Monitoring long data assimilation time-series: a reanalysis perspective with ERA-Interim

P. Poli, D. Dee, P. Berrisford, and S. Uppala

European Centre for Medium-range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, UK
Corresponding author email: paul.poli@ecmwf.int

Abstract

Reanalyses apply sequentially (typically four-times or twice daily) data assimilation schemes in order to derive datasets of the past weather. On any given day, reanalyses aim at producing three-dimensional fields that are spatially and physically consistent. However, as the usage of reanalyses products typically involves looking across the time dimension as well, another important consideration in reanalysis is to produce time-consistent datasets. This requirement on the time dimension is then similar to that in the spatial domain. However, currently there is no data assimilation scheme available that can deal in practice with producing in one single operation a reanalysis at a reasonable resolution (<100km) that would span more than 1-2 days. Consequently, in the absence of forward-backward exchange of information between various analysis cycles, the variations in the observing system tend to propagate quite directly in to the variations of the quality of the reanalysis products. It is then very important for the producers of reanalyses to properly visualize, monitor and detect problems in the data assimilation of a varying observing system in order to mitigate these changes as much as possible, or at least make a note of them to inform data users. We present here a novel approach aimed at setting up a data supply chain to assist in these processes. To this end we build an observation statistics database. Examples of the application of this database are shown based on long time-series (1989-2009) from ERA-Interim. We further postulate that a database based on a similar concept may help deal with other problems of everyday NWP operations and monitoring.

1. Introduction: The crucial role of data assimilation in reanalysis

Reanalyses aim at constructing datasets describing the atmosphere from observations, with the additional goal that the products be consistent temporally, spatially, and physically. To achieve this, the approach so far has been to make use of state-of-the-art weather prediction models and data assimilation schemes to propagate the information in space and in time and ensure physically consistent fields (wind/mass balance for example). In that respect, climatologies produced by averaging observations and interpolating between gaps or data-void areas can also be seen as reanalyses, with the difference that the models used to propagate the information are much cruder (interpolation), while the data assimilation scheme is reduced to a more simple kind of filter (averaging). Conversely, it can be postulated that reanalyses can yield advanced climatologies since they integrate several sources of observations in one unified framework.

Practically, the components of a reanalysis system resemble very much those of a numerical weather prediction (NWP) system. They include elements that can be grouped as follows:

- Data input: observations and forcing data (sea-surface temperatures, sea-ice…),
- Data assimilation scheme (DAS): analysis and bias correction scheme, with ad hoc components for observation handling (quality control, observation error, observation operators including radiative transfer modeling…), and
- NWP model: physical and dynamical models.

In a reanalysis, in order to mitigate as much as possible the fluctuation of the product quality over time, the last two groups of elements are held constant by picking a fixed version of the NWP model and DAS.

On any given day, the role of the DAS is to formulate an analysis that takes into account the various sources of observations that are typically unequally distributed over the globe. For example the collection of in situ measurements has been traditionally much more frequent and widespread in the Northern Hemisphere (NH) than in the Southern Hemisphere. This is primarily because of larger continental masses in the NH, from where these measurements are more easily made; but this also reflects the unequal distribution of wealth in the world. As for a NWP data assimilation scheme, a data assimilation scheme in reanalysis must also be able to come up with an optimal solution that does not suffer too much from such discrepancies in input data coverage.

Finally, the first group of elements mentioned above (time-varying input data) cannot be held completely constant. The observing system changed over time and it is now obviously impossible to go back and collect observations of the past. We will illustrate below the difficulty involved by these changes.

2. Current ECMWF Reanalysis: ERA-Interim

Global reanalyses of the atmosphere have been produced by the National Centers for Environmental Prediction (NCEP) in collaboration with the National Center for Atmospheric Research (NCAR), by the National Aeronautics and Space Administration Data Assimilation Office (NASA DAO), by the Japan Meteorological Agency (JMA), and by ECMWF.

For example, ERA-40 (Uppala et al., 2005) covered the time period extending from 1957 to 2002, building on previous reanalysis experience (ERA-15) and on data gathered by partner institutions.

The current ECMWF reanalysis ERA-Interim covers the time period from 1989 to the present. Unlike ERA-40, this reanalysis uses an adaptive bias correction (variational bias correction, applied to satellite radiances) instead of pre-computed, piece-wise fixed, bias corrections, and a twice-daily four-dimensional variational (4DVAR) assimilation instead of a four-times-daily three-dimensional variational (3DVAR) assimilation.

The main characteristics of ERA-Interim are:

- T255 (~80km) horizontal resolution.
- 60 vertical layers; top level at 0.1 hPa.
- Improved model physics (ECMWF model cycle 31r2).
- Four-dimensional variational (4DVAR) analysis using a 12-hour time window.
- Revised humidity analysis, as compared to ERA-40.
- Wavelet-based background error covariances.
- Variational bias correction of radiance data and assimilation of rain-affected microwave radiances.
- Assimilation of the latest generation of satellite data that were not available/assimilated in ERA-40, such as GPS radio occultation measurements (from 2001 onwards), hyperspectral infrared measurements from the Atmospheric InfraRed Sounder (AIRS, from 2003 onwards), ozone retrievals from MLS and OMI on EOS-Aura (from 2008 onwards), and microwave radiances from SSM/I-S and AMSR-E (from 2009 onwards).

The main improvements as compared to ERA-40 are:

- Better fit to observations.
- Much better hydrological cycle.
- Improved stratospheric transport.
- Improved forecast skill.

As of June 2009, ERA-Interim has reached real-time and now continues as a Climate Data Assimilation System. New data from recent satellites will be added whenever possible if this can be achieved with little disruption to the overall system (e.g., NOAA-19). See http://www.ecmwf.int/research/era/do/get/index for product access and latest updates.

