ECMWF Feature article

from Newsletter Number 107 – Spring 2006

METEOROLOGY

A variational approach to satellite bias correction

www.ecmwf.int/en/about/news-centre/media-resources
doi:10.21957/mn7643f9s8

This article appeared in the Meteorology section of ECMWF Newsletter No. 107 – Spring 2006, pp. 18–23.

A variational approach to satellite bias correction

Tony McNally, Thomas Auligné, Dick Dee, Graeme Kelly

Satellite instruments, like any other measurement system, are not perfect and are prone to error. While errors which are purely random are undesirable (such as noise at the radiation detector) their adverse effect can be significantly reduced within a data assimilation scheme by a combination of spectral, spatial and temporal filtering/averaging. However, errors which are systematic (i.e. a bias) cannot be handled in this way. Any observation which is biased can systematically damage the data assimilation scheme and ultimately the quality of the forecasting system. There are well documented examples of particular countries using certain (biased) radiosonde instrument types which introduce local anomalies in the analysis when the data are assimilated. However, biases in satellite observations are of particular concern as they have the potential, if uncorrected, to damage the NWP system globally in a very short space of time.

At ECMWF considerable effort has gone into dealing with satellite biases and a new system has been developed based on an objective approach. The Variational Bias Correction (known as VarBC) updates satellite bias corrections every analysis cycle. VarBC is being tested for operational implementation.

Diagnosing satellite biases

The bias in a particular satellite observation can only be determined by comparison with some unbiased ground truth. Special measurement campaigns involving surface or aircraft mounted radiometers and balloon borne sensors (launched to coincide with the satellite overpass) provide a highly accurate ground truth, but are inevitably limited to certain locations and times. While they can (and have) been used to expose problems with satellite instruments and/or radiative transfer models, they cannot tell us about potentially important variations in the systematic errors for different areas (e.g. biases over warm humid tropical oceans compared to cold dry polar regions).

Note that the radiative transfer model, henceforth referred to as the RTM, links the atmospheric state to the radiation measured by the satellite and for the purposes of this discussion may be considered as an integral component of the satellite observation.

The limited spatial and temporal availability of campaign data has led to the practice of NWP centres monitoring and diagnosing satellite biases using the NWP assimilation system itself. The obvious advantage of this approach is the in-house real-time availability of what is arguably (in the case of the NWP analysis or short-range forecast) the best estimate of the global atmospheric state. However, this approach has the disadvantage that it does not provide a completely unbiased ground truth. Indeed for some atmospheric variables and regions of the atmosphere, biases in the NWP system can be a significant fraction of (or even exceed) the biases we are attempting to diagnose in the satellite information. Despite these concerns the overwhelming benefits of monitoring satellite biases against the NWP system have led to its widespread adoption.

What biases do we see and where do they come from?

In general the biases we observe when satellite data are monitored against the NWP model are not fixed offsets as the pure use of the word "bias" might suggest. Instead they can vary with time (e.g. diurnal or seasonal changes), with geographic location or air-mass (including changes in the underlying surface e.g. land/sea/ice), and even with the scan position of the satellite instrument. Some examples are shown in Figure 1.

These biases we observe between the data and the model arise due to systematic errors in any one of (but more usually a combination of) the following sources:

- The satellite instrument itself (e.g. due to poor calibration/characterization and adverse environmental effects).
- The radiative transfer model (RTM) (e.g. errors in the physics/specroscopy and non-modelled atmospheric processes).
- Systematic errors in the background atmospheric state provided by the NWP model used for monitoring.

In principle we do not wish to correct the observations for the latter because that would only reinforce the model errors. Given the complicated nature of the various sources of systematic error (and how they may combine in a complicated way), it is not surprising that a considerable amount of effort has been directed towards bias correction at ECMWF and other NWP centres.

