Observing System Experiments (OSE) to estimate the impact of observations in NWP

Peter Bauer

European Centre for Medium-Range Weather Forecasts
Reading, UK
peter.bauer@ecmwf.int

Abstract

This paper gives a short overview of the purpose and design of observing system experiments used to assess the impact of individual or groups of space-based observing system components in NWP at ECMWF. The main categories of experiments are described and examples are presented. The complementary role of observing system monitoring, observing system experiment evaluation and advanced adjoint diagnostic tools is emphasized to understand the complex interaction between the sensitivity of observations to the atmospheric state, the NWP model and the data assimilation system.

1. Introduction

Observing system experiments (OSEs) evaluate the effect of adding or removing individual components of the observing system on NWP analysis and forecast quality. The motivation for running OSEs stems from the need to test system upgrades at operational NWP centres that undergo continuous development and to perform more systematic studies on the value of individual observing system components. The latter objective is complicated by the fact that the role of observing system components is not purely quantifiable in terms of analysis/forecast impact because the role of individual components may also be to provide absolute measurement reference points to stabilize bias correction schemes (Auligné et al. 2007). Additionally, the contribution of individual components to NWP skill may be more strongly driven by the accuracy of the model parameterizations and the flexibility of the data assimilation system rather than the actual quality of the observation.

Table 1 shows a list of the current satellite observing system assimilated at ECMWF. Satellite (radiance) observations provide 95% (90%) of the data volume in the analysis and, at present, observations from more than 50 instruments are used. The large amount of data and the large variety of derived information requires a complex system for the monitoring of data quality and consistency between model and data at the same time.

Both experiment and operational model evaluation is usually performed using established methods. In data space, a comprehensive monitoring suite is operated at ECMWF that allows the routine evaluation of data vs. model statistics. These are obtained from the comparison of short-range forecasts and analyses combined with radiative models (for radiance data) and observations[1]. The monitoring is updated with every analysis and allows immediate response to data/model problems and the evaluation of accumulated statistics, e.g. for the purpose of bias correction tuning and data selection.

Many studies on OSEs performed in the context of model upgrades have been performed at ECMWF (refer to technical memoranda archive[2]). There have been various systematic studies of general observing system impact, for example Kelly and Thépaut (2007) and Bauer and Radnóti (2009). As instructive as these studies are they also demonstrate the limited general validity of OSEs due to (1) insufficient statistical significance

---

[1] http://www.ecmwf.int/products/forecasts/d/charts/monitoring/satellite

[2] http://www.ecmwf.int/publications/library/do/references/list/14

of the results produced by too short experimentation periods, (2) ambiguity of forecast verification from choice of verifying model analyses, (3) significant dependence of results on NWP system configuration that requires repetition of OSEs along with NWP system changes.

In particular the second issue is currently unsolved and requires more sophisticated diagnostic tools while the first and third issue are mainly driven by limited computing resources. This paper briefly summarizes the main areas of observing system experimentation and provides a few examples.

Table 1: Satellite observing system used in ECMWF analysis (October 2009).

Observation type	Satellites/Instruments
Radiances	AMSU-A on NOAA-15/18/19, AQUA, Metop AMSU-B/MHS on NOAA-17/18, Metop SSM/I on F-13/15, AMSR-E on Aqua HIRS on NOAA-17/19, Metop AIRS on AQUA, IASI from Metop MVIRI on Meteosat-7, SEVIRI on Meteosat-9, GOES-11/12, MTSAT-1R imagers
Bending angles	COSMIC (6 satellites), GRAS on Metop
Ozone	Total column ozone from SBUV on NOAA-17/18, SCIAMACHY on Envisat, OMI on Aura
Atmospheric Motion Vectors	Meteosat-7/9, GOES-11/12, MTSAT-1R, FY-2C, MODIS on Terra/Aqua
Sea surface parameters	Near-surface wind speed from Seawinds on QuikSCAT, Scatterometer on ERS-2, ASCAT on Metop Significant wave height from RA-2/ASAR on Envisat, Jason altimeter

2. Observing system experiments

OSEs serve a number of different purposed, possibly the most important of which is related to the testing of new instruments to be assimilated in future model versions. Other applications are described with more detail below. The testing of new instruments is fundamental prior to their activation in an operational system and it occupies the largest resources associated with OSEs.

