CECMWF Feature article

from Newsletter Number 119 – Spring 2009

METEOROLOGY

Variational bias correction in ERA-Interim



This article appeared in the Meteorology section of ECMWF Newsletter No. 119 – Spring 2009, pp. 21-29.

Variational bias correction in ERA-Interim

Dick Dee, Sakari Uppala

Reanalysis has a long history at ECMWF, beginning with the first reanalysis of observations from the First GARP Global Experiment (FGGE) in 1979. Since then, two major reanalysis projects have exploited the substantial advances made in operational weather forecasting at ECMWF. The first of these, ERA-15 (1979–1993), was completed in 1995 (*ECMWF Newsletter No. 73,* page 7) and the second extended reanalysis project, ERA-40 (1957–2002), in 2002 (*ECMWF Newsletter No. 101,* page 4). Products of ERA-15 and ERA-40 have been used extensively by the ECMWF Member States and the scientific community at large (*ECMWF Newsletter No. 104,* page 5). The ERA-40 public data server currently has 12,000 registered users world-wide. Reanalysis is also increasingly important to many of the ECMWF's core activities, for example:

- · Providing the climatology needed for forecast verification.
- Serving as a reference for the validation of long-term model simulations.
- · Allowing the development of a seasonal forecasting capability.
- Establishing the climate of EPS (Ensemble Prediction System) forecasts needed for construction of forecaster-aids such as the Extreme Forecast Index.

ECMWF is now producing ERA-Interim, a global reanalysis of the data-rich period since 1989 based on cycle 31r2 (Cy31r2) of the Integrated Forecast System (IFS).

A key component of ERA-Interim is the variational bias correction system for satellite radiances. In this article we describe the performance of this system based on the first nineteen years (1989–2007) of production of ERA-Interim.

ERA-Interim

ERA-Interim represents a step towards ECMWF's next-generation reanalysis system, to be developed in the next few years if the necessary funding can be obtained. Relative to the ERA-40 system, which was based on IFS Cy23r4, ERA-Interim incorporates many improvements in the model physics as well as in analysis methodology. The configuration of the ERA-Interim system and many aspects of its performance are described in *ECMWF Newsletter No. 110* (page 25) and *No. 115* (page 12). When it reaches real-time in early 2009 ERA-Interim will be maintained as a Climate Data Assimilation System (CDAS), opening new opportunities for climate monitoring.

The variational bias correction system for satellite radiances used in ERA-Interim was developed at ECMWF and implemented in operations in 2006 (*ECMWF Newsletter No. 107*, page 18). This system detects the appearance of a new satellite data stream, and it then initialises, updates, and keeps track of bias estimates for radiance observations from all channels for each sensor flying on the satellite. The bias estimates are updated during each analysis cycle by including parameters for that purpose in the control vector used to minimise the 4D-Var cost function. This ensures that radiance bias estimates are continuously adjusted to optimise the consistency of the corrected radiances with all other information used in the analysis (i.e. the conventional observations as well as the model background). An important practical advantage of this approach is that it removes the need for manual tuning procedures, which are prone to error and simply impractical in the modern age.

ERA-Interim is the first ever reanalysis produced with a fully automated bias correction system. Previous operational experience with variational bias correction of radiance data has been confined to numerical weather prediction (NWP) applications, first at NCEP (National Centers for Environmental Prediction) and more recently at ECMWF. For an NWP system the ability to automatically detect new data and quickly develop bias estimates without human interference is not as crucial as it is for reanalysis, where data events happen much faster than they do in real time. And since the natural mindset in the NWP context is to look forward rather than backward, the long-term performance and stability of the adaptive approach to bias correction has not previously been documented.

A major concern in reanalysis is the ability to accurately represent climate trends and variability. This is a key requirement for the ERA-Interim CDAS and future reanalysis systems, if they are to be useful for climate monitoring and the assessment of climate change. The main difficulty is that changes in the observing system, combined with the presence of biases in models and observations, can cause shifts and trends in

reanalyses that interfere with true climate signals. There is a general tendency over time towards increasing data coverage in all dimensions, but this occurs in bursts and spurts, rather than continuously. Most satellite observations require bias corrections before they can be usefully assimilated, and the biases often depend not only on instrument characteristics but also on atmospheric conditions. The challenge for reanalysis is to smoothly handle data events and bias changes, to minimise their effect on the representation of trends and variability, and, where possible, to quantify the expected uncertainties in the reanalysis products.

