Chemistry data assimilation validation using independent observations

Angela Benedetti[1], Richard Engelen[1], Johannes Flemming[1], Harald Flentje[2], Nicolas Huneeus[3], Antje Inness[1], Luke Jones[1], Eleni Katragkou[4], Julia Marshall[5], Jean-Jacques Morcrette[1], Carlos Ordóñez[6], and Michael Schulz[3]

(1) ECMWF, Shinfield Park, Reading
RG2 9AX, United Kingdom
Corresponding author: A.Benedetti@ecmwf.int
(2) Deutscher Wetterdienst (DWD)
Meteorologisches Observatorium Hohenpeienberg (FEHp)
82383 Hohenpeissenberg, Germany
(3) Laboratoire des Sciences du Climat et de lEnvironnement
91191 Gif-sur-Yvette Cedex, France
(4) National Kapodistrian University of Athens - Aristotle University of Thessaloniki
Athens, Greece
(5) Max-Planck-Institut für Biogeochemie
07745 Jena, Germany
(6) UK Met Office,
Exeter, United Kingdom

ABSTRACT

An analysis system for greenhouse gases, reactive gases and aerosols has been developed at the European Centre for Medium–Range Weather Forecasts, under the Global and regional Earth-system Monitoring using Satellite and in-situ data (GEMS) project. The GEMS modelling and analysis systems are fully integrated in the operational four–dimensional assimilation system. Their purpose is to produce global forecasts and reanalyses of atmospheric composition using satellite data. A multi–year reanalysis has been conducted for 2003-2007. Analysed fields include carbon dioxide, methane, carbon monoxide, ozone, formaldehyde, nitrogen oxides and aerosol optical depth. These fields are being evaluated using independent observations. This study presents a collection of results from the validation activities performed both by ECMWF and by various European institutes, which were part of the GEMS consortium.

1 Introduction: the GEMS and MACC projects

The Global and regional Earth-system Monitoring using Satellite and in-situ data (GEMS) project which finished in May 2009 included thirty-two European partners with expertise in various aspects of atmospheric composition monitoring. GEMS was part of the Global Monitoring for Environment and Security (GMES) initiative and was established under European Commission funding in 2005 to create an assimilation and forecasting system for monitoring aerosols, greenhouse gases and reactive gases, at global and regional scales, through exploitation of satellite and in–situ data (Hollingsworth et al., 2008). An important component of GEMS has been the forecasting of regional air quality at the European scale. This is performed with an ensemble of air quality models from the participating institutes. Boundary conditions for the high–resolution models are provided by the global model.

The basis for the global GEMS forecast and analysis schemes is the operational ECMWF Integrated Forecasting System (IFS) that includes the incremental four–dimensional variational analysis system. The system has been extended to include new prognostic variables for atmospheric tracers (i.e. gases and aerosols). A coupled chemical transport model is also part of the GEMS system and provides tendencies for the chemically–active species which are present in the model (Flemming et al., 2009).

The follow–up to GEMS is the Monitoring of Atmospheric Composition and Climate (MACC) project whose objectives are largely the same as GEMS with a greater emphasis on user–oriented applications.

2 Brief description of the GEMS-ECMWF integrated analysis system

The main characteristics of the integrated GEMS–ECMWF analysis system are:

- T159 (120km) horizontal resolution,
- 60 vertical layers; top level at 0.1 hPa,
- model physics based on ECMWF model cycle 32r3,
- Four-dimensional variational (4DVAR) analysis using a 12-hour time window,
- Wavelet-based background error covariances for all tracers, and background error statistics computed using the NMC method (Parrish and Derber , 1992).

Control variables for the system include: carbon dioxide, methane, carbon monoxide, ozone, formaldehyde, nitrogen oxides and total aerosol mass. Specific observational operators have been developed for the various constituents, taking into consideration the maturity of the observing systems and the availability of observations suitable for assimilation. For example, the analysis of CO_2 is driven by AIRS radiance observations, whereas the reactive gases analysis is driven by retrievals of total column integrated amounts from various space-borne sensors such as MIPAS, OMI, SCHIAMACHY, SBUV, MLS, GOME, MOPITT. For the aerosol analysis, retrieved aerosol optical depths from the MODIS sensor are used. More details on the three sub-components of the GEMS system can be found in Engelen et al. (2009); Inness et al. (2009); Benedetti et al. (2009).

The GEMS reanalysis of the years 2003-2007 was completed in March 2009. As of July 2009, the GEMS reanalysis has been restarted to cover 2008 and 2009, and it has now reached November 2008.

3 Validation metrics

Although there were no prescribed standards for the validation exercise, for the sake of consistency most validating groups chose common validation metrics as well as similar methods to display results. A few of the metrics used are listed below

- Modified normalized mean bias, $B = \left(\frac{2}{N}\right) \sum_{i=1}^{N} \frac{f_i - o_i}{f_i + o_i}$
- Correlation coefficient, $r = \frac{\frac{1}{N} \sum_{i=1}^{N} (f_i - \overline{f})(o_i - \overline{o})}{\sigma_f \sigma_o}$
- Normalized Median Bias, $NMedB = Median \sum_{i=1}^{region} \frac{(f_i - o_i)}{Median(o_i)}$

For the visualization the following methods were used

- Taylor diagrams (standard deviation and correlation)
- Scatterplots (bias and correlation)
- Line plots (time series of bias and Root Mean Square error)

These can be considered to provide quantitative information about the quality of the analysis with respect to the independent observations. Qualitative information is also displayed using profile/cross section plots and maps.

4 Validation of the greenhouse gas analysis

The assimilation system for greenhouse gases include CO_2 and methane (CH_4) as prognostic variables and control variables. The observations assimilated consist of AIRS radiances for CO_2 and SCHIAMACHY retrievals for CH_4. For the CO_2 analysis, the verification observations are ground-based flask measurements and aircraft observations as shown in figure 1 while IASI CH_4 retrievals, that were not assimilated, constitute the validating dataset for the CH_4 analysis.