Experiment analysis verification is mostly based on model vs. observation statistics with the objective of improving the fit of the model to observations with any modification of the system. These statistics are rather robust due to the large data volumes available in the analysis.

Forecast verification is usually performed using analyses as the reference, but for a limited set of parameters radiosonde verification is also available. For the latter, however, representativeness is an issue. Forecast verification with analyses can be carried out in different configurations, namely as a verification against the respective experiment's own analyses. This assumes that the observing system in the experiment is modifying the mean analysis state such that a single (say operational) analysis is not an appropriate reference. Verification against operational analyses is justified if the experiment configuration is obviously inferior compared to the operational system, e.g. in terms of spatial resolution or the observing system. There is, however, a risk of introducing a bias towards the operational observing system. All this has to be kept in mind when forecast scores are assessed.

2.1. Adding (improving) a new observation type

Once new instruments become available at ECMWF, a number of technical modifications have to be performed such as the conversion of the received data format to BUFR according to WMO conventions, the BUFR conversion to the observational database (ODB) that is used in analysis system, the set-up of the management of unique satellite/instrument identifiers in system, the generation of radiative transfer model coefficients (see RTTOV, Matricardi et al. 2004) in case radiance data is used, the definition of data screening (quality control for data problems, clouds or surfaces effects) and the implementation of the new data type in the variational bias correction if desired. Most of these technical changes are rather common among operational centres.

First, monitoring experiments are performed in which observations are routed through the system as all other observations expect for their active assimilation. This allows to tune quality control and to spin up bias corrections without shocking the system. Usually, the activation/deactivation of data is controlled by so-called blacklists that are outside the computer code so that they can be easily modified on short notice.

Only then the actual OSEs are run by activating the new data. The assessment of data impact consists of the analysis and the forecast impact evaluation. The former usually focuses on the impact of the data on the fit between short-range forecast and analysis and all other observations. This evaluation is performed in observation space, i.e. in the units of the respective observation type, say temperature, specific humidity and wind for radiosondes and radiances for most satellite observations. The advantage of having a large set of different data types already available in the system is that very robust statistics across a wide range of parameters, levels and regions can be derived. The aim is to produce a better and more consistent analysis that is usually defined by a better fit of short-range forecasts and analyses to the observations. In certain cases, also the assessment of the impact on the mean analysis state is important if certain observations constrain the model’s climate.

An example for this assessment is given in Figure 1 that shows the dependence of analysis departure standard deviation between model and selected HIRS and MHS channels as a function of weight given to 10 additional IASI water vapour channels. These channels have been introduced in the ECMWF data assimilation system in autumn 2009 and the tests have been performed with the full operational (observing) system. The weight is modified by varying the observation errors associated with these channels. Figure 1 shows that for an observation error of ~1.5K the fit of other moisture sensitive satellite observations is consistently reduced by up to 5%.