 doi:10.21957/mn7643f9s8

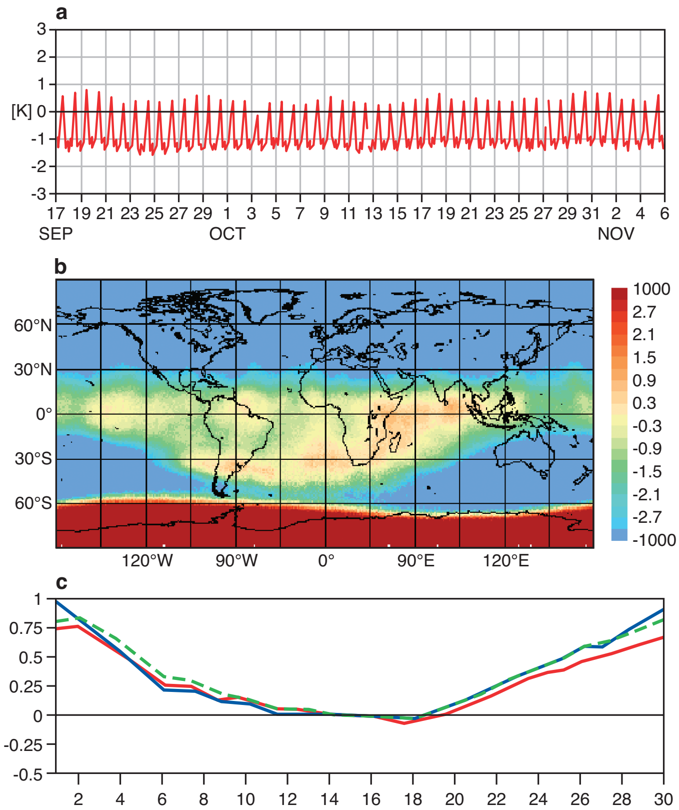

Figure 1 Examples of bias variation. (a) Diurnal varying bias in a METEOSAT window channel. (b) Air-mass (geographically) dependent bias in AMSU-A channel 14. (c) Scan dependent bias in AMSU-A channel 7.

A brief history of satellite bias correction at ECMWF

The very first attempts to assimilate satellite data at ECMWF assumed fixed constant offsets applied to each channel, but these were very quickly exposed as inadequate (particularly for the rather poorly calibrated MSU and HIRS instruments that were available at the time, see *Kelly & Flobert,* 1988) and a more sophisticated correction was needed. A scheme which aimed to apply a geographically varying bias correction (depending on air-mass) was proposed in *Eyre* (1992). The air-mass was characterized by predictors based on the radiance observations themselves and the appropriate bias correction generated by linear regression (having been trained previously on a representative sample of observed minus background radiance departures). A modification of this scheme where the observation based predictors were replaced by NWP model based predictors of air-mass was proposed in *Harris & Kelly* (1998) - this has survived in essence to the present day. The aim of the scheme (and its predecessor) was very clear: to predict and apply a correction to the satellite data that effectively removed every systematic departure between the NWP model and the observed radiances (i.e. correct the data to the model). To this end the regression was given many degrees of freedom (predictors) and the regression coefficients were frequently re-trained (or updated) to capture any possible drift in time. This arguably proved a very successful strategy and played no small part in the operational assimilation of satellite radiance data becoming so well established at ECMWF.

In the late 1990s two coincident events challenged the existing view of how satellite bias correction should be done.

- The arrival of radiance data from the Advanced Microwave Sounding Unit (AMSU-A).
- The extension of the ECMWF NWP model boundary into the upper stratosphere and lower mesosphere.

In the past there were many examples where satellites had been rather poorly calibrated and/or not particularly stable in time, while the model (at least in comparison) was thought to be relatively unbiased. When stratospheric radiance observations from the AMSU-A were monitored against the new NWP model it quickly became clear that situation had been reversed (see Figure 2).

There was strong evidence from a number of independent sources that the stratospheric temperatures provided by the NWP model had large time varying systematic errors whereas the AMSU-A (while not perfect) was rather well calibrated and very stable. Thus the rather courageous decision was taken to assimilate the uppermost channel of the AMSU-A with no bias correction applied.

Having sown the seeds of heresy and questioned the strategy of bias correcting the radiance data completely towards the model, investigations soon started to reveal evidence (although not as clear as the stratospheric example) of other aspects where a significant component of the bias correction being applied to the radiance data was actually due to systematic model error. Over the following years a number of steps were taken to mitigate this problem.

- The practice of frequent updates of the bias coefficients was abolished, acknowledging that the AMSU-A instruments (which formed the backbone of the assimilation system in the early 2000s) were generally stable and that time varying biases were most likely seasonal variations in the model systematic error.
- Some of the air-mass predictors which allowed the bias correction system to absorb known errors in the NWP (e.g. stratospheric and humidity) were removed (at least for some sensors) and partially replaced by new physically based corrections (so called gamma corrections, *Watts & McNally,* 2006) introduced into the AIRS and AMSU-A RTMs.

