Bias Correction for Environmental Monitoring

Richard J. Engelen

ECMWF, Shinfield Park, Reading
RG2 9AX, United Kingdom
richard.engelen@ecmwf.int

1 Introduction

At ECMWF, a 4-dimensional environmental monitoring system is being build by introducing greenhouse gases, reactive gases, and aerosols into the operational 4D-Var data assimilation system. These new variables will be constrained by various satellite observations consisting of both geo-located radiances and level-2 retrieval products. The reactive gases will have an extra complicating factor through the coupling between the data assimilation system and a full-blown version of a chemistry model. The main goal of environmental monitoring is to produce an unbiased analysis rather than an unbiased forecast. This is especially important for the greenhouse gas fields that will be inverted to surface flux distributions. All these new developments require slightly different viewing angles on the concept of bias correction, as will be illustatred with a few examples in the next two sections. The examples are mainly focused on CO_2, because a simple CO_2 data assimilation system already exists at ECMWF. However, most issues equally apply to aerorols and reactive gases as well.

2 Observation bias

Within the GEMS project a wide range of satellite data will be used to constrain the atmospheric distributions of the various trace gases and aerosols, as shown in Table 1.

Table 1: ***Satellites to be used within the GEMS project***

Greenhouse Gases	Reactive Gases	Aerosols
AIRS[a]	SCIAMACHY[c]	MODIS[e]
IASI[a]	SBUV[c]	SAGE[e]
CrIS[a]	OMI[c]	
Mopitt[b]	TOMS[c]	
SCIAMACHY[b]	GOME[c]	
OCO[b]	MIPAS[a]	
GOSAT[b]	MLS[d]	

[a] Infrared observations
[b] Near-infrared observations
[c] Ultraviolet observations
[d] Microwave observations
[e] Visible observations

The instruments observe the atmosphere in various parts of the electromagnetic spectrum and will therefore have different sources of systematic errors. Most common sources of observation bias are spectroscopy er-

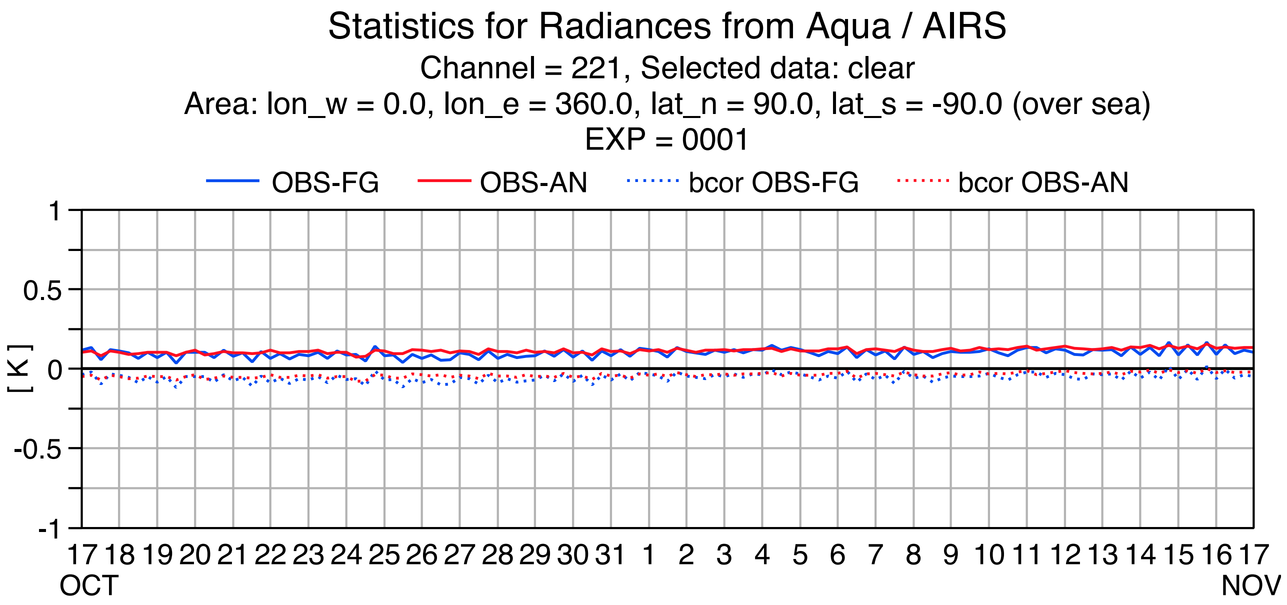

Figure 1: AIRS monitoring.