3. The ever-changing observing system

For a given reanalysis time period, one approach could be to consider only the common dominator of all the observation instruments or sources. The 20th Century Reanalysis project (Compo *et al.* 2006), currently being conducted in the United States, follows that approach (http://www.cdc.noaa.gov/data/gridded/data.20thC_Rean.html). In their reanalysis, Compo et al. assimilate, with an ensemble filter, surface pressure observations only, in an attempt to avail themselves from an otherwise largely changing observing system. However, even only with surface pressure observations, the challenges of producing time-consistent reanalysis products remain considerable, because the surface observing system did change significantly over the century, both in quantity and in quality.

The reanalysis approach being pursued at ECMWF (e.g., ERA-40, ERA-Interim) is slightly different in that it tries to assimilate as many sources of available data as possible. In fact, this acknowledges that we already have to deal with observations of irregular quality (over the globe) and irregular spacing; the irregularity in the time dimension is simply another complication that adds to the irregularities in the spatial and physical domains.

Figure 1 shows the timeline of surface observations assimilated in the ERA-Interim. Except for the METAR data (meteorological reports collected at airports) that were only available from 2004 onwards, most of the SYNOP types of observations were available nearly continuously since 1989 (we are not yet discussing the coverage, the quality or the quantity of observations behind each bar). This picture is to be compared with the timeline of satellite radiance instruments shown in Figure 2. As such, these two figures only present a partial view of the whole observing system which also includes radiosondes, aircraft, drifting buoys, wind profilers, satellite imagers (from which atmospheric motion vectors (AMVs) are derived), satellite scatterometers (from which sea-surface winds are derived), satellite ozone instruments, and also GPS radio occultation receivers in recent years.

These separations by observation type are important but somehow artificial. It can be more revealing to separate the various observations by type of physical observable. Figure 3 shows the total counts of data assimilated by observable type. The number of radiance observations is by far larger than any other type of observables and increases over time. There are several reasons for this: first, the number of satellites increased (for example satellites originally intended for research such as NASA EOS provide data in real-time in the same fashion as the operational NOAA satellites), second, the resolution of radiometers improved (higher spectral resolution for example), and third, the reliability of the satellite instruments improved (longer lifetime -- this is most readily visible from Figure 2). These radiances contain primarily information on the temperature (i.e., mass) fields, while a small number of channels are also sensitive to water vapour information.

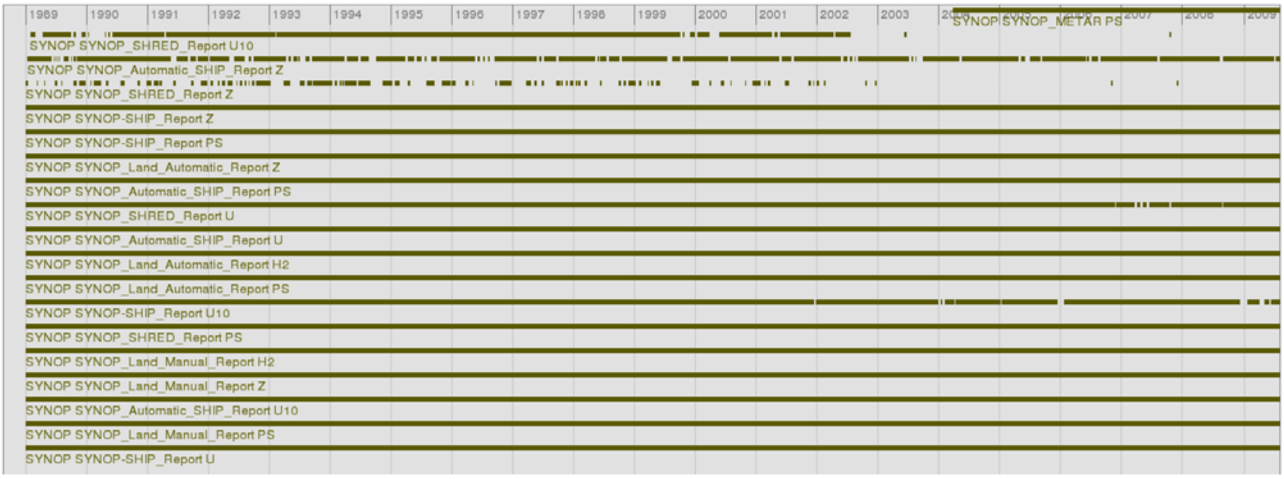

Figure 1: Timeline of the SYNOP observations assimilated in ERA-Interim. Each bar represents a different type of observation.