Basic performance aspects

As in any modern NWP system, the quality of a reanalysis increasingly depends on the way that satellite data are handled. This is certainly the case for ERA-Interim which covers the data-rich period from 1989 onward. The ability of the analysis system to generate optimal bias corrections for satellite observations should therefore be reflected in standard quality measures, such as the fit to conventional observations and the quality of forecasts initialised with ERA-Interim analyses.

Figure 1 compares the fit to radiosonde temperature observations in the southern hemisphere of ERA-Interim with that of ERA-40. These results pertain to a particular month (August 2001), but are representative for the entire period 1989–2002 when ERA-Interim and ERA-40 overlap. Figure 1b shows that the mean fit of ERA-Interim reanalysed temperatures to radiosondes is better than that of ERA-40. This indicates that the variational analysis is able to generate bias corrections for the radiance data that render them consistent with the radiosondes, even though the former greatly outnumber the latter. In terms of the root mean square fit, Figure 1a shows that the ERA-40 analysis fits the radiosonde data slightly better than ERA-Interim, but the background errors for ERA-Interim are much smaller. This implies that the time-consistency of the assimilation has improved in ERA-Interim; ERA-40 draws closer to the individual station data but requires large corrections (analysis increments) to do so.



Figure 1 (a) Root-mean-square and (b) mean fit to southern-hemisphere radiosonde temperature reports for August 2001. Black curves are for ERA-Interim and red for ERA-40; dashed curves for analysis fit and solid for background fit. Figure 2 shows the quality improvements in ERA-Interim as measured by forecast skill. For reference, Figure 2a shows the evolution since 1989 of the skill of the ECMWF operational forecasting system, in terms of anomaly correlations of 500 hPa height forecasts for both hemispheres. The corresponding results for ERA-Interim and ERA-40 are shown in Figure 2b. By this measure, the quality of the ERA-Interim reanalysis is impressively uniform in time and space. Throughout the reanalysis period the forecast skill is similar to that of the operational system around 2002, when the spectral resolution of the model was T511 (compared to T255 for ERA-Interim).

Many other aspects of ERA-Interim product quality have improved substantially relative to ERA-40, as described in *ECMWF Newsletter No. 110* (page 25) and *No. 115* (page 12). These include well-documented difficulties such as the representation of the hydrological cycle and the strength of the Brewer-Dobson circulation. Since the observations assimilated in ERA-Interim are largely those used in ERA-40, much of the improvement is due to progress in modelling and data assimilation achieved at ECMWF since the production of ERA-40. The precise contribution of improved radiance bias corrections to this general picture is not known. However, as we shall see below, it can be clearly demonstrated in some situations that the variational approach to bias correction results in better use of observations, and is therefore beneficial to the reanalysis.



Figure 2 Mean anomaly correlations for 3-day, 5-day, and 7-day forecasts of 500 hPa geopotential height in northern and southern hemispheres for (a) operational forecasts and (b) ERA-Interim and ERA-40 forecasts. All forecasts are verified against their own analyses.



2001

2003

2005

2007

1999

NOAA-10

NOAA-11

NOAA-12

NOAA-14

1989

1991

1993

1995

1997

MSU instrument errors

The long record of radiance bias estimates produced in ERA-Interim for multiple sensors flown on different satellites provides a wealth of information. Here we specifically focus on radiance data from MSU channel 2, which measures temperatures in a deep layer of the middle troposphere, with maximum sensitivity near 600 hPa. Figure 3 shows the global mean bias estimates for this channel for 1989–2007, for each of the four NOAA satellites that carried the MSU sensor during this period.

