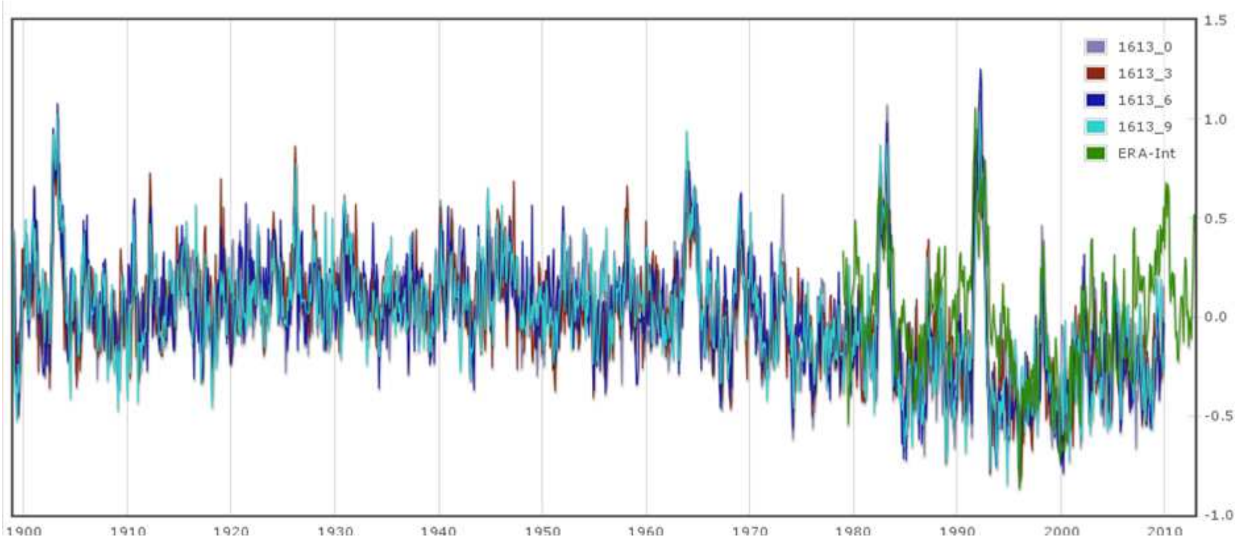


Figure 10: Timeseries of the global mean temperature anomaly at 100 hPa from ERA-Interim (green) and four out of ten ensemble members of the new ERA-20C reanalysis (see key).

2. A tool to assess the long term consistency among variables

An important feature of the CMF is the ability of comparing physically correlated variables. This helps assessing the reliability of changes in the time series of a given variable and identifying the possible causes. Using the example in figure 10 above, over-plotting e.g. aerosol timeseries (provided their availability over such long period of time) to the temperature anomaly could confirm that the largest changes observed in the temperature anomaly near the tropopause are actually the response to volcanic eruptions.

3. A tool to assess the correctness of long-term trends

One of the most critical aspects in climate assessments is the detection and validation of long term variability and trends in CDRs. Any error in the production of a CDR can result in artificial changes that in turn can provide wrong trend estimates. In the case of reanalyses it was discussed in Part 1 that changes in the global observing system represents the first cause for erroneous trends. Figure 11 shows once again the global mean temperature anomaly at 100hPa as presented in figure 10 but extended to account for other reanalysis streams. It is clear from figure 11 that the NRA-2 temperature anomaly (red line) is substantially different from all the other datasets. At a first, naïve look of the NRA-2 time series, it could appear that the suggested trend is a still-in-act, strong global mean temperature reduction around the tropopause during the last thirty years (as suggested by the fit labelled as A)). At a closer look, it is more likely that changes in the observing system may have caused a shift of the global mean temperature after 1993 to a new regime value (during the last decade), as suggested by pattern labelled as B). Paltridge et al. (2009), for example, casted doubt on the general consensus that the global water vapour feedback was strongly positive based on NCEP/NCAR NRA-2 trend analysis. Dessler and Davis (2010) analysed several reanalysis datasets and found that the NRA-2 reanalysis was the only one affected by such a negative trend, ascribing that negative trend to changes in the NRA-2 observing system.