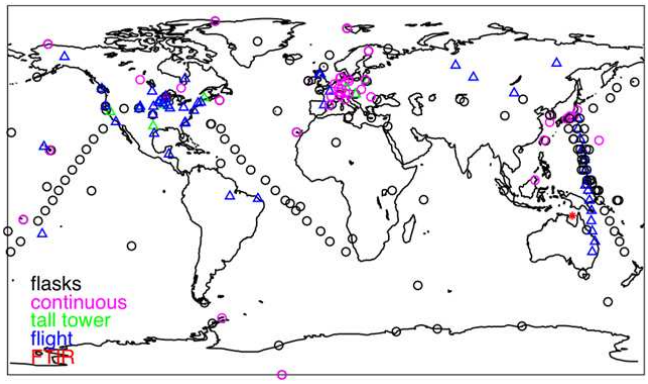

Figure 1: Map of observations for validating analysed CO_2.

4.1 Verification of the CO_2 analysis

Four-dimensional analysed CO_2 fields were sub-sampled to match available surface, tower, ship-based, and flight data The resultant time series were compared to available observations. Statistical results were summarized by means of a Taylor diagram (figure 2), and by using maps of modified mean normalized bias and correlation coefficient (figure 3).

The performance of the CO_2 analysis is detailed in the Taylor diagram of figure 2 which shows only a few stations for which the standard deviation is large and the correlation is poor. Otherwise most stations crowd around the reference point, indicating a good agreement of the modelled CO_2 with the observations. The Taylor diagram does not show, by construction, the analysis bias which is instead illustrated in the top panel of figure 3. The figure shows that the analysed CO_2 has up to a 10% positive bias over Europe. In general, the Southern Hemisphere is well-constrained, while a slightly positive tendency can be seen in the Northern Hemisphere. This seems related to a seasonal cycle which is too weak. The bottom panel of the same figure shows a map of the correlation coefficient. Remote stations show good agreement while poor correlation are seen over highly populated regions with heterogeneous fluxes such as Europe.

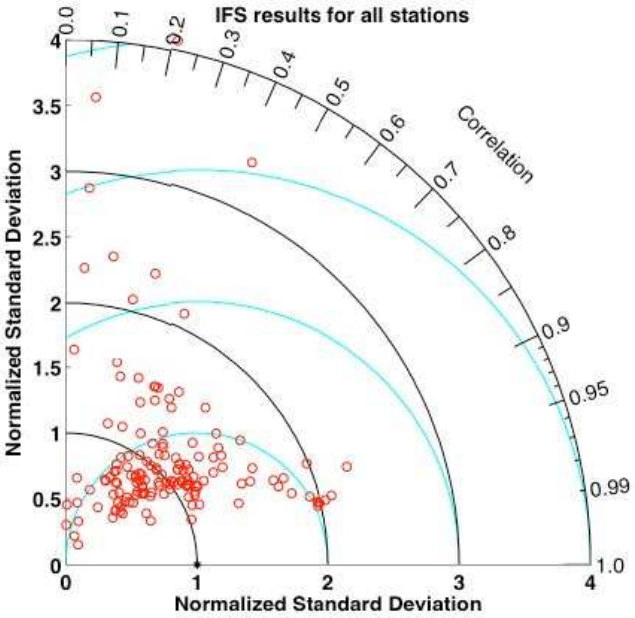

Figure 2: Taylor diagram for analysed CO_2 with respect to surface observations.

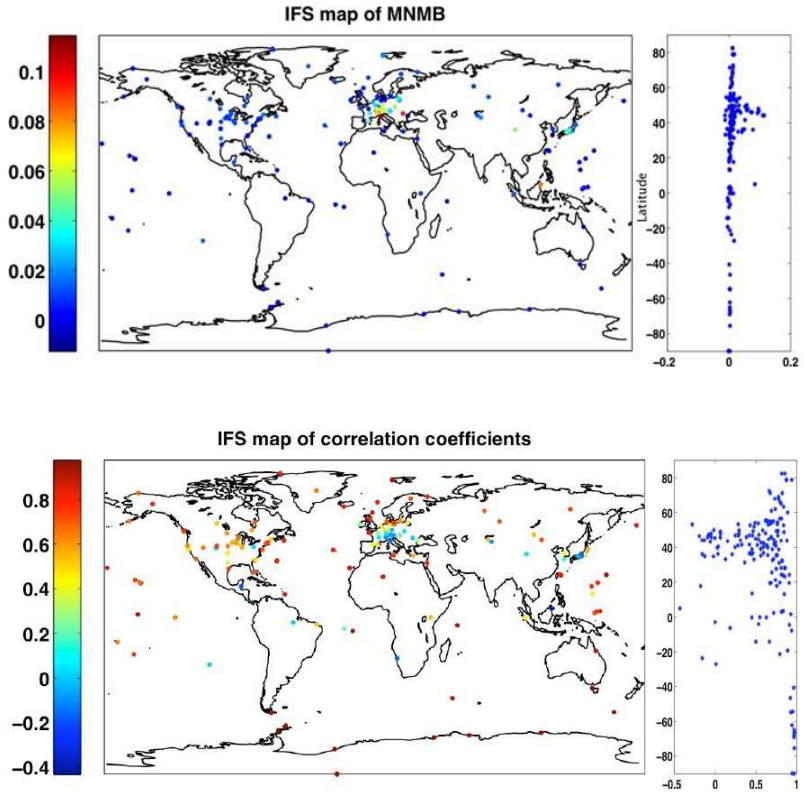

Figure 3: Top panel: modified normalized mean bias of analysed CO_2 with respect to ground–based data. Bottom panel: correlation coefficient.

4.1.1 Comparison with CO_2 observations from the Mauna Loa observatory

Figure 4 shows the performance of the analysis in comparison with observations of CO_2 surface concentration taken at the Carbon Cycle In Situ Observatory of Mauna Loa (19 N, 155W, 3397 m ASL). The plot (which can easily be considered the most infamous plot of the century) shows the rapid increase of CO_2 over the last decade. The analysis does a good job in reproducing such trend especially in more recent years, possibly due to general improvements in the observing system (for example the inclusion of GPS radio occultation data) leading to a better CO_2 analysis.