The results in Figure 3 have several notable features. There is considerable variability in the bias estimates on monthly and interannual time scales. The results for hemispheric averages (not shown) are very similar, implying that the variation in time is mainly due to global changes in the bias (i.e. the spatial structure of the bias as shown in Figure 3 is approximately stationary). There is some drift, especially for NOAA-11 during its first five years of operation, but this feature is not shared by other satellites and is therefore most likely due to an instrument-specific calibration issue. The global mean bias estimates for each satellite are stable, in the sense that they do not appear to drift indefinitely. Instruments on different satellites are biased relative to each other; for example, the offset between NOAA-11 and NOAA-12 is about 1.2 K on average.

The remarkable pattern of variability in the bias estimates begs an explanation. The MSU record, which extends back to late 1978, is considered a key data set for the assessment of climate change in the free atmosphere. MSU data have been used by various research groups to reconstruct the tropospheric temperature record in order to estimate trends and other climate signals. This involves the application of corrections to the data that account for calibration errors associated with each sensor, due to, for example, drift and/or decay of the satellite orbits. There is no universal agreement on the optimal method of correction, but each method relies to some extent on overlaps between pairs of satellites, comparisons with radiosonde observations, and modelling of physically-based calibration errors.

In deriving their corrections to the MSU record, *Grody et al.* (2004) used a calibration model for the instrument that includes the effect of orbital drift of the satellite. The change in equator crossing time due to the drift causes a variation in the total heat budget of the spacecraft, which in turn affects the temperature of the on-board warm target used for calibration. Figure 4b, taken from *Grody et al.* (2004), shows the NOAA-14 MSU warm target temperature changes during the lifetime of the instrument. These changes are remarkably similar to the bias estimates obtained in ERA-Interim, duplicated in Figure 4a to facilitate the comparison. It appears that the reanalysis is quite successful in detecting and correcting the complex calibration errors in the MSU observations.

This example clearly demonstrates the power of the adaptive approach to bias estimation, which can account for large and relatively rapid changes in the error characteristics of the instrument. Such an approach needs lots of information in order to succeed. In reanalysis, the variational bias correction is fundamentally a statistical method for cross-calibration of observations. It uses all the available data from multiple instruments, in addition to physical constraints provided by the forecast model.



Figure 4 (a) Global mean bias estimates from ERA-Interim for NOAA-14 MSU channel 2, as in Figure 3. (b) Recorded variations of the warm-target calibration temperature on board NOAA-14 from *Grody et al.* (2004).

Response to the Pinatubo eruption

A well-known problem with the ERA-40 reanalysis is the excessive precipitation over the tropical oceans after 1991. This was partly due to the method used for analysing humidity at the time, combined with the effect of assimilating increasing numbers of humidity-sensitive radiance observations from HIRS and SSM/I during the 1990s. The humidity analysis methodology used in the IFS (and, consequently, in ERA-Interim) was completely revised as a result of the lessons learned in ERA-40.

The tropical precipitation problem in ERA-40 was exacerbated by effects of the eruption of Mt. Pinatubo in June 1991. Large amounts of aerosol were injected in the lower stratosphere, resulting in significant cooling of the HIRS infrared radiances. During the weeks following the eruption, the radiances in the water vapour band (channels 11 and 12) changed by approximately 0.5 K when averaged over tropical latitudes. The aerosols from the eruption persisted in the atmosphere for several years. However, the radiative transfer model used for the data assimilation does not account for this type of change in aerosol concentration, nor does the assimilating forecast model. This means that the large signal seen by the HIRS data cannot be properly represented in the analysis. The response of the ERA-40 analysis was to adjust the humidity field in order to maintain the fit to HIRS data, causing a large injection of excess moisture in the tropical atmosphere. Introduction of NOAA-12 on 1 July 1991 with a second HIRS sensor made matters worse.

This situation presents an interesting challenge for the variational bias correction. In the absence of a realistic representation of the aerosols in the radiative transfer model, the only way to properly use the remaining information in the HIRS observations is to absorb the aerosol signal in the bias estimates. Figure 5 shows that this is in fact what happens in ERA-Interim. In the tropics, the bias estimates for HIRS channel 11 on NOAA-10 and NOAA-11 drop swiftly during the second half of June 1991 to remove the aerosol effect from the signal. The estimates for NOAA-12 HIRS when introduced immediately reflect the prevailing situation. A gradual return of the bias estimates to normal (pre-Pinatubo) values then takes place during the next few years.