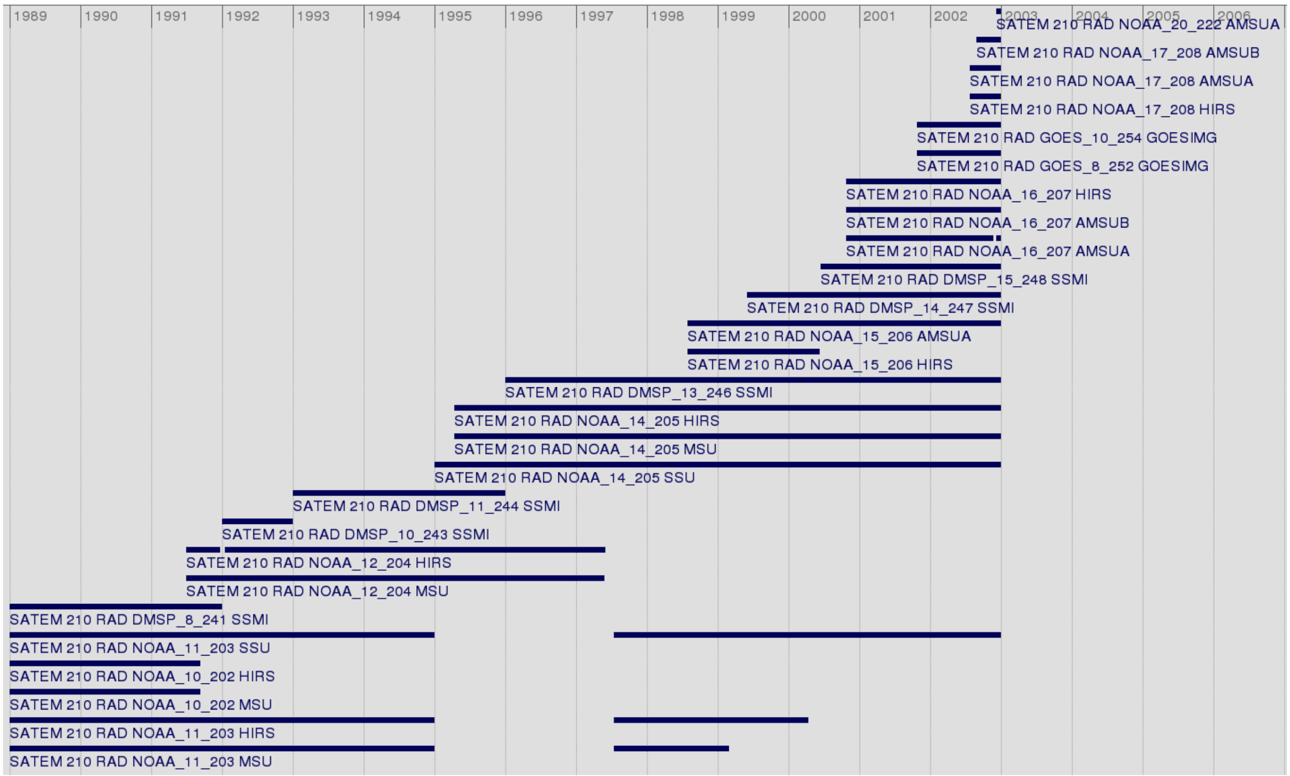

Figure 2: Timeline of the satellite radiances assimilated in ERA-Interim. Each bar represents a different combination of satellite and instrument. Note that NOAA-20 refers in fact to NASA EOS Aqua. The list of instruments includes HIRS, SSU, MSU, AMSU-A, AMSU-B, MHS, GOES imagers, METEOSAT imagers, MTSAT-1R imager, SSM/I, AIRS, SSM/IS, and AMSR-E.

The number of upper-air temperature, surface pressure or geopotential height observations (i.e., observations of mass) is also found to be highly increasing over time. By contrast the number of observations of upper-air humidity only increased slowly, though the number of surface humidity data increased in equal proportion to that of surface pressure data. Overall the quantity of humidity information still remains much smaller than that of the mass fields, although it is known that humidity presents a higher variability and exhibits smaller time- and space-scales. The number of wind observations increased significantly over the period in a proportion that is similar to the increase seen in radiances, because most of the wind observations are derived from satellite imagery, while only a fraction is collected in situ. As of 2009 the number of assimilated observations exceeds 8 million per day, while the total number of observations ingested daily for screening (i.e., before quality control and thinning) exceeds 200 million.

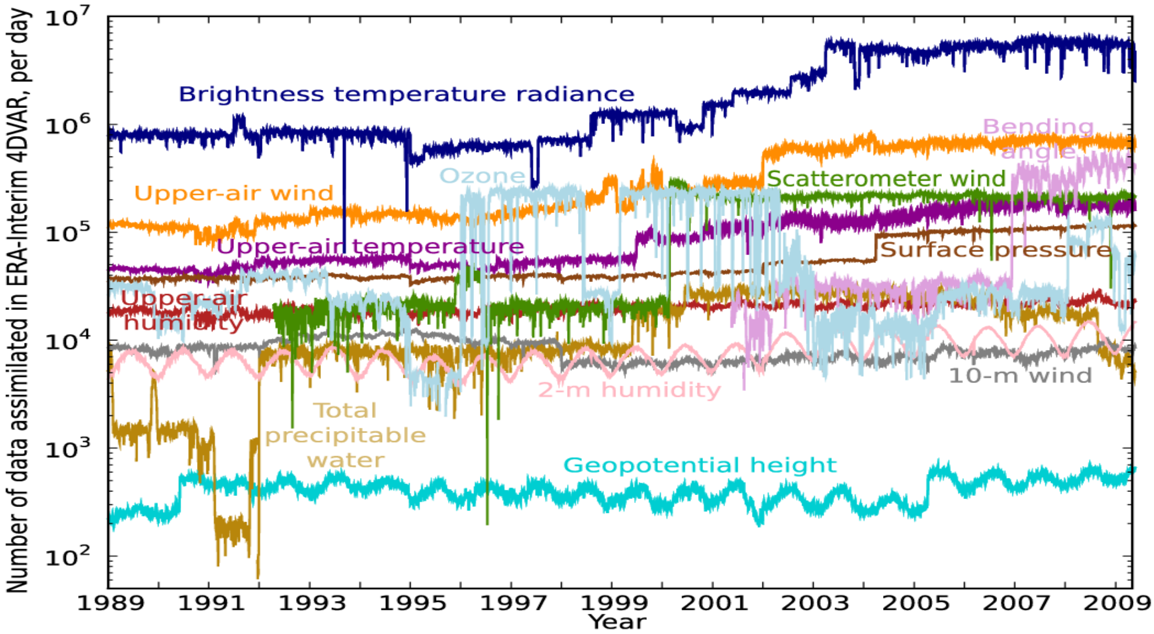

Figure 3: Number of individual observations assimilated in ERA-Interim 4DVAR, per day, grouped by observable type.