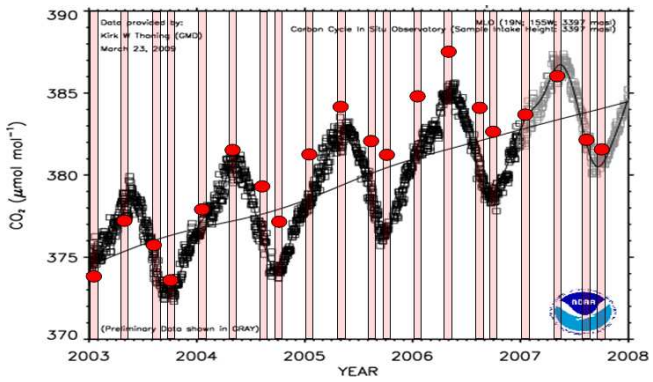

Figure 4: Comparisons of CO_2 concentration observed at the Mauna Loa observatory (black) with the GEMS analysis (red).

4.2 Verification of the methane analysis

The analysed methane fields were sub-sampled in time and space to match individual retrievals from an independent satellite The appropriate weighting function was applied (not shown). Monthly mean maps were compared for spatial and temporal correlation (figure 5). Qualitatively, the analysed methane compares well with IASI retrievals, but tends to be higher on average, especially in the Indian Ocean and Indian Sub-continent regions.

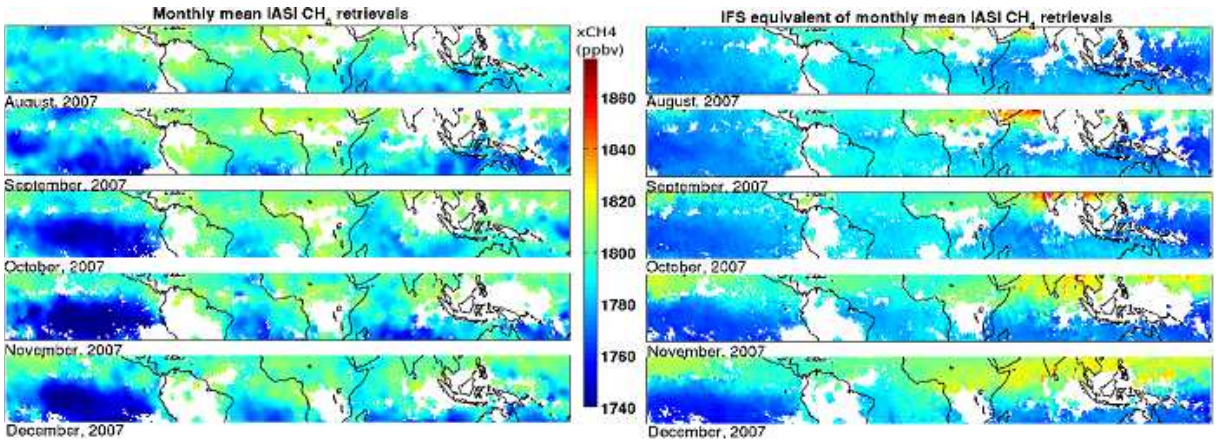

Figure 5: Left: methane fields retrieved from IASI observations; Right: methane fields from the GEMS reanalysis.

5 Validation of the aerosol analysis

Aerosol prognostic variables include 3 bins for desert dust, 3 bins for sea-salt, hydrophobic and hydrophilic organic matter, hydrophobic and hydrophilic black carbon, and sulphate. The control variable is formulated in terms of the total aerosol mixing ratio. Assimilated observations are the MODIS Aerosol Optical Depths (AODs) at 550 nm over land and ocean. Observation errors over ocean are prescribed as functions of the satellite scattering angle. Errors over land are assigned as 50% of the optical depth value. For more details see Benedetti et al. (2009).

Validation datasets used are optical depths from AERONET, AEROCE (U. of Miami), and compilation datasets. These are shown in figure 6.

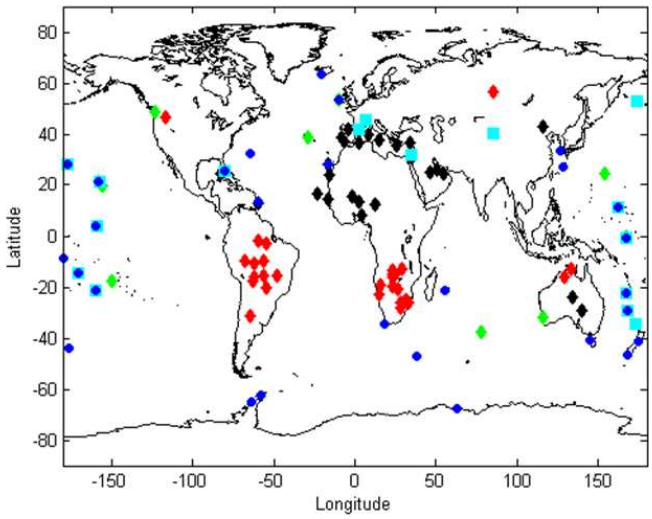

Figure 6: Map of observations for the validation of the aerosol analysis. Optical Depth 550 nm & Angstrm Exponent (AERONET) (diamonds); Surface Concentration DD (AEROCE, U. de Miami) (stars, not discussed here) and Total Deposition by Ginoux et al, 2001 (squares, not discussed). AERONET Stations where a species dominates the total optical depth for at least 4 months are marked with the following symbols: black diamonds = desert dust stations; red diamonds = biomass burning stations, green diamonds = sea salt stations.

The verification has focused so far on the AOD at 550 and 865 nm and the Angström exponent, α, defined from the relationship $\frac{\tau_1}{\tau_2} = \left(\frac{\lambda_1}{\lambda_2}\right)^{-\alpha}$. A validation effort is currently under way to validate also the surface concentrations of the aerosol species.