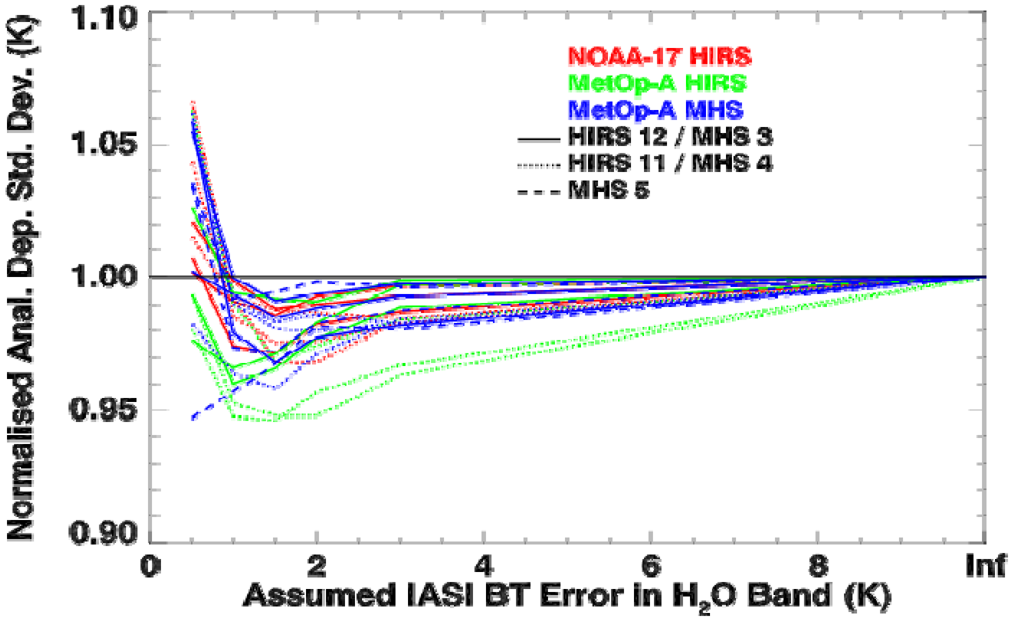

Figure 1: Normalized analysis departure standard deviation between model and different HIRS and MHS channels as a function of weight given to ten additional IASI water vapour channels in the assimilation system (Courtesy Andrew Collard).

However, adding these channels also proved difficult to verify in forecast terms despite the clearly positive impact on the analysis. Only by verifying both experiment and control with the experiment's analysis, positive forecast scores could be obtained. This underlines the ambiguities associated with forecast verification by analyses in general. This is in particular problematic for moisture related changes that affect the forecast only for a few days knowing that the choice of verifying analysis mainly affects the first days of the forecast range.

2.2. Investigating fundamental observation impact

In certain cases OSEs with rather basic observing systems are performed if the purpose is the understanding of fundamental mechanisms in 4D-Var. Also, the effect of removing individual observation types from a rich observing system does not necessarily produce the same relative ranking as adding those observation types to a poor system. At ECMWF, the 'poor' system usually contains all conventional observations and, to not disadvantage the Southern hemisphere, Atmospheric Motion Vectors (AMV) mainly derived from geostationary imagery and, if necessary, a single AMSU-A. The latter has proven to be important in cases where humidity observations are investigated that require a higher degree of temperature/dynamics analysis quality than provided by conventional observations and AMV (Kelly et al. 2007).

Recent examples for this kind of OSEs are the investigation of the impact of GPS radio occultation observations on the variational bias correction of radiances and the understanding of the main analysis mechanism responsible for producing wind increments from geostationary satellite radiance observations mostly sensitive to moisture (Peubey and McNally 2009).