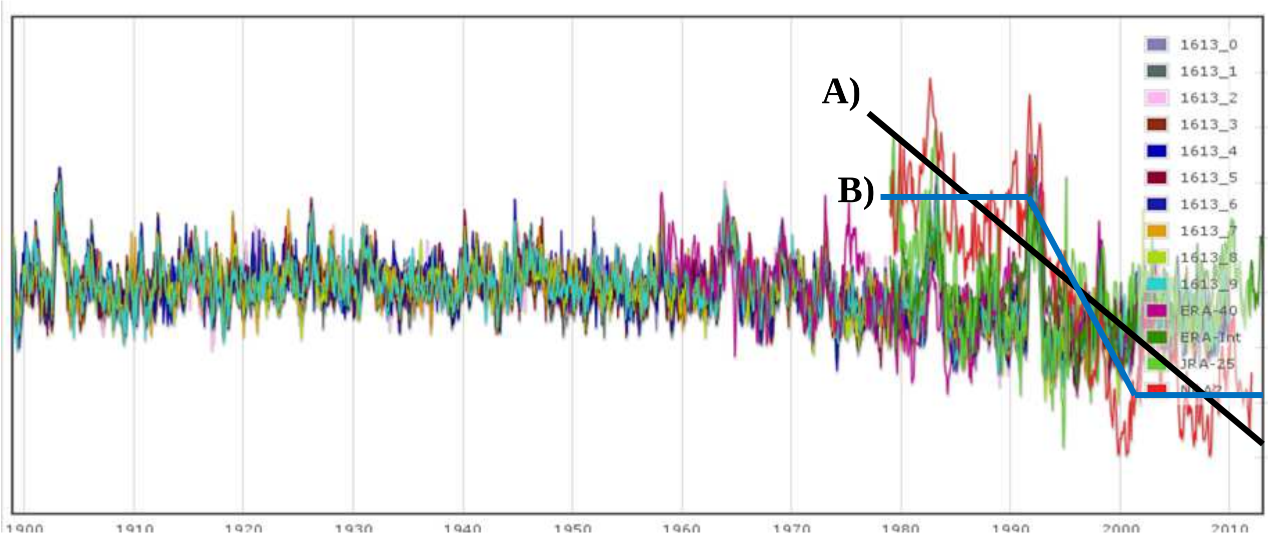


Figure 11: Timeseries of the global mean temperature anomaly at 100 hPa from the ten ensemble members of ERA-20C, ERA-40, ERA-Interim, JRA-25, and NRA-2(see legend).

This type of issues is not just common of reanalysis products. Observational datasets can also be affected by a number of problems that could lead to wrong conclusions about their trends and variability. It was shown in Part 1, that even when observations are measured with instruments using the same manufacture and then processed with the same algorithms, inter-instrumental biases can exist and be severe. These biases should be accurately removed before generating a data record long enough to permit trend analysis. Furthermore, even when limiting the attention to a single instrument, changes in the instrument measuring mode - that could affect its calibration - or errors related to the normal “wear and tear” of the instrument - that can only be modelled once identified - can all lead to erroneous conclusions. A tool, like the CMF, able to ingest many data streams and visualize statistics of all of them at once represents a very efficient way to provide a first assessment of the long term variability of CDRs while detecting potential issues and inconsistencies.

4. A tool to assess the observation uncertainties

CMF can also be used to assess the quality and reliability of the observation uncertainty. The anomaly (i.e. the variability around its mean) of the model equivalent of an observation can be used as a proxy of the natural variability of the parameter under assessment and so of the observation uncertainty. Another method consists in comparing the observation uncertainty with the spread of an ensemble of reanalysis realisations, each produced with slightly different, but equally plausible conditions, as currently done for the ERA-20C reanalysis.