5.1 Observation statistics from the analysis of MODIS data

The aerosol optical depth from the analysis for the whole month of May 2003 was used to investigate the analysis performance with respect to the assimilated observations. In a successful analysis the departures should always be smaller than the first guess departures (and the analysis should better match the observations, at least in a statistical sense). Figure 7 shows scatterplots of assimilated aerosol observations versus first guess (top left) and analysis (top right). By visual inspection, it is apparent that the scatter in the analysis is smaller than in the first guess. The root mean square error with respect to the MODIS data is lower for the analysis (0.122) than for the first guess (0.168) while the correlation coefficient is higher for the analysis (0.888) than for the first guess (0.757), indicating a good performance of the analysis. However, while we did not expect the analysis to improve on the first guess biases, it was surprising to notice that the analysis effectively has a larger bias than the first guess. The distribution appears to be skewed and it is evident from the shape of the scatterplot that the analysis is more efficient

in increasing low values of optical depth than in reducing high values. This asymmetrical behaviour can be partially attributed to the prescription of the observation error. In the current analysis configuration, the error on the observed AOD is prescribed as a percentage of the acutal AOD value. As a result, large values of AOD have correspondingly higher errors, although these values may actually be observed better than low values of AOD, when the contrast with the surface reflectance is large enough. We repeated one month analysis with the observation errors for optical depths larger than 1 capped to 0.4 and were more successful in fitting the observations with a reduction of bias in the analysis to 0.009 and of RMS to 0.092, and an increase of the correlation coefficient to 0.927 (see bottom panels of Figure 7). The first guess fit is also improved considerably.

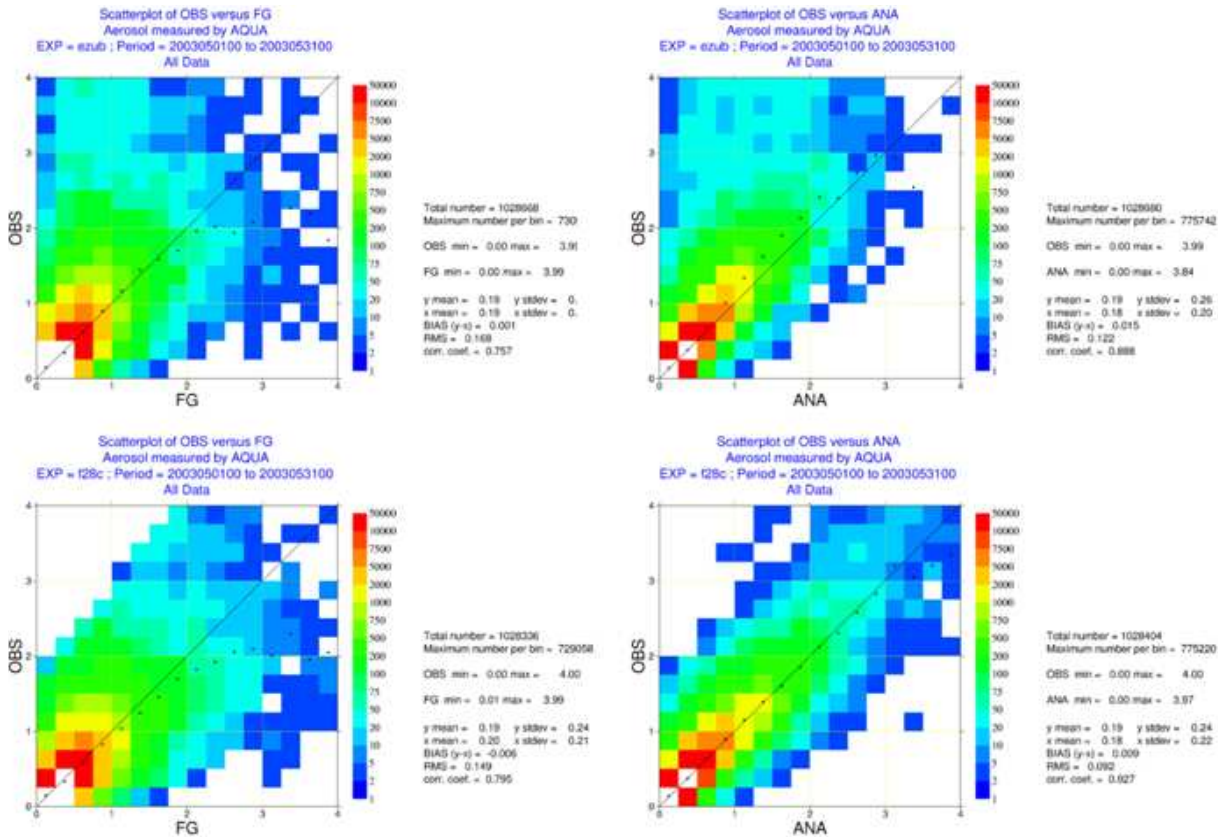

Figure 7:

5.2 Validation of the aerosol analysis using AERONET data

Global statistics show that the analysis (ASSIM) has a positive bias with respect to the AERONET data which is larger than that of the forecast without assimilation of MODIS data (FCST) while having much higher correlation and lower RMS with respect to the same dataset. This refers to tha assimilation system where the observation error was prescribed as a percentage of the actual AOD value and highlights the problems with this error specification. However, there could also be a modeling component to the analysis bias. This is under investigation.

Good performance of the analysis is shown in terms of seasonal and spatial correlation. Results are summarized for the whole year 2003 and 2004 in table 1. For comparison purposes, the MODIS Aqua values are also shown.

	FCST 2003	FCST 2004	ASSIM 2003	ASSIM 2004	MODIS Aqua 2003	MODIS Aqua 2004
AERONET AOD	0.22	0.22	0.22	0.22	0.22	0.22
# N months	1125	1422	1225	1422	1143	1292
AOD	0.24	0.26	0.27	0.27	0.20	0.20
Corr	0.68	0.69	0.82	0.82	0.79	0.78
RMS	0.13	0.14	0.11	0.12	0.11	0.12
Std Dev	0.76	0.73	0.81	0.79	0.89	0.93
Seasonal r	0.75	0.76	0.80	0.80	0.81	0.79
Spatial r	0.71	0.73	0.78	0.81	0.80	0.81

Table 1: Global statistical comparison of aerosol free-running forecast and analysis with AERONET data.

Figure 8 shows scatterplots and pdfs of optical depth at 865 nm which compare the global AERONET data with the free-running forecast (no assimilation) and the analysis. The correlation between the observations and the analysis is greater, but the bias in the analysis is larger than that of the forecast. As highlighted when discussing the analysis fit to the assimilatedassimilated observations (see subsection 5.1), the analysis does not seem to be able to correct the large AOD values.

Figure 9 shows the Angström coefficient for the free-running forecast and the analysis as they compare with AERONET data. It can be seen that there is no improvement between the free–running experiment and the analysis. In general there is an over-estimation of the amount of coarse aerosol and underestimation of fine aerosol in the model.