rors, (undetected) clouds and/or aerosols, surface reflectivity errors, and errors in climatological fields used in stand-alone satellite retrievals. For the purpose of data assimilation we can divide the satellite observations into two distinct categories: radiance observations that are directly assimilated and retrieval products. An example of the first category are the observations from the Advanced Infrared Sounder (AIRS) that are directly assimilated to constrain temperature, humidity, ozone, and CO_2. An example of the second category are the total ozone products from the Total Ozone Mapping Spectrometer (TOMS). Careful monitoring of the radiances and independent validation of the retrieval products will be required to minimize the observation bias.

2.1 Effect of bias correction on the CO_2 estimation from AIRS

A standard method to estimate observational bias in satellite observations is to monitor first-guess departures for a certain period of time, as shown in Figure 1. First-guess departures are calculated as the observed brightness temperatures for a certain spectral channel minus the model simulated brightness temperatures. When we select clear data only, we obtain a representation of the systematic difference between the observations and the model. These differences can be caused by both observation error (including radiative transfer error) and model error. The model error component is often minimized by only gathering statistics close to accurate independent observations (e.g., radiosondes). The systematic differences in Figure 1, which shows the statistics for an AIRS channel sensitive to mid- and upper tropospheric CO_2, are relatively small and constant for this particular month. Therefore, this channel could be easily corrected for the bias. However, the CO_2 signal we are trying to extract is of the same order of magnitude as the bias, as is shown in Figure 2. This means that there is potentially a danger of removing the signal through the bias correction or interpreting a bias as a CO_2 signal, as will be illustrated in the following example.

Figure 3, shows averaged CO_2 results using two different bias correction methods. The CO_2 estimates come from the single column implementation described by *Engelen and McNally* (2005). The figure on the left used a flat (global mean value) bias correction for each individual channel, while the figure on the right used a gamma bias correction (see elsewhere in these proceedings). Both methods did not take CO_2 variations into account. It is immediately clear that large differences can result in the CO_2 estimates from only small differences in the bias correction. Only independent validation data can show which result is closer to the