4. Towards a data supply chain for monitoring and exploring the observations in reanalysis

Generating time-series as shown in the Figure 3 is typically achieved by first gathering ad hoc data and then plotting the series. While the plotting can be easily trimmed down to rely on a common resource, the data gathering part still typically requires some doctoring to gather and organize the statistics at the level at which one ultimately wants them to be plotted. The real issue is that today no flexible data supply chain provides statistics on observations from the reanalysis to users. All current monitoring mechanisms are built around the target plots that users might want to see -- and not around the data that enter the reanalysis.

The solution that we have developed removes the so-far inevitable repeats of the earlier necessity of "statistics calculation and organization doctoring". Briefly, (1) gathering observation statistics and counts for all the observations is performed according to a set of user-defined specified "keys", and these granular statistics are stored in a database; (2) a second "key" specification, relying on the same algorithm, allows further extracting aggregate of statistics along any dimension(s), from which plots can be generated. We now describe how the overall scheme is constructed.

4.1. Methodology

The data input to our scheme is the SQL (Structured Query Language)-based Observational DataBase (ODB) used for observation handling in the ECMWF Integrated Forecast System (IFS).

First the user has to specify the "keys" or the granularity according to which observation statistics are to be computed. Then a recursive engine generates corresponding SQL queries. This engine simply identifies the various cases as defined by the user. In practice this means for example separating between the SQL queries depending on whether we need to bin by channel for satellite radiances, or we need to bin by pressure levels. Note that the engine does not need to "know" these details (or contain observation-specific code); it simply applies the defined keys and in that sense is totally generic and could have many other applications. It is really the nerve centre of our system. As discussed below, this engine has more far-reaching implications, especially when it comes to "drilling" into the database and visualizing information.

The application of the produced SQL queries to the SQL-based ODB then calculates the statistics according to the selected keys, using SQL native aggregate statistical functions. At this point we get statistics that we can organize in a tree (again, using the early definition) and further store in a second SQL-based database, the observation statistics database. This database is for now relying on POSTGRESQL but any other SQL-based database could do.

Extracting statistics is then simply a repeat of all this; first we define a list of keys according to which we wish to group the statistics, say, for satellite radiances we want to see the timeseries of mean innovations (observation minus first-guess) before and after bias correction, binned by channel, while for other data we only want the innovation, and binned by pressure if the data are on pressure level. The tool as described above then performs the calculations and aggregations, except that this time it acts on the observation statistics database instead of the observation database; the final result is a tree of statistics that we can then browse and plot.

4.2. Technical implementation

In practice, the tool relies on open-source tools and no single line of code was written to recalculate standard statistics such as minimum, or mean, etc... The idea here was really to externalize as much as possible all the low-level tasks that have been coded many times since the invention of the computer; the best versions of these algorithms are more likely to lie in the open-source community software (which constantly get improved), than in any package of our own device.

The data supply chain is written in Python™ to benefit from the many packages that have been developed by the wide community of Python™ users, estimated to exceed 1 million. Examples of such packages that are used here are:

- MetPy: a meteorological package developed at ECMWF that interfaces with ODB (and also the standard meteorological formats, BUFR and GRIB, though we do not use them here).
- Psycopg: an interface to POSTGRESQL.
- simpleJSON: an interface to load/write JavaScript Object Notation, JSON. JSON is a text-based serializable language to hold data. These data are readily interpretable by JavaScript and also viewable in Firefox©, effectively allowing rapid debugging of the data flow.
- Turbogears: a web server.
- Matplotlib: a plotting package.

- flot: a web plotting JavaScript resource that we call via an Asynchronous JAvascript eXtensible Markup Language --XML-- (AJAX) layer, or rather AJAJ since we exchange light-weight JSON data instead of unnecessarily cumbersome XML.

Putting this all together allows serving the data to a web page featuring zoom-enabled plots. All the figures hereafter are snapshots of such web pages.

Although the list above may seem a bit long, it really means that we could spend our efforts in exploring the "keys" or dimensions of the observing system statistics, instead of repeating writing ad hoc codes to read/write/calculate/display statistics.

4.3. Limitations

As indicated above, our scheme works on the SQL-base ODB. It could work equally well in any environment where the observations are arranged in a SQL-type database. However, for other types of database, the engine could probably be re-written to generate ad hoc queries, but the important point to remember here is that the scheme externalizes the statistics computation to the database engine.

5. Selected observation statistics time-series from ERA-Interim

The present section contains a series of examples of observation statistics time-series that were generated by our data supply chain from the ERA-interim ODBs.

5.1. Bias correction time-series for AMSU-A channel 10

The AMSU-A instrument is a microwave sounder that has been flying on various polar-orbiting platforms since 1998. The channel 10 of this instrument senses the atmospheric emission in the 57 GHz oxygen band and the corresponding weighting function peaks between the 30 hPa and the 100 hPa pressure levels. Because it senses the lower stratosphere, this channel is useful to monitor variations in the temperature in the layer immediately above the troposphere, and can as such be an indicator of long-term atmospheric changes (the nearby channel 9 is also often used since it offers a near-continuation of the MSU channel 4).

Figure 4 shows that the bias corrections for AMSU-A channel 10 calculated by the variational bias correction in ERA-Interim are not identical for all satellites. There are clear offsets between the various platforms. Ideally, the bias correction calculated in this manner should only reflect the contributions to the biases from the observations and from the observation operator (RTTOV-8 in this case). The first part should be uncorrelated between the various sensors while the second part would induce correlations between identical sensors flying on different satellites. However, it is likely that the variational bias correction also acts to correct (up to a certain extent) for the model (background) bias, which, again, would result in correlated bias corrections between the satellites. Given the very similar variations observed between different satellites carrying the same sensor, it seems indeed clear that the bias correction strongly contains either the observation operator bias, or the model bias (or both).

A recent event that was known to have been only poorly captured by the model is also shown in January-March 2009. A sudden stratospheric warming event took place in the Northern Hemisphere. All the brightness temperatures for AMSU-A channel 10 then became a little bit warmer than the conditions created by the model; the variational bias correction reacted by increasing the bias correction for all satellites in a similar manner. Overall the time-series shown here suggests that the variational bias correction did a proper job of capturing observation events (see for example the recalibration that occurred on METOP-A around mid-2007).