Figure 10 summarizes results with the aid of a Taylor diagram. The free–running forecast without assimilation is marked in black, while the analysis is in red. The following fields are plotted: aerosol optical depth at 550 nm; aerosol optical depth at 865 nm; Angström exponent fine mode aerosol optical depth at 550 nm; coarse mode aerosol optical depth at 550 nm.

Preliminary conclusions from the comparison with AERONET observations can be drawn as follows:

- Significant improvement in column integrated aerosol variables in terms of correlation and standard deviation.
- A positive bias is present in the analysis.
- Assimilation of AOD at 550 nm improves also AOD at 865 nm which is not assimilated.
- Improvement of AOD at 550 and 865 nm does not translates into improvement of Angström exponent suggesting that assimilation acts on correcting total aerosol burden rather than size distribution.
- Overestimation of the Angström exponent for coarse aerosols indicates smaller particles in the model.
- Too much fine mode sea salt represented in the model (not shown).
- Not enough Desert Dust is emitted and too much fine Desert Dust is transported far off source regions in the forecast model (not shown).

5.3 Another independent comparison with AERONET data

Another independent comparison with AERONET data was performed at ECMWF. The aerosol optical depth data used in this comparison are the Level 2.0 (cloud-screened and quality-assured) product. The

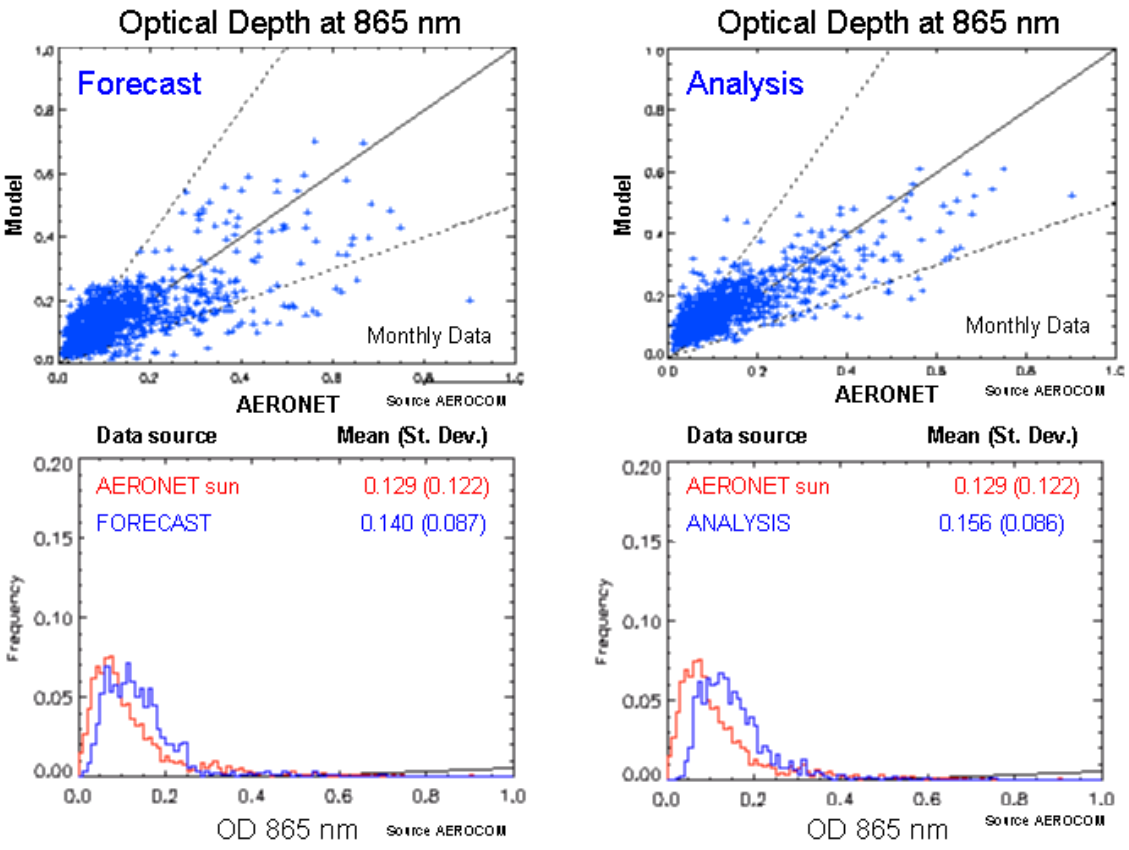

Figure 8: Optical depth at 865 nm: free-running forecast (left) and analysis (right). AERONET sunphotometer data are shown in red in the pdf plot; model values are in blue.

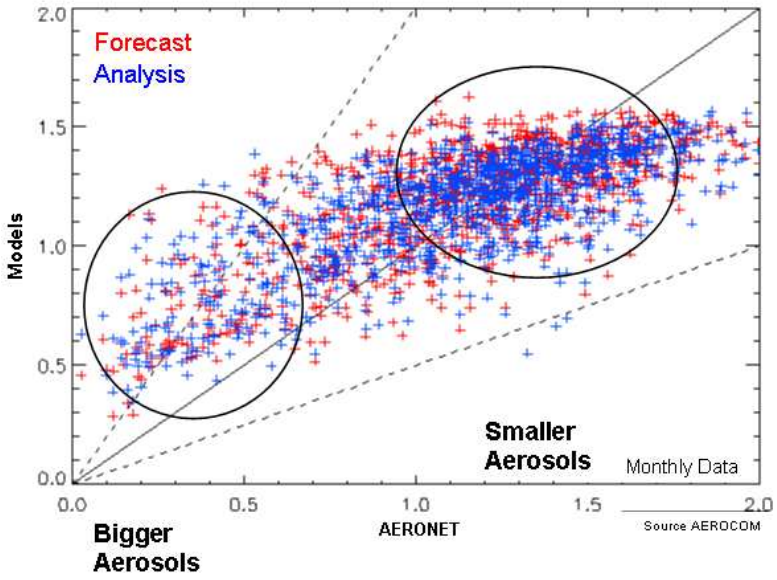

Figure 9: Scatterplot of Angström coefficient for the free-running forecast (red) and the analysis (blue) compared with AERONET data.