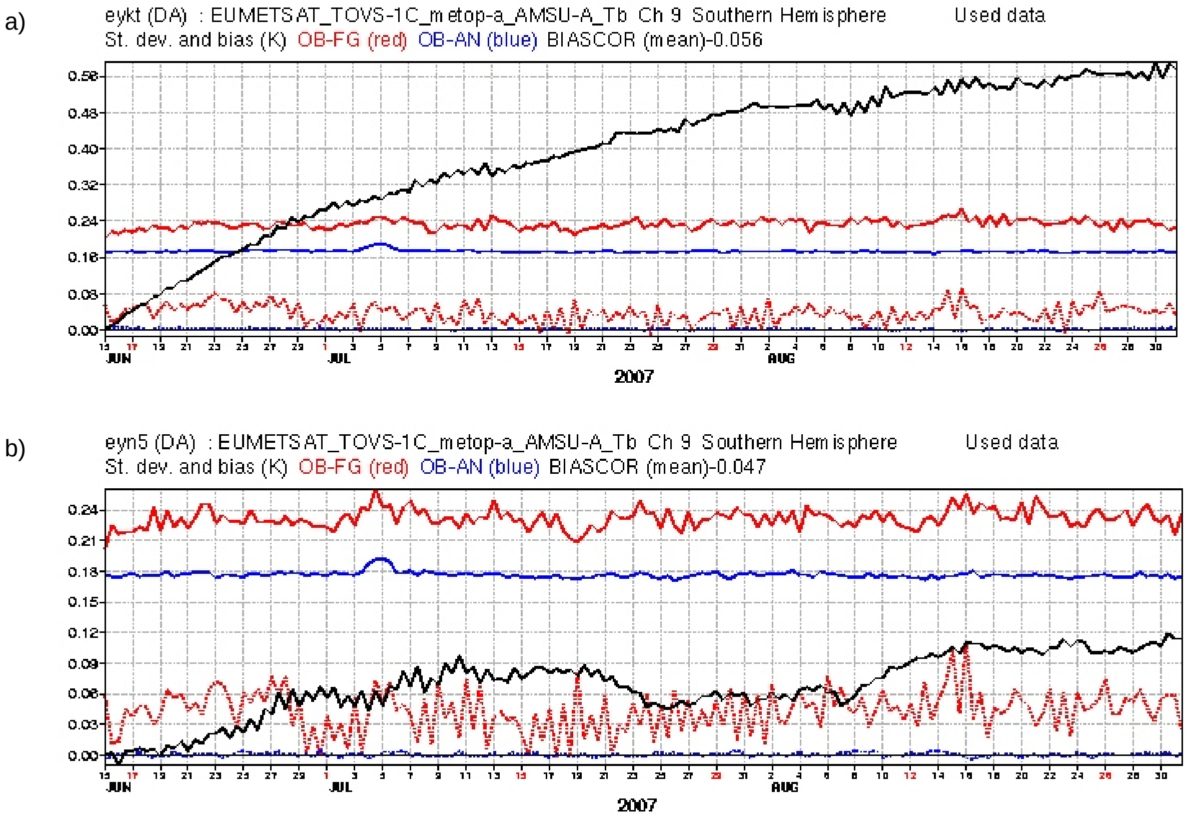

Figure 2: Time series of AMSU-A channel 9 departure statistics and bias corrections: first-guess departure standard deviation (red solid) and mean (red dotted), analysis departure standard deviation (blue solid) and mean (blue dotted), mean bias correction (black solid). Units are degrees Kelvin; data is limited to the Southern hemisphere; control experiment (a) and control experiment plus COSMIC GPSRO observations (b) (Courtesy Sean Healy).

The former OSE is illustrated in Figure 2. An experiment with a basic observing system has been set up that only contained conventional data, 1 AMSU-A and 1 MHS and that had been initialized with the operational model. A second experiment has been run in which GPSRO observations from 6 COSMIC satellites have been added to this system. In both cases, the variational bias correction was activated for the radiance data but GPSRO observations are considered bias-free.

Figure 2 demonstrates that GPSRO observations greatly reduce the amount and spin up of the radiance bias correction; here of AMSU-A channel 9 that is most sensitive to temperatures near 100 hPa. It can also be concluded that a large part of the bias correction is actually due to model biases in the upper atmosphere that is stronger constrained by the GPSRO observations in the analysis. The example therefore demonstrates (1) the interaction between model biases and observation bias corrections and (2) the complementarity between different types of observations.

2.3. ‘Continuous’ observation impact assessment

The assessment of all individual components of the observing system is a very costly exercise and is only rarely performed. Examples have been mentioned already in the Introduction. Again, these experiments can be performed by withdrawing individual types from the complete system or by adding them individually to a poor system. This can be done for groups of observations, for example, those constraining temperature, moisture, ozone, surface or cloud/precipitation analysis.

The continued assessment of the individual components of the global observing system is currently only performed through operational radiance monitoring (departures, biases) and, in the near future, adjoint diagnostic studies.

3. OSEs and adjoint diagnostics

Recently, new diagnostic methods for the assessment of the observational impact on both analyses and forecasts have been developed (Cardinali et al. 2004, Langland and Baker 2004, Cardinali 2009, Gelaro and Zhu 2009). These methods employ the basic operators available in 4D-Var data assimilation systems to quantify the sensitivity of the analysis state or forecast error to individual observations. These tools are attractive in that they measure the impact of observations in the context of all other observations present in the assimilation system while OSEs require the removal of observations. Adjoint methods are therefore the better option for routine monitoring because they can be run alongside sequential analysis/forecast cycles.