5. Summary

A reanalysis is a statistical method for constructing high-quality climate records that represent our best estimate of how the climate system has evolved over time. This is achieved by combining a diverse set of past observations together within a model. The output provides comprehensive, consistent, and reliable long-term sets of numerous variables (e.g. temperature, precipitation, winds, etc…) able to characterize the state of the climate system. Thanks to these qualities, reanalysis products can be used to either produce or assess the quality of Climate Data Records (CDRs).

Reanalysis can contribute to the production of CDRs in two different ways:

1. In the production of climate quality Level 2 observations: reanalysis products can be used as auxiliary information / *a priori* to constrain the retrieval algorithms.
2. As CDRs themselves.

When a reanalysis stream and a set of observations are independent (i.e. neither the reanalysis were used in the observation retrieval algorithms nor the observations were assimilated in the reanalysis), their comparisons can infer useful information about the quality of the observations using their reanalysis equivalent as a reference, assuming that the reanalysis is accurate enough.

Whether reanalysis products are used to produce CDRs or to assess their quality, the first and foremost requirement is that the reanalysis production is done in such a way to guarantee climate quality of its products. Here, a climate quality dataset is regarded as a record of data of sufficient length, consistency, and continuity to permit climate variability and change assessments, and with accurate information about its uncertainty.

Part 1 of the present document has then focussed a) on overviewing the past, and present reanalysis productions, as well as the near future plans in the field, b) on detailing the issues that can limit the reanalysis quality - these are often related to changes in time of the observing system -, and c) on discussing how the reanalysis quality is normally monitored and what tools are available and can be used to that end.

Part 2 of the present document focussed instead on the value of reanalysis in assessing the quality of observations. It was stressed that this model-observation confrontation can be performed either within the data assimilation system (possible for Level 1b and Level 2 data records) or outside it (this, instead, applies to Level 2 and Level 3, i.e. monthly mean gridded, data records).

The model-observation confrontation up to Level 2 is most useful for detecting fast changing situations over very short period of time. The model-observation confrontation based on monthly mean gridded fields is, instead, most useful to detect climate signals and to assess the long-term variability and trends. In the context of this long-term variability and quality assessment and in support of the ESA CCI activity, ECMWF has developed the prototype of a Climate Monitoring Facility interface that facilitates the inter-comparisons of long time series of data from numerous sources. The aim is to use this facility to assess the long-term consistency and homogeneity of the ESA CCI products once available.

6. Appendix A: The Climate Monitoring Database content

The Database content evolves and increase with time. The tables below refer to what available at the time of the writing.

- **Data product streams:**

Name	Short name	Period
ECMWF ERA-Interim	ERA-Int	Jan 1979 - onwards
ECMWF ERA-40	ERA-40	Jan 1957 - Aug 2002
JMA JRA-25	JRA-25	Jan 1979 - Dec 2011
NCEP NRA-2	NRA2	Jan 1979 - Dec 2011
ECMWF ERA-20C (10 members)	1613_0-1613_9	Jan 1899 - Dec 2011
ECMWF CERA (prototype, 2 members)[1]	1644_0-1644_1, 1667_0-1667_1	Jan 1899 - Dec 2009
Hadley Centre SST[2]	HadISST2	Jan 1899 - Dec 2007

[1]Test of the coupled atmosphere-ocean reanalysis. [2]Preliminary version of the Hadley Centre SST dataset.

- **Region**

Available regions			
20N-60N	Antarctica	Europe	Oceans
20S-20N	Arctic	Global	Southern Hemisphere
30S-90S	Asia	India	South America
60S-20S	Australia	Indonesia	Southern Europe
30N-90N	Britain	Land	Scandinavia
60N-90N	Central Europe	Northern Hemisphere	Siberia
90S-60S	China	North America	Southern Oceans
Africa	Congo	North Atlantic	Tropical Oceans
Amazon	Euro-Russia	North Pacific	U.S.A.