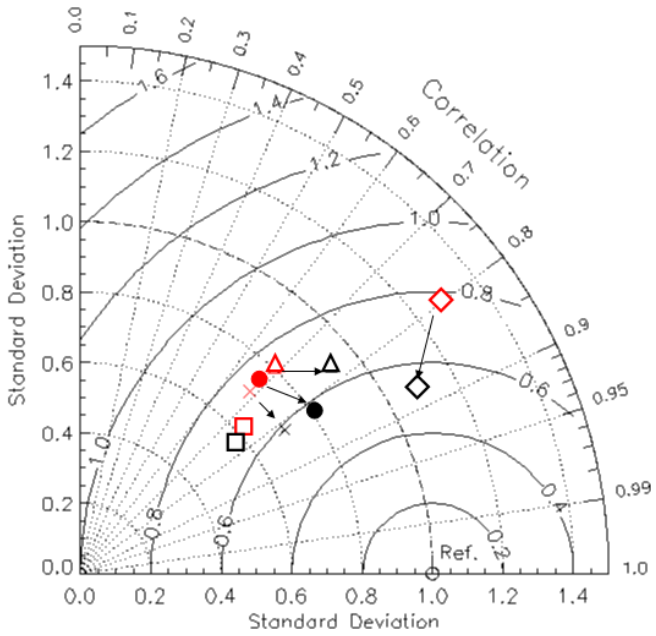

Figure 10: Taylor diagram for model AOD and Angström coefficient with respecto to AERONET observations. Aerosol optical depth at 550 nm is shown with full circles; aerosol optical depth at 865 nm is shown with crosses; the Angström exponent is marked with squares, the fine mode aerosol optical depth at 550 nm with triangles; and the coarse mode aerosol optical depth at 550 nm is indicated with diamonds.

site selection was thinned using an algorithm which looped through all available sites, checking each for proximity to others. If two sites were found within 700 km of each other, then the site with greater data availability (measured as the number of 6 hour periods with at least one observation at 500 nm during January 2003) was kept and the other discarded. This resulted in a selection of 41 stations shown in figure 11.

Figure 12 shows some comparisons with AERONET independent data for the month of May 2003. The AODs from the model are averages over 6 hours, whereas the AERONET observations are instantaneous. To make them comparable, the AERONET observations are averaged over the same period. Because the observations are unevenly spaced in time, a weighted mean is computed in such a way that it is equal to the mean of the series of straight lines that join neighbouring observations over the period. Forecast AODs from the free–running experiment and the analysis are bilinearly interpolated to the observation location in space.

The analysis is shown in red and the free–running forecast in blue. Both plots show that the analysis is on average closer to the AERONET observations displaying a lower bias (left panel) and RMS error (rigth panel) than the forecast.

From this comparison with AERONET data it appears that the validation outcome is subject to the choice of the validating dataset (or a subset of). This highlights the fact that the quality of the analysis varies on a spatial, possibly regional scale, depending on the dominating aerosol species. Although in apparent contradiction with the results displayed in table 1, it shows the fact that the global spatial redistribution of AOD operated through the assimilation of MODIS data into the model was successful. It also emphasizes that global averages, although very useful for a quick assessment of the analysis quality, may hide some important features. The issue of the global analysis bias is, however, still to be addressed.

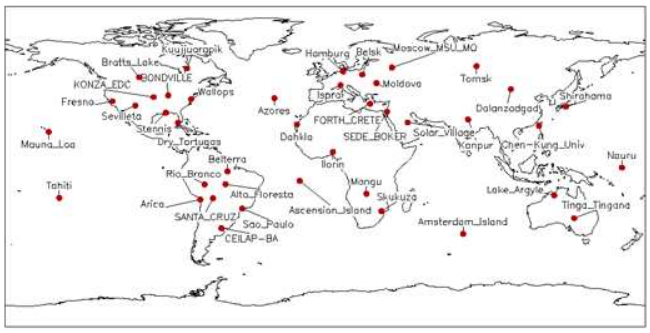

Figure 11: Map of the AERONET stations used for the verification of the aerosol analysis performed at ECMWF

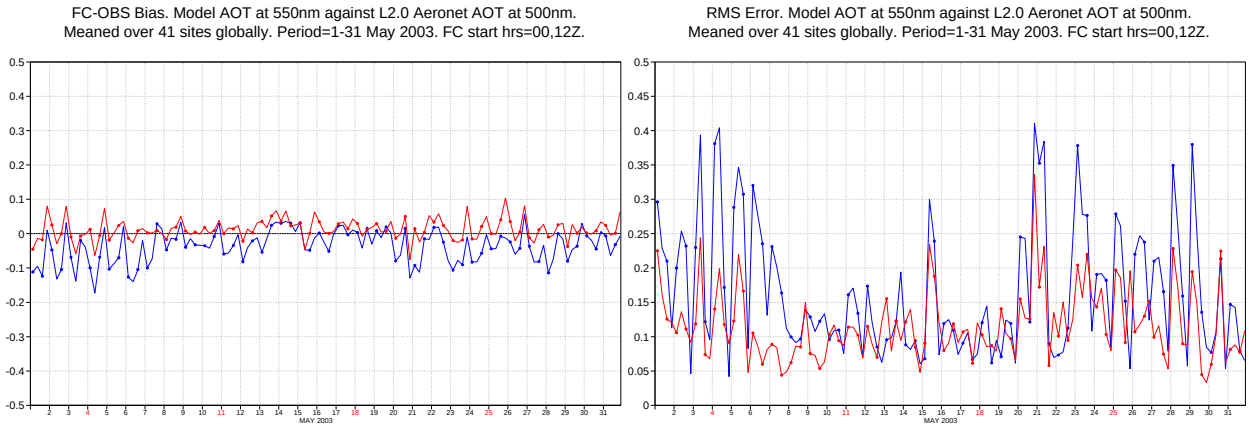

Figure 12: Bias (left) and RMS (right) of the AOD at 550 nm from the free–running forecast (blue) and analysis (red) with respect to AERONET ground–based observations at 500 nm for May 2003.