The effect of the Pinatubo eruption on MSU and SSM/I radiances is different. Due to the long wavelengths in the microwave spectrum these instruments are not directly sensitive to the stratospheric aerosols produced by the eruption. On the other hand, absorption of radiation by the aerosols causes an increase in lower-stratospheric temperatures by several degrees. This signal is accurately measured by MSU channel 4, which has its peak sensitivity slightly above the tropical tropopause. The forecast model does not know about the anomalous stratospheric aerosol in this situation and therefore cannot predict its effect on temperature. As a result a slight cold bias develops in the model background, resulting in systematic departures from the MSU channel 4 radiances in the tropics.

The appropriate response in this case would be to correct the model bias in order to improve the agreement with the radiance observations, but the analysis system is not equipped to do this. Instead the system gradually increases the radiance bias estimates for MSU channel 4 during the second half of 1991, by approximately 0.45 K in the tropics, as shown in Figure 6. The amplitude of the signal in the uncorrected radiance departures during this period is nearly 3 K, so that the true signal in the data is reduced by about 15%. This has a small, but nevertheless adverse, effect on the representation of the temperature signal in the lower stratosphere. The impact is ultimately limited by the other assimilated data, including the lower-peaking MSU channels, temperature observations from radiosondes, and the HIRS radiances as previously discussed.







Figure 6 Tropical averages of variational bias estimates for MSU channel 4 radiance data from (a) NOAA-10, (b) NOAA-11 and (c) NOAA-12.

Impact of model errors

The adjustment of the MSU data following the Pinatubo eruption illustrates a potential weakness of the variational bias correction scheme in the presence of systematic model errors. It shows that the variational analysis adjusts the observations in order to control the bias in the departures, regardless of its source. The obvious danger is that the data are falsely corrected to compensate for model bias, which could then cause the assimilation to drift toward the model climate. However, there are two distinct factors that limit the potential for such a scenario.

First, the nature of the bias model (i.e. the choice of bias predictors) determines the types of corrections that can be made, and this can restrict the possibilities for aliasing with systematic model errors. For example, the use of scan bias predictors for radiance data that depend only on the viewing angle of the instrument is likely to produce corrections that truly reflect biases in the data and/or in the radiative transfer model. On the other hand, the air-mass dependent predictors often used for radiance bias correction could potentially explain model biases as well.

Second, the cost of making adjustments to any subset of observations depends on the resulting fit of the analysis to all other observations. The bias corrections produced for different channels and sensors must therefore be consistent with each other as well as with any other data used in the analysis. This is a powerful feature of the variational approach to bias correction, which provides it with a major advantage over alternative schemes that estimate biases relative to a fixed reference state.

The risk of contaminating observations by the effect of model biases on the variational bias corrections is therefore highest in sparsely observed situations where there are large-scale errors in the model background. Given the reality that forecast models do – and probably always will – have biases, the assimilation system requires a certain amount of anchoring information to remain stable, in the form of uncorrected (and preferably unbiased) observations. It is not clear how much and what kind of anchoring information is needed for this.

In ERA-Interim, the effect of model biases on the data assimilation is most clearly seen in the stratosphere. To avoid excessive drift, it was decided (see *ECMWF Newsletter No 111*, page 5) to constrain the upper layers of the model by using uncorrected radiance observations from the highest-peaking channels on successive Stratospheric Sounding Units (SSU) and Advanced Microwave Sounding Units (AMSU-A). Since each instrument has slightly different characteristics, this has resulted in spurious shifts in the reanalysis of the upper stratosphere that are visible, for example, in time series of global mean temperatures at 5 hPa and higher.

Drift in AMSU-A radiance observations

The benefits of microwave radiance data from the AMSU-A sensor for NWP at ECMWF and elsewhere are well known. AMSU-A is a 15-channel sounder measuring atmospheric temperature and humidity profiles. It represents an improvement over MSU in terms of spatial resolution, both horizontally and vertically. At the time of this writing AMSU-A sensors on five polar orbiting satellites are available for NWP, providing almost complete coverage of the earth every four hours. These data will become a mainstay of climate monitoring information for the ERA-Interim CDAS. However, since the AMSU-A record is still relatively short its suitability for this purpose has not yet been carefully assessed.