- **Levels**

Available levels			
1 hPa	20 hPa	150 hPa	925 hPa
3 hPa	30 hPa	250 hPa	1000 hPa
5 hPa	50 hPa	300 hPa	Undefined*
7 hPa	70 hPa	500 hPa	
10 hPa	100 hPa	850 hPa	

**This is valid for surface or integrated fields.*

- **Geophysical parameter:**

Name	Description	Name	Description
2D	2m dewpoint temperature	SSCF	Surface solar radiative cloud forcing
2T	2m temperature	SSHF	Surface Sensible Heat Flux
AM	Angular momentum	SSR	Surface Net Solar Radiation
ASRU	Solar radiation reflected by the atmosphere	SSRC	Surface Net Solar Radiation, Clear sky
BLD	Boundary Layer Dissipation	SSRD	Surface Solar Radiation Downwards
CI	Sea-Ice cover	SSRU	Surface Solar Radiation Upwards
CP	Convective precipitation	SST	Sea Surface Temperature
E	Evaporation	STR	Surface Thermal Radiation
EWSS	East-West surface turbulent stress	STRC	Surface Net Thermal Radiation, clear
GWD	Gravity wave dissipation	STRD	Surface Thermal Radiation Downwards
HCC	High cloud cover	STRU	Surface Thermal Radiation Upwards
LCC	Low cloud cover	T	Temperature
LNSP	Logarithm of surface pressure	TB	BL stress torque
LSP	Large-scale precipitation	TCC	Total Cloud Cover
MAGSS	Magnitude of surface turbulent stress	TCDA	Total column Dry Air
MAM	Mass angular momentum	TCO3	Total column ozone
MASS	Mass	TCW	Total column water
MASSC	Mass convergence	TCWV	Total column water vapour
MASSD	Dry mass	TCWVC	TCWV Convergence
MASSDC	Dry mass convergence	TE	Total Energy
MASSP	Mass production	TEC	Total Energy Convergence
MCC	Medium cloud cover	TEI	Total Energy Input
NAO	North-Atlantic Oscillation	TG	GW stress torque
NI34	Nino 3-4 SST Index	TH	Thermal energy
NetS	Net surface energy exchange	THC	Thermal energy convergence
O3	Ozone	TLCF	Top-of-atmosphere thermal radiative cloud forcing
OLR	Top-of-atmosphere thermal radiation (net)	TNCF	Top-of-atmosphere radiative net cloud forcing
P-E	Precipitation minus Evaporation	TP	Total precipitation
PNA	Pacific-North America Oscillation Index	TSCF	Top-of-atmosphere solar radiative cloud forcing
Q	Specific humidity	TSR	Total Net Solar Radiation
QBO	Quasi-Biennial Oscillations	TSRC	Total Net Solar Radiation, Clear sky
RO	Runoff	TSRD	Total Net Solar Radiation Downwards
SD	Snow Depth	TSRU	Total Net Solar Radiation Upwards
SF	Snow Fall	TTRC	Top net thermal radiation clear sky
SKT	Skin Temperature	U	Zonal wind
SLCF	Surface thermal radiative cloud forcing	V	Meridional wind
SLHF	Surface Latent Heat Flux	WAM	Wind angular momentum
SOI	Southern Oscillation Index	Z	Geopotential height
SRC	Skin reservoir content		

The highlighted variables are those currently available that will also be available as CCI ECV products.