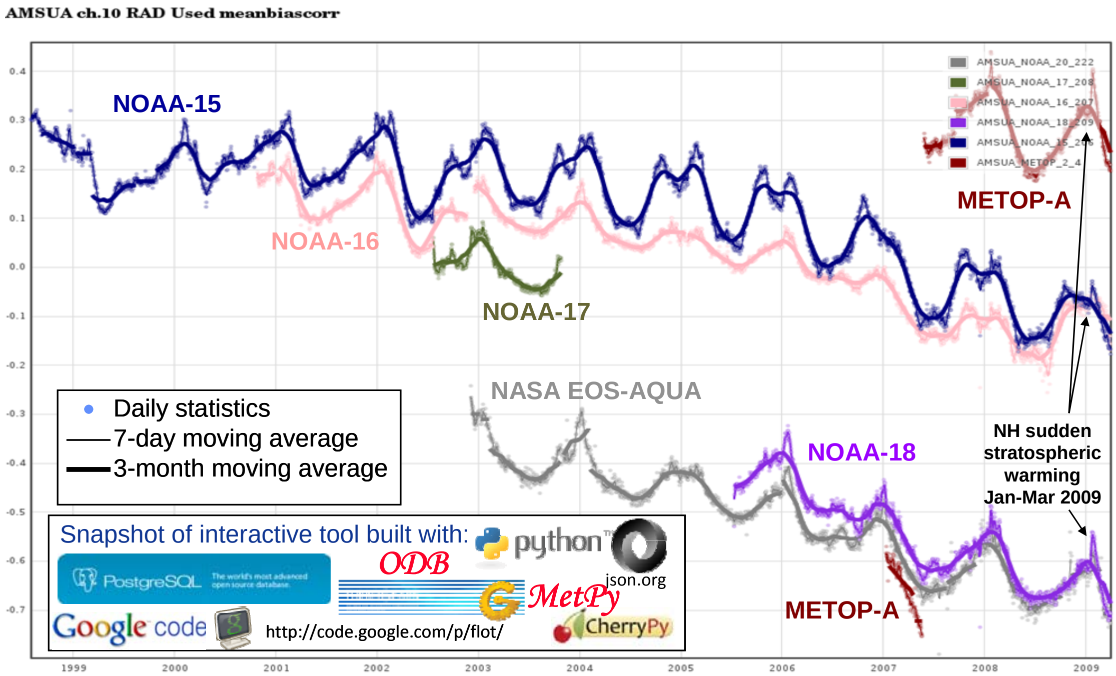

Figure 4: Time-series of AMSU-A channel 10 (peaking at 30-100 hPa) bias corrections, as calculated by ERA-Interim variational bias correction.

The slow drift apparent for all satellites after 2002-2003 seems to be the result of biases induced by more numerous, warmly biased, aircraft observations, as discussed now. Figure 5 shows the number of temperature observations from aircraft assimilated in ERA-Interim in the layer between the 225 hPa and the 175 hPa pressure levels. These temperature observations are not bias corrected and so are used as references (like radiosondes) by the variational bias correction. They increase drastically in number after 2002-2003. However, as shown in Figure 6, these observations are not exactly bias-free as compared to the background, but are typically warmer by 0.1-0.5 K. The mean of the residuals (i.e. observation minus analysis) is reduced to approach zero (not shown here). This adjustment in the analysis, coupled with the earlier increase in data counts, results in the entire system adjusting itself (including all the bias corrections) to fit the warmer aircraft observations. Obviously the aircraft temperature observations would need to be bias-corrected in a future reanalysis (and ideally one should also address the model bias issue).