There are other fundamental differences between adjoint methods and OSEs, for example, that the modification of the observing system in OSEs affects analyses and forecast cumulatively while adjoint systems perform an evaluation within the system for each cycle independently (see Gelaro and Zhu for more details).

It can be demonstrated that the results of OSEs and adjoint systems show similar features but that they are not identical due to the above mentioned differences. They should therefore be regarded as complementary tools.

4. Summary and conclusions

In NWP, OSEs are continuously performed for the assessment of new (revised) observation impact along model updates, the study of basic impact features (often using poor observing system baselines) and the systematic assessment of all observing system components. The impact is currently evaluated using the fit to short-range forecast/analysis model fields (consistency, reference observations) and the model forecast skill using standard scores.

While this represents a well established framework there are fundamental shortcomings of current observation impact assessment, namely:

- the evaluation of individual observation type impact on fit of model fields to other observation types is only available for analyses but not forecasts;
- standard diagnostics for tuning/optimization of the observing system is not available for the purpose of thinning, channel selection, definition of observation errors;
- general overview diagnostics require large and costly sets of OSEs, no continuous built-in evaluation is available yet;
- standard forecast scores often contradict the results from the analysis evaluation, e.g. when new observations add noise to the analysis (i.e. from performing more work in the analysis) and therefore may increase root-mean-square difference between forecasts and analyses used for verification.

In particular the latter issue inhibits a consistent evaluation of performance between analyses and forecasts and may require alternative approaches to forecast error evaluation, for example, through probabilistic measures (Tan et al. 2007).

In the future, the most potent evaluation tool is expected from the comparison of forecasts against all observations that are also available in the analysis through comparison in observation space. This will provide sufficient statistical significance, the observations can be considered independent with increasing forecast range and the available observations will provide a vast amount of information on the atmospheric state given the wide range of employed satellites/instruments/data types. Further, the combined use of OSEs and adjoint diagnostics is expected to provide complementary information on data impact in both analysis and forecasts. Adjoint diagnostics are computationally cheap and allow a continuous monitoring of performance.

Another important aspect is the link between observational impact and model physics. This link can be rather complex if complex physical processes propagate the observational information through analysis and forecast.

Acknowledgements

The presented work was compiled with contributions from Niels Bormann, Carla Cardinali, Andrew Collard, Mohamed Dahoui, Alan Geer, Ron Gelaro, Sean Healy, Tony McNally, Gabór Radnóti, Jean-Noël Thépaut.

References

Auligné, T. A.P. McNally, and D.P. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. Roy. Meteor. Soc.*, **133**, 631-642.

Bauer, P. and G. Radnóti, 2009: Study on Observing System Experiments (OSEs) for the evaluation of degraded EPS/Post-EPS instrument scenarios. Report available from ECMWF, Reading, UK, 99 pp.

Cardinali, C., S. Pezzulli, and E. Andersson, 2004: Influence-matrix diagnostic of a data assimilation system. *Q. J. Roy. Meteor. Soc.*, **130**, 2767-2786.

Cardinali, C., 2009: Monitoring the observation impact on the short-range forecast. *Q. J. Roy. Meteor. Soc.*, **135**, 239-250.

Gelaro, R., and Y. Zhu, 2009: Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus*, **61A**, 179–193.

Kelly, G. and J.-N. Thépaut, 2007: Evaluation of the impact of the space component of the Global Observing System through Observing System Experiments. *ECMWF Newsletter*, No. 113, 16-28.

Kelly, G., P. Bauer, A.J. Geer, P. Lopez, and J.-N. Thépaut, 2007: Impact of SSM/I observations related to moisture, clouds and precipitation on global NWP forecast skill. *Mon. Wea. Rev.*, **136**, 2713-2726.

Langland, R.H., and N.L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189–201.

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

Peubey, C. and A.P. McNally, 2009: Characterization of the impact of geostationary clear-sky radiances on wind analyses in a 4D-Var context. *Q. J. Roy. Meteor. Soc.*, submitted.

Tan, D.G.H., Andersson, E., Fisher, M., and L. Isaksen, 2007: Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. *Q. J. Roy. Meteorol. Soc.*, **133**, 381–390.