- **Quantity**

Available statistics		
Mean datum	Standard deviation datum	RMS datum
Mean anomaly datum	Standard dev. anomaly datum	RMS anomaly datum
Mean ensemble spread datum	Stand. dev. ens. spread datum	RMS ensemble spread datum

- **Geophysical parameters and data streams being processed at the time of writing:**

Name	Data stream	Description
STL1-4	ERA-40, ERA-Interim	Soil Temperature Level 1 to 4
SWVL1-4	ERA-40, ERA-Interim	Soil Water Vapour Level 1 to 4
TCO3	MACC Reanalysis	Total Column Ozone
O3	MACC Reanalysis	Ozone mass mixing ratio
TCCO2	MACC Reanalysis	Total column carbon dioxide
TCCH4	MACC Reanalysis	Total column methane
AOD469	MACC Reanalysis	Total AOD at 469nm
AOD550	MACC Reanalysis	Total AOD at 550nm
AOD670	MACC Reanalysis	Total AOD at 670nm
AOD865	MACC Reanalysis	Total AOD at 865nm
AFM	MACC Reanalysis	Aerosol Fine Mode
ACM	MACC Reanalysis	Aerosol Coarse Mode

The highlighted variables are those that will also be available as CCI ECV products.

7. Appendix B: List of acronyms

AIRS	Advanced Infrared Sounder
CERA	Coupled ERA
CIRES	Cooperative Institute for Research in Environmental Sciences
HIRS	High-resolution Infrared Radiation Sounder
IASI	Infrared Atmospheric Sounding Interferometer
JMA	Japan Meteorological Agency
GCOS	Global Climate Observing System
MERRA	Modern Era Retrospective-analysis for Research and Applications
MSU	Microwave Sounding Unit
NCEP	National Center for Environmental Prediction
NCEP CFSR	NCEP Climate Forecast System Reanalysis
NOAA	National Oceanic and Atmospheric Administration

8. References

Auligné, T., A. P. McNally, and D. P. Dee (2007), Adaptive bias correction for satellite data in a numerical weather prediction system, Quart. J. R. Meteorol. Soc., 133, 631–642.

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, ECMWF Technical Memorandum, ERA Report Series n. 13, available from www.ecmwf.int/publications/library/do/references/list/782009.

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: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res., 114, D13205, doi:10.1029/2008JD011115, 2009.

Bengtsson L., M. Kanamitsu, P. Källberg and S. Uppala, 1982. FGGE 4-dimensional data assimilation at ECMWF, B. Am. 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. R. Meteorol. Soc.*, **137** (659), 1381–1399.

CMUG, 2013a: Status of the ECMWF Climate Monitoring Database Facility. Technical Note/Deliverable 3.5, under revision.

CMUG, 2013: Uncertainties in Climate Data Records. Technical Note/Deliverable 3.3, under revision.

Compo, G.P., J.S. Whitaker, P.D. Sardeshmukh, N. Matsui, R.J. Allan, X. Yin, B.E. Gleason, R.S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R.I. Crouthamel, A.N. Grant, P.Y. Groisman, P.D. Jones, M. Kruk, A.C. Kruger, G.J. Marshall, M. Maugeri, H.Y. Mok, Ø. Nordli, T.F. Ross, R.M. Trigo, X.L. Wang, S.D. Woodruff, and S.J. Worley, 2011: The Twentieth Century Reanalysis Project. *Quarterly J. Roy. Meteorol. Soc.*, **137**, 1-28. DOI: 10.1002/qj.776

Dee, D.P (2005), Bias and data assimilation, *Q. J. R. Meteorol. Soc*, **131**, 3323-3343.

Dee, D. P., E Källén, A J Simmons, and L. Haimberger, 2011a: Comments on “Reanalyses Suitable for Characterizing Long-Term Trends” *Bull. Amer. Meteorol. Soc.*, **91**, 66-70.

Dee, D. P. and S. 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, M.A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A.C.M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Källberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011b). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553-597.

Dee, D.P., M. Balmaseda, G. Balsamo, R. Engelen and A. Simmons (2012). Toward a consistent reanalysis of the climate system, ECMWF Memorandum 687. Available from www.ecmwf.int/publications/library/do/references/list/14.

Dessler, A. E. and S. M. Davis (2010). Trends in tropospheric humidity from reanalysis systems. *J. Geophys. Res.*, **115**, D19, 1984–2012.