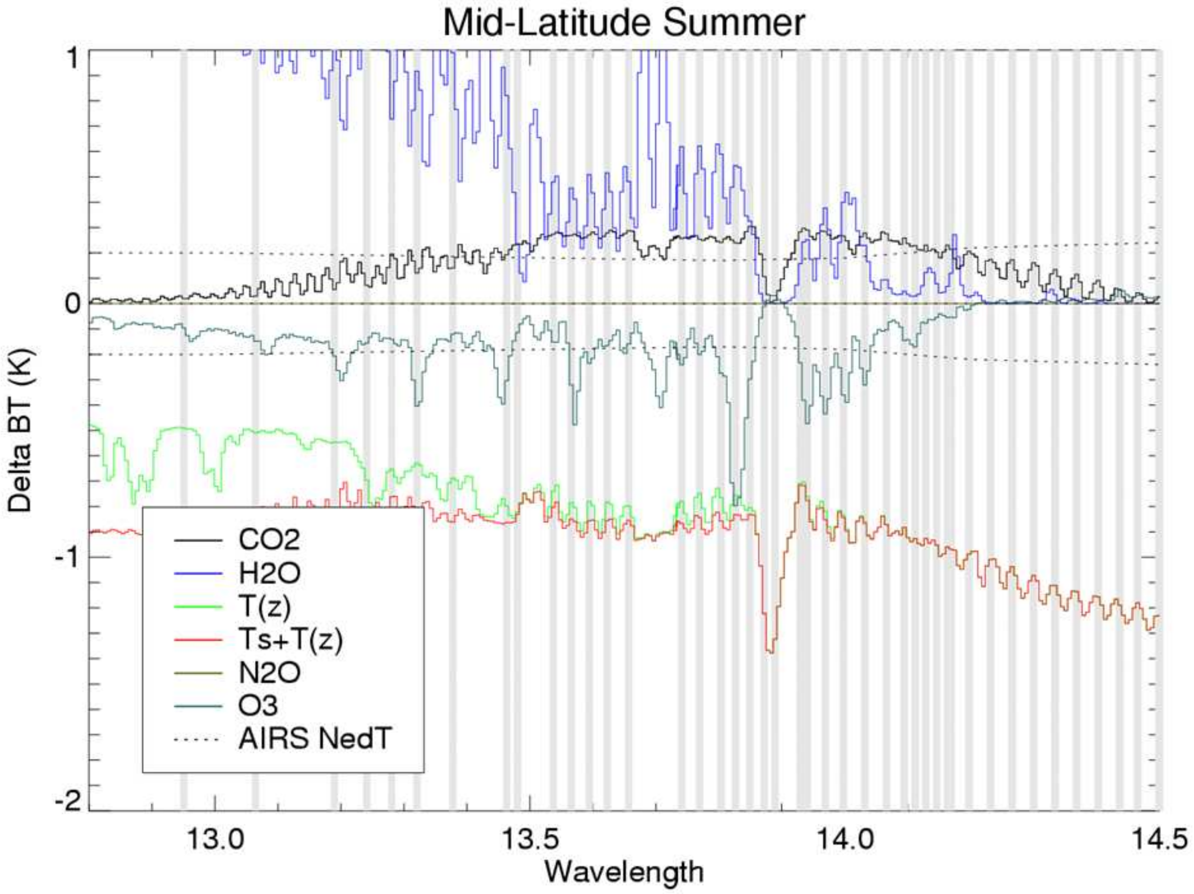

Figure 2: Sensitivity of AIRS longwave channels to background error perturbations of temperature, water vapour, and ozone as well as to seasonal perturbations of CO_2. Vertical grey lines indicate AIRS channels that are being used in the operational data analysis system.

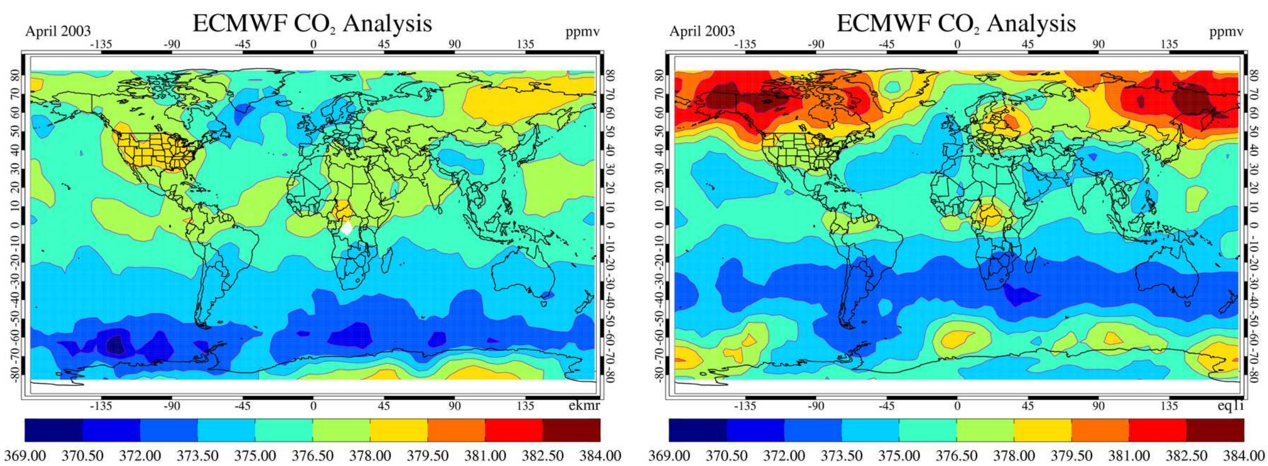

Figure 3: AIRS CO_2 results using a flat bias correction (left) and a gamma bias correction (right).