5.4 Case study: Saharan dust outbreak of March 2004

To further assess the performance of the analysis we looked at a case study relative to a major Saharan dust storm recorded in early March 2004. The storm was detected by several satellite sensors and ground–based sites. Very large values of AOD were recorded. Figure 13 shows comparisons between AODs from the free-running model and the analysis compared to MODIS observations for 6 March 2004. The shape of the dust outflow is well represented in both free–running model and analysis, but the magnitude of the AODs is much larger in the latter in better agreement with the observations. This is also confirmed by looking at the AERONET data at three key stations (see Figure 14). The peaks shown in the AERONET data are well captured by the analysis, despite the lack of MODIS data over the in-land desert sites.

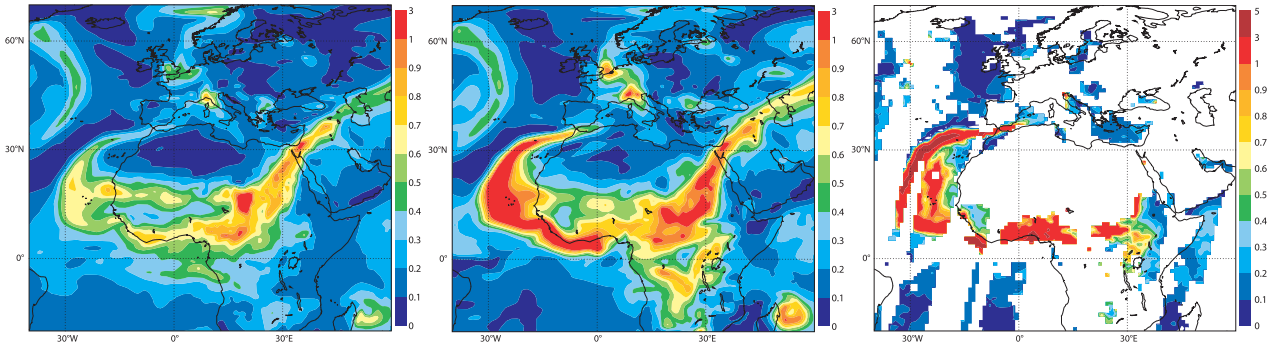

Figure 13: 6th March 2004 Saharan dust outbreak: comparisons of free-running model and analysis 550 nm AODs with MODIS (assimilated) observations: (a) free running model ; (b) analysis ; and (c) MODIS observations.

5.5 Vertical profiles comparison using CALIPSO data

Data from CALIPSO were used to qualitatively assess the vertical distribution of the aerosol in the analysis. Generally, a good agreement is achieved on the vertical (see figure 15) but there is no improvement with respect to a forecast without assimilation of AOD observations. This is not unexpected since the AOD observations cannot constrain the vertical profile of extinction but only the model total optical depth.

It appears that too much aerosol is present in the upper troposphere in the analysis. This is likely to depend on interaction between convection/vertical diffusion and the aerosol transport. We plan to compare extinction profiles and obtain more quantitative profile information.

6 Validation of the reactive gas analysis

The reactive gases included in the GEMS analysis are ozone (O_3), carbon monoxide (CO), nitrogen oxides (NOx), and Formaldehyde. The chemical model MOZART is coupled with IFS which provides the meteorological forcing to the CTM (Flemming et al., 2009). Chemical tendencies are provided to IFS every hour.

Observations used for ozone and carbon monoxide are shown in figure 16 along with the timeline of usage.

The verifying observations for the reactive gases analysis are provided by TOMS, SCHIAMACHY,GAW

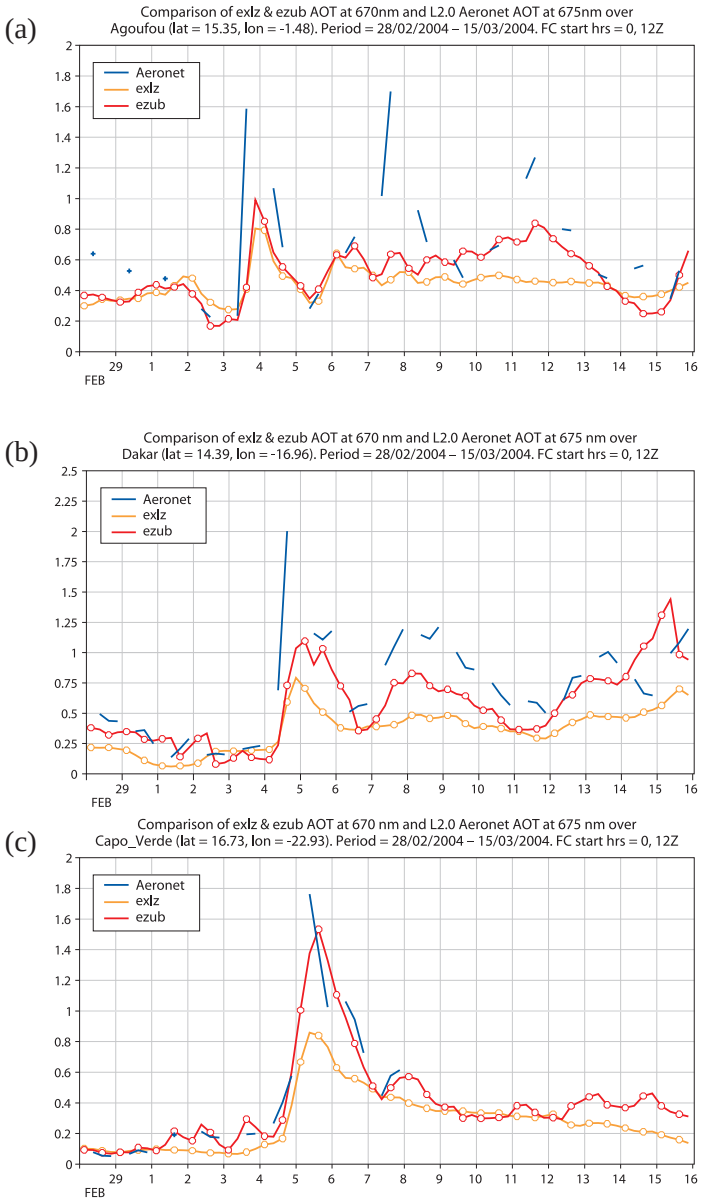

Figure 14: Comparisons of analysis aerosol optical depth at 675 nm with AERONET observations for the Saharan dust outbreak of March 2004: (a) Agoufou (Mali) ; (b) Dakar (Senegal); and (c) Cape Verde. AERONET data are shown here in light blue, the analysis in red (experiment ezub) and the free–running forecast in dark yellow (experiment exlz).