Dragani, R. and A.P. McNally (2013). Operational assimilation of ozone-sensitive infrared radiances at ECMWF. *Q. J. Roy. Meteorol. Soc.*, DOI: 10.1002/qj.2106.

Ebita, A., S. Kobayashi, Y. Ota, M. Moriya, R. Kumabe, K. Onogi, Y. Harada, S. Yasui, K. Miyaoka, K. Takahashi, H. Kamahori, C. Kobayashi, H. Endo, M. Soma, Y. Oikawa and T. Ishimizu (2011), The Japanese 55-year Reanalysis “JRA-55”: An Interim Report, *SOLA*, **7**, pp.149-152, doi:10.2151/sola.2011-038.

Gibson JK, Källberg P, Uppala S, Nomura A, Hernandez A, Serrano E. (1997). ERA Description. ERA-15 Project Report Series 1, ECMWF. Available from www.ecmwf.int/publications.

Grody, N. C., K. Y. Vinnikov, M. D. Goldberg, J. T. Sullivan, and J. D. Tarpley (2004), Calibration of multi satellite observations for climatic studies: Microwave Sounding Unit (MSU), *J. Geophys. Res.*, **109**, D24104.

Hollingsworth, A., Engelen, R. J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J. W., Morcrette, J.-J., Rayner, R., Peuch, V.-H., Rouil, L., Schultz, M. G., Simmons, A. J., and The GEMS Consortium: Toward a monitoring and forecasting system for atmospheric composition: the GEMS project, B. Am. Meteorol. Soc., 89, 1147–1164, doi:10.1175/2008BAMS2355.1, 2008.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team (2013): The MACC reanalysis: an 8-yr dataset of atmospheric composition, *Atmos. Chem. Phys.*, **13**, 4073-4109

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds R, Chelliah M, Ebisuzaki W, WHiggins, Janowiak

J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, **77**, 437-471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II reanalysis (R-2). *Bull. Amer. Meteor. Soc*., **83**, 1631–1643.

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.
Reichle, R. H., R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and A. Toure (2011), Assessment and enhancement of MERRA land surface hydrology estimates, Journal of Climate, 24, 6322-6338, doi:10.1175/JCLI-D-10-05033.1.

Paltridge, G., A. Arking, and M. Pook (2009), Trends in middle‐ and upper‐level tropospheric humidity from NCEP reanalysis data, *Theor. Appl. Climatol*., **98**, 351–359.

Rienecker, M.M., M.J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M.G. Bosilovich, S.D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, et al., 2011. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624-3648.

S. Saha, S. Moorthi, H-L Pan, X Wu, J Wang, S Nadiga, P Tripp, R Kistler, J Woollen, D Behringer, H Liu, D Stokes, R Grumbine, G Gayno, J Wang, Y-T Hou, H-y Chuang, H-M H. Juang, J Sela, M Iredell, R Treadon, D Kleist, P Van Delst, D Keyser, J Derber, M Ek, J Meng, H Wei, R Yang, S Lord, H van den Dool, A Kumar, W Wang, C Long, M Chelliah, Y Xue, B Huang, J-K Schemm, W Ebisuzaki, R Lin, P Xie, M Chen, S Zhou, W Higgins, C-Z Zou, Q Liu, Y Chen, Y Han, L Cucurull, R W. Reynolds, G Rutledge, and M Goldberg (2010). The NCEP Climate Forecast System Reanalysis, *Bull. Am. Meteor. Soc*., **90**, 1015-1057.

Schubert SD, Pfaendtner J, Rood R. 1993. An assimilated data set for Earth Science applications. *Bull. Am. Meteor. Soc*., **74**, 2331-2342.
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 data sets. *J. Geophys. Res*., **115** D1, 1984–2012.

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

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*, **24**, 4907-4924.

Uppala SM, Källberg 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, Holm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. (2005). The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc*., **131**, 2961-3012.