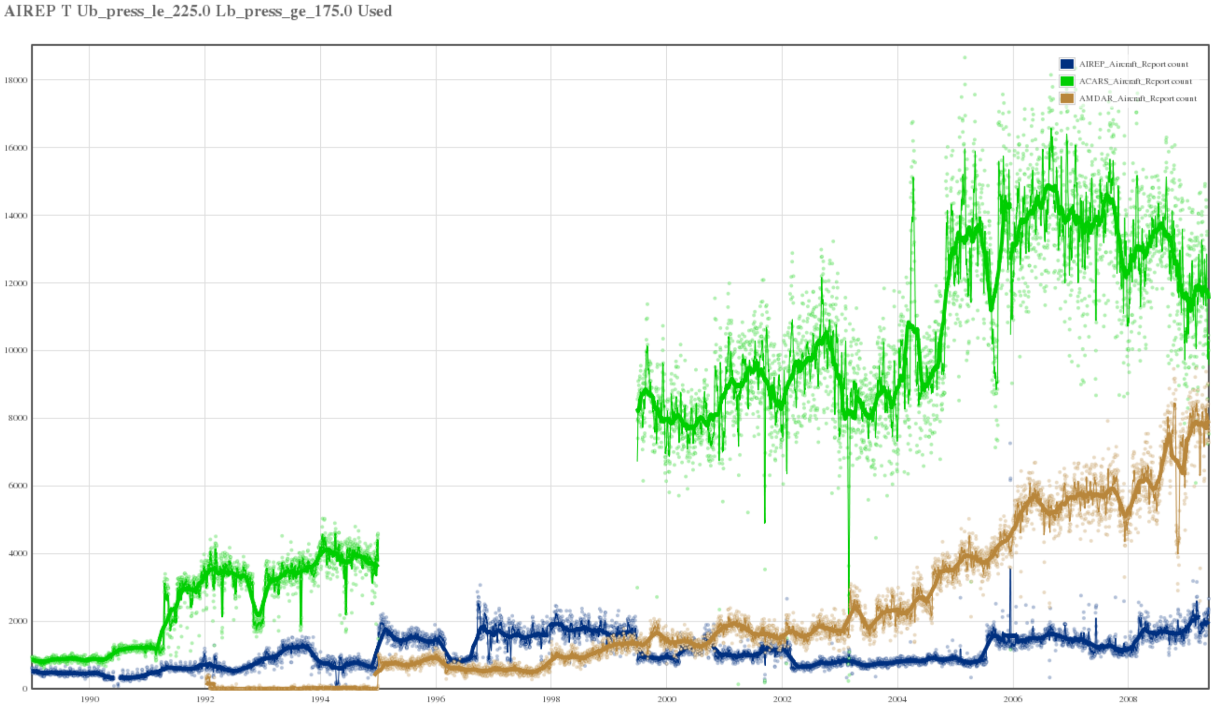

Figure 5: Time-series of the daily number of temperature observations assimilated from aircraft in the layer 175-225 hPa

Figure 6: Time-series of the observation minus first-guess departures for temperature observations from aircraft (layer 175-225 hPa).

5.2. Comparison of bias corrections for HIRS channel 7 and corresponding AIRS channels

Figure 7 shows the bias corrections computed for the infra-red HIRS channel 7, which senses emission/absorption from the CO2 15 microns band in the lower troposphere. The figure also shows the mean bias correction of the AIRS channels assimilated in ERA-Interim that fall within the spectral region of HIRS channel 7. The spin-up of the bias correction for AIRS appears to have been unusually long (this is confirmed by looking at other AIRS channels). Furthermore, once the bias correction for AIRS is established, the slow variations appear to be well correlated with those observed for HIRS. The biases in the fast radiative transfer model would have no reason to be different between HIRS and AIRS, except if the line-by-line calculations on which the fast model RTTOV is built are biased. Assuming then uncorrelated error biases between HIRS and AIRS, this suggests that the variational bias correction does correct somehow for the model biases.

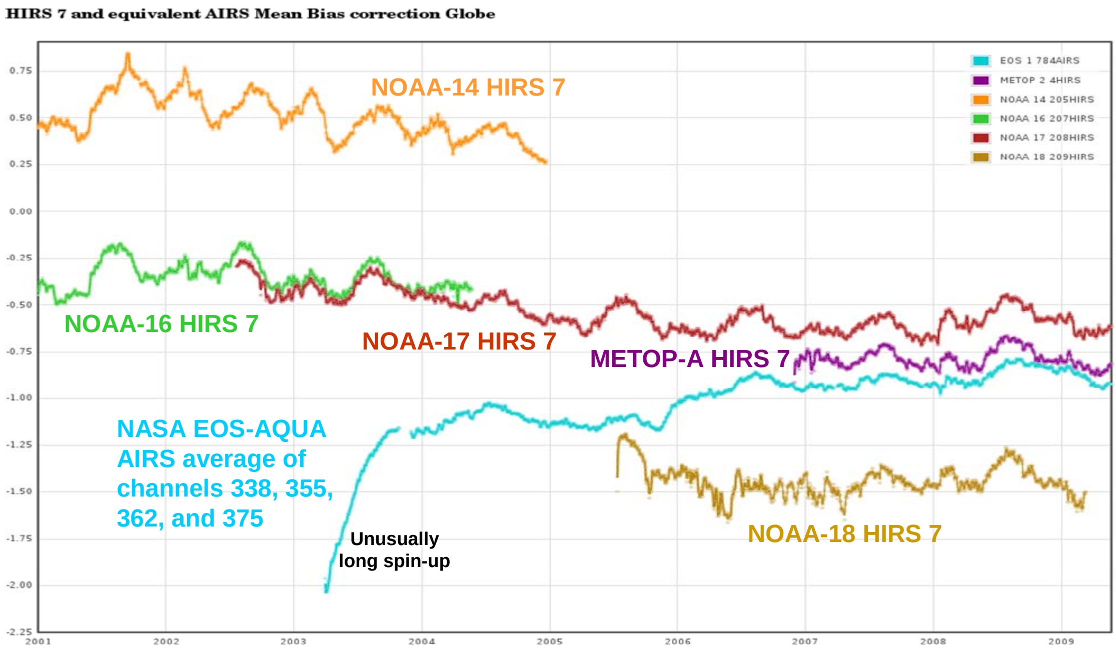

Figure 7: Time-series of bias corrections for HIRS channel 7 and mean of the AIRS channels in the same spectral region

5.3. Change in the quality of SSM/I DMSP F-13 data

Figure 8 shows the standard deviation of the first-guess departure (observation minus first-guess, also called innovation) for the precipitable water content (PWC) observations assimilated in the 4DVAR. These observations are obtained from a 1DVAR retrieval from SSM/I radiances. The figure suggests a drop in innovation standard deviation for the PWC retrievals from DMSP F-13 around mid-1999 when the data from DMSP F-14 also started being assimilated. This is a bit surprising as usually the quality of data from a given satellite degrade over time (with sensor degradation). However, mid-1999 was also a time when the source of SSM/I data changed, from the dataset acquired from RSS for ERA-40 until mid-1999 to the dataset as received by ECMWF operations afterwards. The application of Desroziers et al. (2005) diagnosis is then used to understand hereafter whether the quality of the data inherently changed at that time.