There has thus been a slow evolution towards a more minimal and constrained bias correction of the satellite radiances which aims to correct only biases in the observations and RTM, and not remove NWP model error (although there are some examples of pragmatic departures from this). However, to this day any constraint placed upon the bias correction has always been rather subjective and ad hoc. It has ultimately depended on which predictors we choose to make available to the bias correction scheme and how frequently we choose to update the corrections.

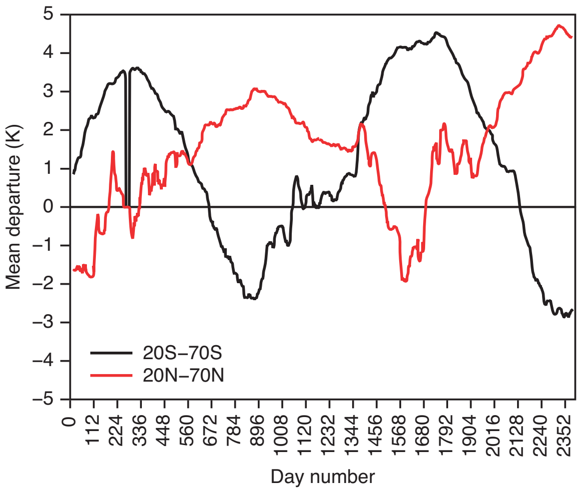

Figure 2 Time varying systematic errors in the temperature of the model stratosphere for the Northern Hemisphere from 20°–70°N (red line) and Southern Hemisphere from 20°–70°S (black line) between 1 November 1998 and 26 June 2000 compared against radiances in channel 14 from the NOAA-15 AMSU-A instrument.

A new way of computing and updating satellite bias corrections

When changes to the bias correction are made, the usual measure of success is whether the new system improves the fit to other assimilated observations (such as radiosondes which have not been bias corrected to the model). While the changes in the past have always been motivated by sound reasoning and often in response to an identified problem, the process of updating the satellite bias correction has effectively been on a trial and error basis. A new system has been developed within the Integrated Forecasting System (IFS) which aims to put this process on a more objective footing, based on ideas developed at NCEP and implemented in their operational 3D-Var system in the mid-nineties (*Derber & Wu,* 1998).

The Variational Bias Correction (VarBC, *Dee*, 2004 and *Auligné & McNally,* 2005) updates the satellite bias correction (not the choice of predictors, but the regression coefficients) every analysis cycle (e.g. every 12 hours). However, it does this *inside* the assimilation system by finding corrections which minimise the systematic differences between the satellites and model while simultaneously preserving (or improving) the fit to other observed data inside the analysis. This is achieved by including the regression coefficients used for bias correction in the control vector of the variational analysis, so that the bias estimates are adjusted simultaneously with the model trajectory based on all information available to the analysis. The fitting is optimal in that it respects the uncertainty of the observations and any background or inertia constraints we wish to impose on changes to the satellite bias estimation. In essence the VarBC can objectively estimate what proportion of the systematic differences between the satellites and the model should be corrected (and what should not) on the basis of all the other information we have available within the assimilation system.

 doi:10.21957/mn7643f9s8

Another significant benefit of the VarBC approach is that it goes a long way toward *automating* the updating and management of satellite bias corrections. The need for an adaptive bias correction system became painfully obvious during the production of ERA-40, which had to be interrupted and restarted on many occasions for manual retuning of the bias corrections. In the early days of satellite data assimilation we were typically using two satellite instruments which together provided radiance information in about 20 channels. Today in operations the 30 or so satellite instruments provide radiance information in more than 500 channels. When we consider that all of these may require different bias corrections depending on the environment in which they are used (e.g. the operational assimilation suite, the experimental *e-suites* and general research department experiments) it can be seen that even technical management of these is very difficult. Add to this the effort involved to monitor possible changes and manually decide (in the event of a change) what new bias correction should be applied, we see that automation of the process is practically unavoidable.