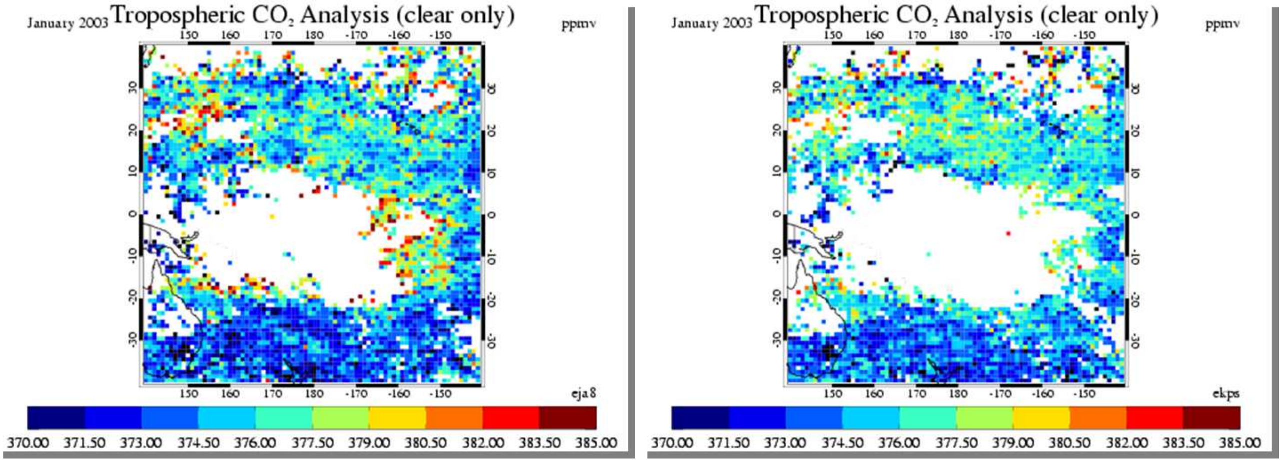

Figure 4: CO_2 results over the west Pacific using a cloud detection algorithm that fails to detect very thin cirrus (left) and using a corrected cloud detection algorithm.

truth. However, the amount of accurate validation data is often limited. Although surface networks are still expanding, they are often not sufficient, especially when considering that only aircraft profile data can be used for proper bias correction.

2.2 Cloud detection bias

Another bias in the CO_2 estimates is caused by problems with the cloud detection for the AIRS radiances. Although the cloud detection works very well in most cases (*McNally and Watts*, 2003), the assumption that there are no systematic errors proved to be a problem in the tropical convective areas. The detection of thin cirrus (thin enough to still show the atmosphere and/or surface underneath it) was compromised by large systematic errors in the background water vapour profiles that affect the lowest peaking channels in the long-wave band (sensitive to water vapour). This meant that cloud-affected channels were slipping through at the edges of tropical convective systems and creating anomalous high CO_2 estimates. The effect is shown in Figure 4 where the left panel shows the results using the original cloud detection algorithm and the right panel shows the results using a corrected cloud detection algorithm. The problem with these kind of effects is that it is often not clear if the anomalous CO_2 values represent the real variability or if they are caused by problems in either the bias correction or the cloud detection.

3 Model bias

Systematic model errors are an even harder problem to deal with. For the greenhouse gases, for example, we know that the forecast model is biased, because we do not have perfect knowledge of the surface fluxes. We are actually trying to estimate these surface fluxes from the analyses through off-line inversion models (e.g., *Gurney et al.*, 2002). This is illustrated by Figure 5 that shows the north-south gradient of zonal mean CO_2 at the surface modeled by the ECMWF model and observed by NOAA/CMDL surface flasks from the GLOBALVIEW data set (*GLOBALVIEW-CO_2*, 2003) on the left. The right panel shows the northen hemisphere seasonal cycle for the ECMWF model and the surface flasks. Although the model compares generally quite well with the observations, there are systematic differences. It is therefore crucial to minimize the effect of the model bias on the analysis by optimally using the information of the obervations. It is also not clear how well