Figure 7 shows globally averaged bias estimates produced in ERA-Interim for AMSU-A channels 7, 6, and 5, for four different satellites. These three channels with overlapping weighting functions provide the bulk of information about tropospheric temperature; channel 7 peaks in the upper troposphere (400–150 hPa), channel 6 in the middle troposphere (600–300 hPa), and channel 5 in the lower troposphere (850–500 hPa).

The most noticeable feature in Figure 7 is the collective, nearly linear, downward trend in the bias estimates. When averaged over either hemisphere or tropical latitudes rather than globally the curves look similar. This implies that the trends and other variations in time of the bias estimates reflect global shifts rather than seasonal or regional changes. The downward trend is largest for channel 6 on NOAA-15 (nearly 0.5 K per decade), and this appears consistent with the other satellites. For channel 5 the bias estimates reduce by approximately 0.15 K per decade for NOAA-15, with less consistency among the different satellites. The bias estimates are positive for all channels, implying that the measured radiances are biased warm relative to the analysis. The estimates for the AQUA satellite are of the opposite sign, probably because of different pre-processing algorithms used by the data provider.



Figure 7 Global mean bias estimates for (a) channel 7, (b) channel 6 and (c) channel 5 of AMSU-A from NOAA-15, NOAA-16, NOAA-17 and AQUA.



Figure 8 Global mean tropospheric temperature anomalies at (a) 250 hPa (b) 500 hPa and (c) 850 hPa from ERA-Interim (red), ERA-40 (black), JRA-25 (blue), and NRA2 (purple). ERA-Interim anomalies are defined relative to the ERA-40 climatology; anomalies for all other products are relative to their own climatologies. JRA-25 is the Japanese 25-year reanalysis and NRA2 is the NCEP/Department of Energy reanalysis.

In view of the reputation of the AMSU-A instrument, these results are unexpected and, at first glance, rather disconcerting. The first worry is that the trend in the bias estimates reflects a slow drift of the reanalysis towards the model climate, similar to the drift in the upper stratosphere mentioned in the previous section. However, the model is known to have a slight cold bias in the troposphere. If the true bias in the AMSU-A data were constant in time, then a drift toward a colder model climate could only be accomplished by increasing the bias corrections, because the data would appear increasingly warm relative to the analysis. Instead, the corrections are decreasing, so that the analysis moves closer to the uncorrected data in a direction that opposes the model bias. This is confirmed by the globally averaged temperature increments produced in the troposphere, which are systematically positive in the reanalysis.

We can state with confidence, therefore, that there is no insidious drift toward the model climate. An alternative explanation is that the changes in the tropospheric AMSU-A biases found in ERA-Interim are real and reflect actual instrument errors. This is partly supported in a recent study by *Mears & Wentz* (2008), in which they attempted to merge recent AMSU-A data with the MSU record from earlier NOAA satellites in an effort to extend their MSU-based climate trend analysis for tropospheric temperatures. They found similar trends and inconsistencies in the AMSU-radiances from NOAA-15 and NOAA-16, and, in fact, most of these data were not included in their merged dataset.

The question remains: which information in the reanalysis is responsible for warming the troposphere? Figure 8 shows global mean temperature anomalies obtained from four different reanalyses (ERA-Interim, ERA-40, JRA-25, and NRA2) at three pressure levels (850, 500, and 250 hPa) for the period 1989–2007. ERA-Interim is consistently warmer in the lower troposphere, and its rate of warming during the AMSU-A period is at least equal to (and perhaps slightly exceeds) that in the other reanalyses. All radiance data (from AMSU-A, but also from HIRS, SSM/I, and GOES) are subject to variational bias correction, which effectively removes their mean signal and calibrates them to all non-radiance data used in the analysis. For tropospheric temperatures in particular, the latter primarily consist of radiosonde and aircraft observations. It must therefore be the case that the tropospheric warming in ERA-Interim largely responds to information from radiosondes and aircraft.