Early successes

The VarBC has been tested within the ECMWF 4D-Var assimilation system (but can be applied inside a 3D-Var system equally well). It has proved to be technically very robust and produced some dramatic improvements. Figure 3 shows the mean fit of the Cy30r1 assimilation system to radiosonde temperature observations averaged over the Northern Hemisphere for a period in August 2005.

It can be seen that the radiosonde temperature data suggest there is a cold bias in the short-range forecast background and in the analysis for the lower stratosphere. Similar statistics for radiances from the AMSU-A in channels sensitive to the lower stratosphere show no such disagreement. However, closer inspection reveals that this agreement is only by virtue of their bias correction towards the model. In the absence of this correction the radiance data would also suggest a similar cold bias. This situation has been known about for some time, and an obvious interpretation is that the short-range forecast does have a cold bias in the stratosphere, which is sustained in the analysis by the assimilation of wrongly bias corrected radiances. Attempts in the past to manually resolve this problem (e.g. by completely removing bias corrections from some of the stratospheric AMSU-A channels) have failed to achieve an appropriate balance between different overlapping AMSU-A channels. The same radiosonde data fits after the VarBC has been allowed to adjust the satellite bias correction are shown in Figure 4.

The time evolution of the bias correction in AMSU-A channel 10 and the fit to radiosonde data at 50 hPa are shown in Figure 5. The striking improvement in the radiosonde agreement is achieved by the VarBC progressively reducing the amount of bias correction applied to the radiance data. This results in more of the information from the AMSU-A forcing mean increments and warming the analysis accordingly. However, the successful removal of the cold bias has not been achieved quickly. The VarBC has taken several weeks of assimilation to gradually reduce the original satellite bias correction to a more appropriate level.

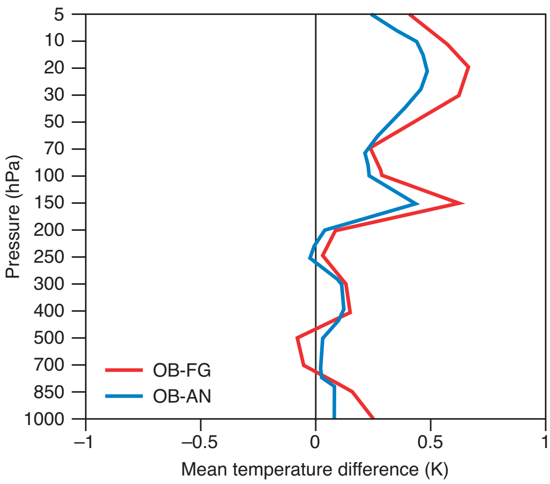

Figure 3 Mean observed minus background (red) and observed minus analysis fit (blue) for radiosonde temperatures in the Northern Hemisphere for August 2005 using the Cy30r1 assimilation system.

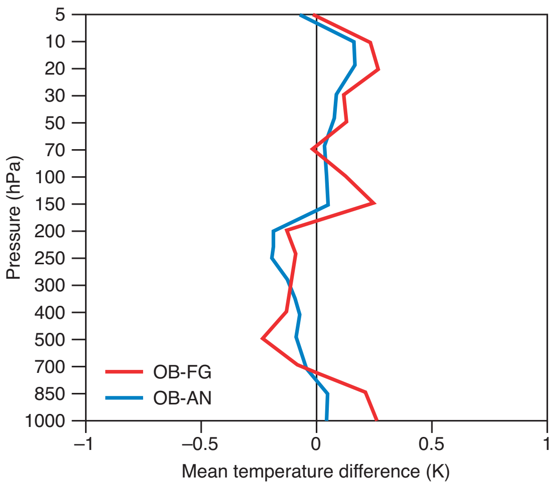

Figure 4 Mean observed minus background (red) and observed minus analysis fit (blue) for radiosonde temperatures in the Northern Hemisphere for August 2005 after VarBC has adjusted the satellite bias correction.

Another success of the VarBC is in the handling of satellite instrument changes. From previous reanalyses there are a number of well documented cases where a satellite instrument has suddenly degraded (or been contaminated by an extreme event such as volcanic emmisions). The usual result, if the event is not known about in advance, is a serious contamination of the analysis (as happened in ERA-15). If the event is known about and expected, blacklisting the affected channel can still disturb the time consistency of the analysis. However, in this context the VarBC has demonstrated an ability to handle sudden systematic changes to the data and minimise damage to the analysis. An extreme example of this is shown in Figure 6 for November 1986 when a cosmic storm event changed the response of the microwave (channel 3) detector on the NOAA-9 satellite.