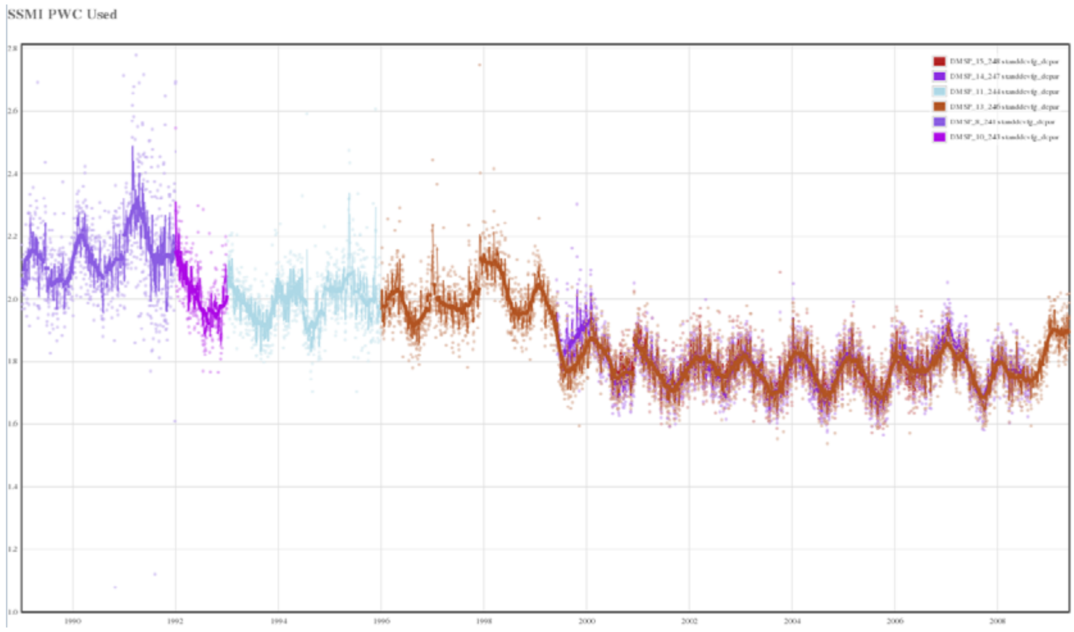

Figure 8: Standard deviation of the PWC innovations (observation minus first-guess) in ERA-Interim, grouped by satellite. Data from DMSP F-14 were assimilated between 1 June 1999 and 24 August 2008 and data from DMSP F-15 were assimilated between 13 June 2000 and 11 August 2006.

To this end, Figure 9 shows the ratio of the observation error as inferred from the formula of Desroziers et al. divided by the actual observation error assumed in the 4DVAR assimilation, for all the channels of SSM/I DMSP F-13. A ratio larger than one indicates that we usually under-estimate the observation error. A ratio smaller than one suggests that we follow a "conservative" approach (we assign observation errors that are probably larger than what they ought to be). The hypotheses in Desroziers' method are that the background and observation present very different observation error correlation patterns, so that the two errors can be disentangled in the formula. The usual hypotheses of unbiased normal errors for all information sources also apply. The diagnostic does seem to indicate that the quality of data did change sharply around June 1999, with apparently smaller random errors in the operational dataset than in the RSS dataset.

In fact, looking at Figure 10 for the SSM/I DMSP F-13 channel 3 (22GHz, the primary source of water information among the 7 SSM/I channels), we notice that the quality of the data changed in a particular way. The bias corrections were small in the RSS dataset, probably because the data had been purposely bias-corrected by the data producer. The bias corrections calculated by the variational bias correction and as a result of the 4DVAR assimilation increased suddenly when the data received in operations started being assimilated. From that point on, the standard deviation of the residuals (observation minus analysis) also decreased. The residual fit after analysis is the best indicator of data quality in the absence of other systematic verification data to compare with; the quality improvement is in line with the result shown in the previous figure. Overall this illustrates that a data assimilation system that can directly deal with features in the observations can extract more information from the raw data (by fitting them better) than using an a priori method to correct the data "outside" a data assimilation system. Of course such comments only apply now that the variational bias correction was incorporated in ERA-Interim; there are many areas of the pre-assimilation processing that are of course still better left to the data producers. We can conclude from this section that the diagnostic from Desroziers et al. gives meaningful results in the example shown here.

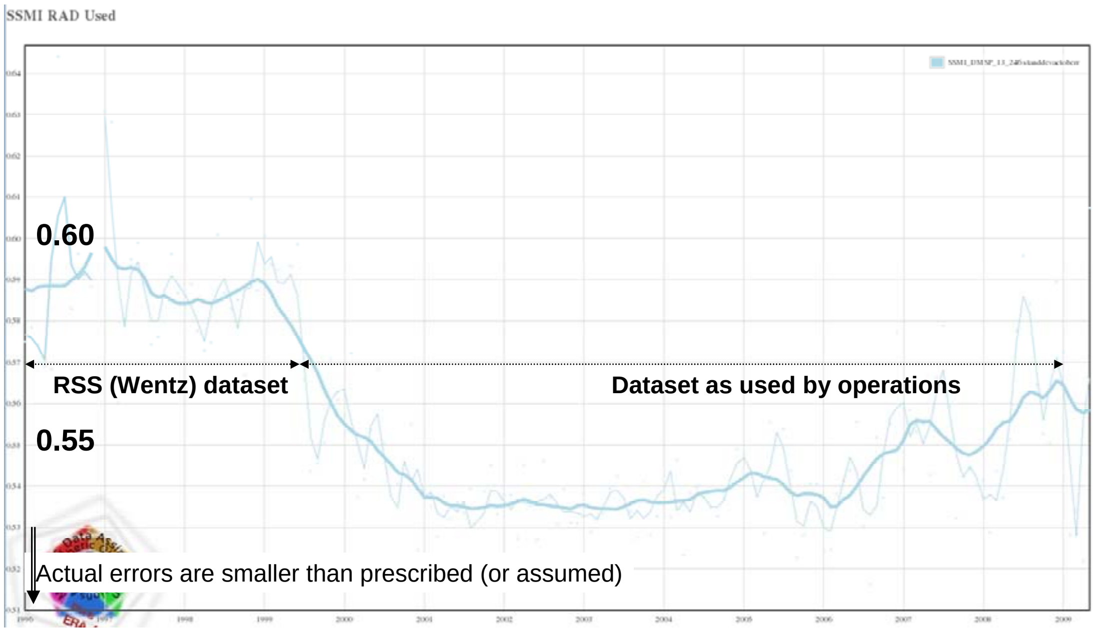

Figure 9: Ratio of the observation error calculated using the diagnostic of Desroziers et al. divided by the observation error used in the 4DVAR assimilation, for all channels of SSM/I DMSP F-13.