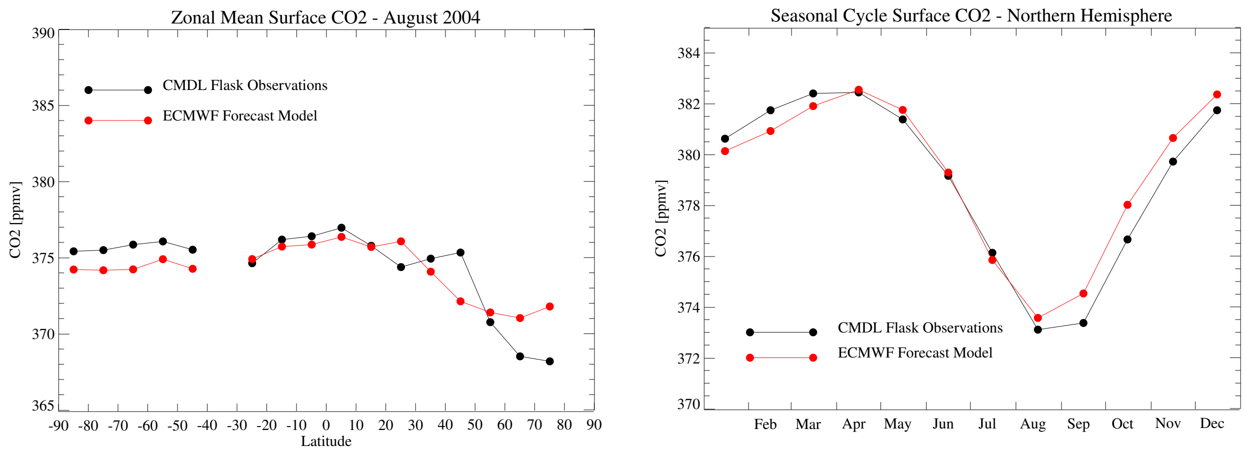

Figure 5: Comparisons between the ECMWF CO_2 model and surface flask observations from the NOAA/CMDL network. Zonal mean values for August on the left and northern hemisphere seasonal cycle on the right.

tracers are transported with the vertical diffusion and convection parameterisations. The effect of these biases on the area of the atmosphere where the current AIRS observations are sensitive to CO_2 is even less certain. If the analysis CO_2 fields, used as substitute for individual observations in the flux inversions, are biased due to forecast model bias, flux inversions will produce the wrong fluxes. Therefore, model bias should be corrected as wll as possible and its effect on the analysis should be minimal. This all requires careful inspection of the analysis results to spot any significant model biases. For aerosols and reactive gases, where we do not invert the analysis fields to estimate surface fluxes, this is all less critical. On the other hand, the model errors are likely to be much larger. For aersols, assumptions need to be made to describe the size distributions, and for the reactive gases we will have a complicated coupling between the assimilation model and detailed atmospheric chemistry models.

Acknowlegdements

The author would like to thank Soumia Serrar, Antje Dethof, Angela Benedetti, and the rest of the GEMS team for their contributions to this work. The work is funded by the EU FP6 project GEMS.

References

Engelen, R. J., and A. P. McNally (2005), Estimating atmospheric CO_2 from advanced infrared satellite radiances within an operational four-dimensional variational (4D-Var) data assimilation system: Results and validation, *J. Geophys. Res.*, *110*, D18305, doi:10.1029/2005JD005982.

GLOBALVIEW-CO_2 (2003), Cooperative Atmospheric Data Integration Project - Carbon Dioxide, CD-ROM, NOAA CMDL, Boulder, Colorado [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW].

Gurney, K. R., et al. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, *415*, 626–630.

McNally, A. P., and P. D. Watts (2003), A cloud detection algorithm for high-spectral-resolution infrared sounders, *Q. J. R. Met. Soc.*, *129*, 3411-3423, doi:10.1256/qj.02.208.