The sudden shift in values initially results in almost all of the observations being rejected from the analysis. However, as the VarBC system is progressively exposed to more of the shifted data it automatically adjusts the bias correction (black line) to compensate. After a few days a completely new bias correction is established allowing all the radiance data to be used again with very little disruption to the analysis system (as seen in the radiosonde data fits).

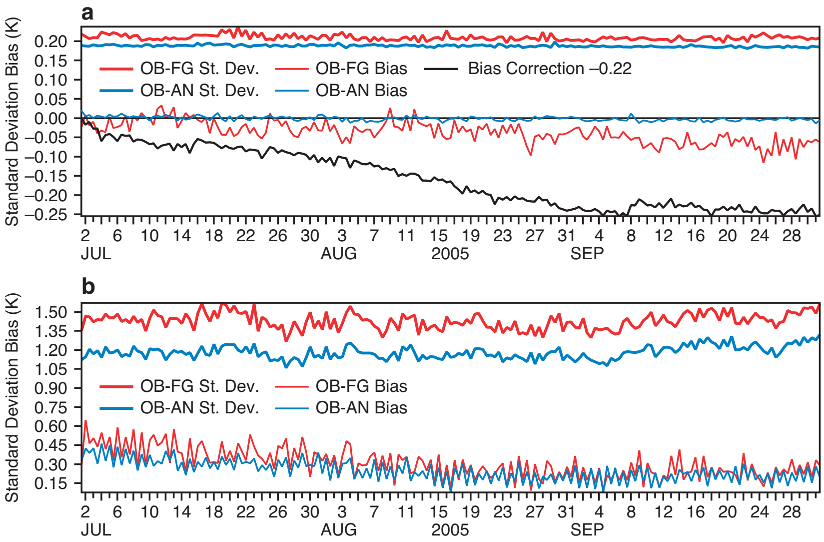

Figure 5 (a) Standard deviation (full line) and bias (dotted line) of observed minus background (red) and observed minus analysis fit (blue) for NOAA-16 AMSU-A channel 10 temperatures in the Northern Hemisphere, with the bias correction shown in black (offset by 0.22 K). (b)As (a) but for 50 hPa radiosonde temperatures in the Northern Hemisphere.

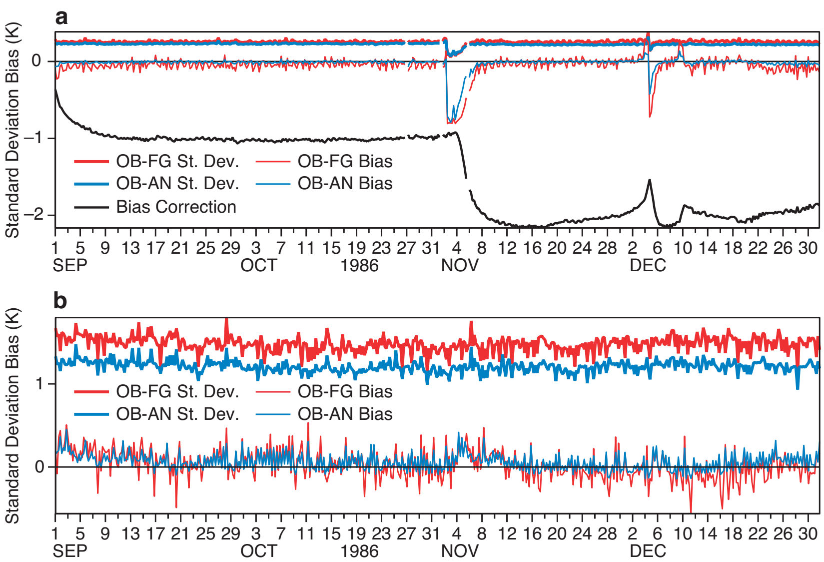

Figure 6 (a) Standard deviation (full line) and bias (dotted line) of observed minus background (red) and observed minus analysis fit (blue) for NOAA-9 MSU channel 3 temperatures in the Tropics, with the bias correction shown in black. (b) As (a) but for 200 hPa radiosonde temperatures in the Tropics.