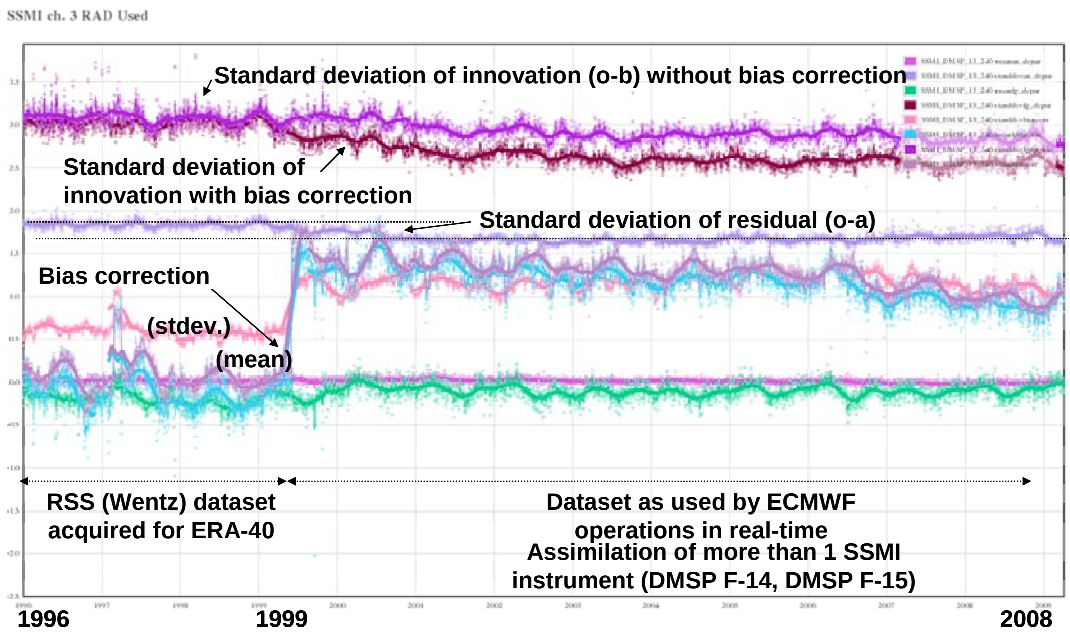

Figure 10: Statistics of SSM/I DMSP F-13 channel 3 departures (observations minus first-guess and analysis) and bias correction.

6. Conclusions and future work

The ERA-Interim reanalysis has now completed more than 20 years of assimilation using a modern data assimilation system (4DVAR with variational bias correction). The focus is now to investigate the quality of the products to appropriately notify users and to prepare for a next reanalysis over a longer time period. In order to achieve this we need to be able to drill into the observation statistics before and after assimilation (innovation and departure, with and without the bias correction calculated by the assimilation), in all the parts of the observing system.

The difficulties in setting up exploration tools for such statistics are that the observing system changed significantly over the period, and also that the observing system presents irregular dimensions (channels for radiances, pressures for radiosondes, moving station positions for drifting buoys, etc…). A first experimental data supply chain was built up from the observations, feeding and interrogating an observation statistics database. The core of the system is a tree-making algorithm that converts sets of keys into SQL queries whose results can then be organized in a tree as well.

The present paper illustrated with a series of examples that the statistics that can easily be accessed via this system contain useful information to prepare for the next reanalysis. For example we find that the future reanalysis should either contain an automatic bias correction convergence scheme in order to avoid using new data whose bias corrections are still changing, or that we need to investigate better ways to initialize the variational bias correction so that data from new sensors can be readily useable. This may be achieved for example by running a lower-resolution reanalysis 1 or 2 years ahead of the full resolution reanalysis, so as to spin up the variational bias correction coefficients.

We also showed that the diagnostic of Desroziers et al. (2005) can be applied easily to monitor the variations in observation errors. In the example investigated here (data from the SSM/I DMSP F-13 satellite) the diagnostic yields results that are consistent with other statistics.

Future work for our experimental data supply chain include increasing the granularity of the database so as to separate the statistics by station when appropriate (e.g., radiosondes, surface stations). The reason for this is the population of sensors vary greatly over time (e.g., huge increase in the number of automatic stations over land in the last decade).

Finally, the observation statistics database constructed here could be used to derive and fine-tune an automatic warning system embedded in the assimilation. As such, data assimilation in reanalysis (and also possibly in operations) could greatly benefit from such an embedded database. The monitoring and hard decisions over quality control (e.g., to activate or remove particular stations or channels) could probably be better controlled through such an integrated mechanism than through ad hoc projects that work separately and outside the data assimilation system.

References

Compo, G. P., J. S. Whitaker, and P. D. Sardeshmukh (2006), Feasibility of a 100 year reanalysis using only surface pressure data. *Bull. Amer. Met. Soc.*, **87**, 175--190, DOI: 10.1175/BAMS-87-2-175

Desroziers, G., L. Berre, B. Chapnik, and P. Poli (2005), Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Royal Meteorol. Soc.*, **131**, 3385-3396, DOI: 10.1256/qj.05.108

Uppala, S. M., P. W. Kallberg, A. J. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. A. Janssen, R. Jenne, A. P. McNally, J. F. Mahfouf, J. J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen (2005), The ERA-40 re-analysis. *Quart. J. Royal Meteorol. Soc.*, **131**, 2961--3012, DOI: 10.1256/qj.04.176.