To check this, Figures 9 and 10 show the time series of global mean departures for temperature data from radiosondes and aircraft, for three tropospheric layers that approximately correspond to AMSU-A channels 5–7. When interpreting statistics for conventional observations one needs to consider their numbers and locations that are highly irregular in space and time. Both radiosonde and aircraft reports are concentrated in the northern hemisphere. Most of the temperature measurements from aircraft occur at the jet-stream level; lower-level reports are predominately over land and especially in the vicinity of airports. Global radiosonde data counts have declined somewhat in the 1990s, but are relatively steady compared with the large variations in aircraft data. The number of temperature data from aircraft is small during the first few years (hence the noisy departure statistics), but increases dramatically in 1999. This explains the sudden shift in mean departures with respect to the higher-level aircraft data noticeable in Figure 10.



Figure 9 Global mean departures from radiosonde temperature observations at (a) 200 hPa, (b) 500 hPa and (c) 850 hPa. The thin curves show the mean background (red) and analysis (blue) departures for each analysis cycle; thick curves are 30-day running means; values in degrees Kelvin indicated on the left axes. Normalised data counts in green; in units of 10,000 per day indicated on the right axes.



Figure 10 As Figure 9, but for aircraft temperature reports.

It has been known for some time that temperature measurements for many types of aircraft are biased warm relative to radiosondes, and this has been confirmed in a number of studies performed at ECMWF and elsewhere. Together with the increasing number of aircraft reports, this explains the opposing mean departures for the two types of data, both at the 200 hPa and 500 hPa levels. It is not clear, however, why the increase in bias with respect to radiosondes at higher levels apparently precedes the major increase in aircraft data counts by about 6–9 months. After 1999 the mean analysed temperatures are increasingly determined by aircraft data, which greatly outnumber radiosonde reports at all levels. This also affects the anchoring of the radiance data from the AMSU-A tropospheric channels and may explain the slow decrease in bias corrections for these channels. The problem can be alleviated by correcting the aircraft temperature data in order to render them consistent with the information from radiosondes. This would have a global impact on reanalysed temperatures via the variational bias corrections of the AMSU-A channels. The net effect would be to cool the reanalysis in the upper troposphere, possibly by as much as a few tenths of a degree Kelvin, and by a lesser amount in the lower troposphere.

Long-term assessments and climate monitoring

The hardest challenge in reanalysis is to properly manage all the changes in the observing system, and to avoid any negative effects that these changes may have on product quality. In ERA-Interim much of the data handling has been automated, including the detection and smooth introduction of new satellite data, estimating and correcting the biases in these data, managing data gaps, etc. This has allowed the reanalysis production to proceed with minimal interruption at a steady pace, without major mishaps.

Having completed nearly 20 years of ERA-Interim, we can now assess the long-term behaviour of the variational bias correction system. We have seen that the system is quite good at removing relative biases among different types of observations, which has helped reduce the occurrence of spurious shifts and other artefacts in the reanalysed fields. On the other hand, model errors can affect the bias corrections, especially in sparsely observed situations. The importance of anchoring data – trusted observations that do not require bias correction – is evident, as is the need to further improve the forecast model, especially in the stratosphere.

Is it possible to obtain accurate trend estimates for climate monitoring from a reanalysis system? We think yes and, in fact, reanalysis offers the best approach to this end. Any study involving the analysis of long data records requires observational quality control and bias correction, based on an assessment of uncertainties in the data. This can only be done by making use of additional information, from independent observations or in the form of physical laws as expressed in a forecast model. Reanalysis provides just the framework for integrating and reconciling all this information.

Further Reading

This article is an abridged version of a paper prepared for the 37th Session of the ECMWF Scientific Advisory Committee. The complete paper is available as *ECMWF Tech. Memo. No.* 575 (www.ecmwf.int/publications/library/do/ references/show?id=88715).

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

Mears, C.A. & F.J. Wentz, 2008: Construction of climate-quality atmospheric temperature records from the MSU and AMSU microwave sounders. Submitted to *J. Atmos. Ocean. Technol.*

© Copyright 2016

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

The content of this Newsletter article is available for use under a Creative Commons Attribution-Non-Commercial-No-Derivatives-4.0-Unported Licence. See the terms at https://creativecommons.org/licenses/by-nc-nd/4.0/.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error or omission or for loss or damage arising from its use.