doi:10.21957/mn7643f9s8

Some concerns

The two examples of the previous section show that the VarBC can successfully correct satellite biases even in the presence of systematic errors in the model. The ability to discriminate between different sources of bias (and apply appropriate corrections to the satellite data) clearly depends on the nature of the bias model used, as well as on the availability of some type of anchoring data such as radiosonde observations. For example, in the upper stratosphere where no observations exist (other than satellite data) the VarBC is prone to drifting with the NWP model error. Simulations predict that a completely unconstrained adaptive bias correction will cause the assimilation system to drift, ultimately to the climate of the NWP model by progressively applying larger and larger corrections to the data (to make the data look like the biased model). We have evidence of only a very small and slow drift in the experiments performed to date. This may indicate that the VarBC is still partially constrained by the chosen bias predictors and the indirect influence of radiosondes. However, this potential drift is a cause for concern (particularly in the reanalysis environment) and will require a robust solution.

The future for VarBC

The technical and scientific benefits of VarBC are such that it is currently being tested for operational implementation at ECMWF as part of the Cy30r2 upgrade. It is also being tested for application to the reanalysis project where, on the one hand, the automation of satellite bias correction is an inescapable practical necessity, but, on the other, the possible effects on the representation of the climate signal are a serious concern.

As a precursor to implementing the fully evolving scheme, static bias correction coefficients derived from a long offline run of the VarBC have already been implemented into operations on 8 February 2006. This has allowed an early exploitation of the improvements obtained in lower stratosphere that have been described in this article (i.e. the improved fit to radiosonde temperature data).

More work will be needed to address the concerns of drift in parts of the atmosphere unconstrained by conventional observations. The imminent availability of high quality temperature information from GPS radio occultation data (e.g. CHAMP) will provide more constraining observations for bias correction in the lower stratosphere. However, in the upper stratosphere and mesosphere we may revisit the choice of parameters available to the bias correction as this has proved, in the past, to be a powerful constraint upon the correction process. It remains to be seen any drift can be sufficiently controlled to satisfy the requirements of climate studies with reanalysis.

Having effectively automated the bias correction process, possibly the greatest threat with VarBC in the future is complacency. While the VarBC will always optimise on the basis of information within the assimilation system, we must ensure that we continue monitoring the bias corrections that are being applied to the satellite data (e.g. to ensure that they are physically reasonable and not very different to what other NWP centres apply). To support this, we also plan to automate the monitoring process to detect changes in the bias correction and send alerts when pre-defined thresholds are violated.

We close with two quotations which will either concern or reassure the reader as required.

"Nothing is more fatal to happiness or virtue, than that confidence which flatters us with an opinion of our own strength, and, by assuring us of the power of retreat, precipitates us into hazard" (Samuel Johnson)

"Pessimists have already begun to worry about what is going to replace automation" (John Tudor)

Further reading

Auligné, T. & **A.P. McNally,** 2005: Adaptive bias correction of satellite data at ECMWF. In *Proc. ECMWF/ EUMETSAT NWP-SAF Workshop on Bias Estimation in Data Assimilation,* 8–11 November 2005, ECMWF, Reading, UK, 127–141.

Dee, D.P., 2004: Variational bias correction of radiance data in the ECMWF system. In *Proc. ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP,* 28 June–1 July 2004, ECMWF, Reading, UK, 97–112.

Derber, J.C. & **W.S. Wu,** 1998: The use of TOVS cloud cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.,* **126,** 2287– 2299.

Eyre, J.R., 1992: A bias correction scheme for simulated TOVS brightness temperatures. *ECMWF Tech. Memo No. 186.*

Harris, B. & **G.A. Kelly.,** 2001: A satellite radiance bias correction scheme for radiance assimilation. *Q.J.R. Meteorol. Soc.,* **127,** 1453– 1468.

Kelly, G.A. & **J.F. Flobert,** 1988: Radiance Tuning. *Technical Proceeding of the Fourth International TOVS Study Conference,* 16–22 March 1988, Igls, Austria, 99–117.

Watts, P.D. & **A.P McNally,** 2004: Identification and correction of radiative transfer modelling errors for atmospheric sounders: AIRS and AMSU-A. In *Proc. ECMWF Workshop on the Assimilation of High Spectral Resolution Sounders in NWP,* 28 June–1 July 2004, ECMWF, Reading, UK, 23–37.

 doi:10.21957/mn7643f9s8