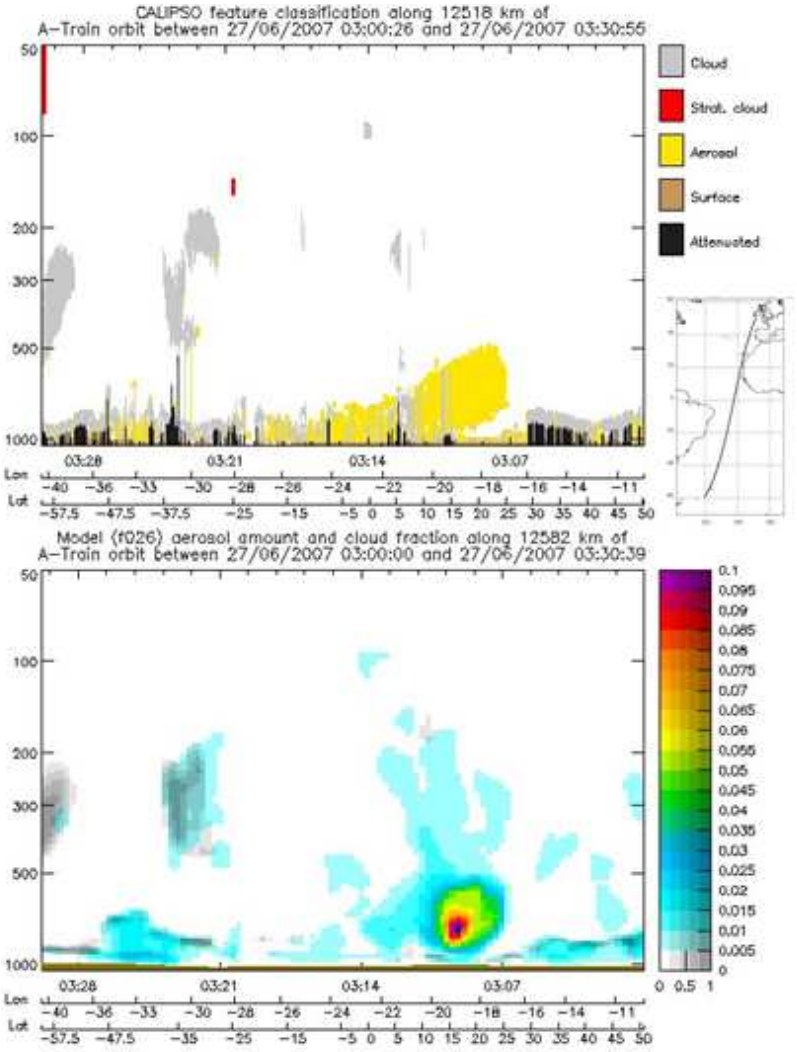

Figure 15: Qualitative comparison of aerosol occurence from the CALIPSO lidar (top panel) with the analysis fields from the GEMS reanalysis (bottom panel).

surface O_3 and CO, and MOZAIC profiles.

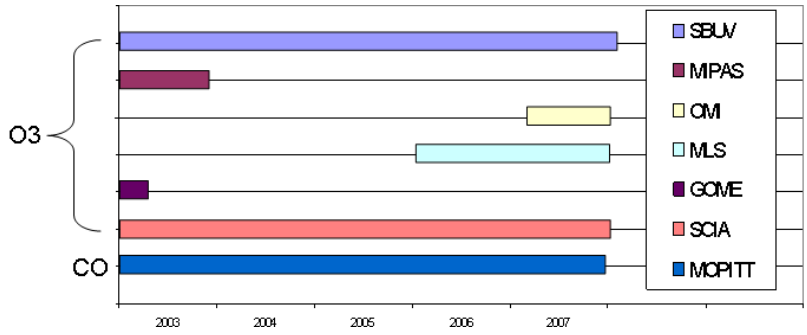

Figure 16: Observation usage in the O_3 and CO GEMS analyses.

6.1 Verification of the ozone hole prediction

Time series of zonal mean total column ozone are shown in figure 17. The top panel shows TOMS data, the middle panel shows the analysed total column ozone from the assimilation run (assimilating SCIAMACHY, MIPAS, GOME, SBUV), and the bottom panel presents the total column ozone field from a control run in which no ozone data were assimilated. Plots show improved agreement of analysed ozone field with independent TOMS data in the assimilation run. The ozone hole is not deep enough in the control while it has a reasonable extent in the analysis.

Figure 18 shows an ozone cross section from 40S across the South Pole back to 40S along 35E from the assimilation run and the control run on 4 October 2003, 12z when an ozone hole was observed. Ozone profiles at the Belgrano station (78S, 35E) from an ozone sonde launched on 4 October 2003 are also shown. Profile and cross section show again the lack of ozone hole in the control.

6.2 Verification of CO profiles with MOZAIC observations

During summer 2003 unusually large values of chemical tracers were observed due to the extremely hot conditions over most of Europe. Data from MOZAIC were used to verify predictions of CO from several models before, during and after the heat wave. Models involved in the comparisons were MOZART, at different resolutions and coupled with IFS, MOCAGE, and TM5, at two different resolutions. The analysis from the GEMS-GRG coupled system (MOZART+IFS) was also included in the comparisons. Figure 19 shows profiles of CO from the different models as they compare with the MOZAIC observations and the Modified Normalized Mean Bias. Most of the models do not reproduce well the observed CO profile close to the surface, and show a large bias in the lower troposphere. Better results are achieved in the middle and upper troposphere, especially by the run with assimilated CO MOPITT data.

6.3 Verification of surface ozone using GAW data

Verification of surface ozone and carbon monoxide was performed for the whole 2003 using data from the GAW network. Sites representative of different regions around the globe were included in the comparison with analysis data. The normalized median bias was used to evaluate the results (see figure 20). Overall both the model run without assimilation and the analysis show large biases at the surface for all regions. Better performance of the analysis with respect to the simulation without assimilation is shown in the CO field.

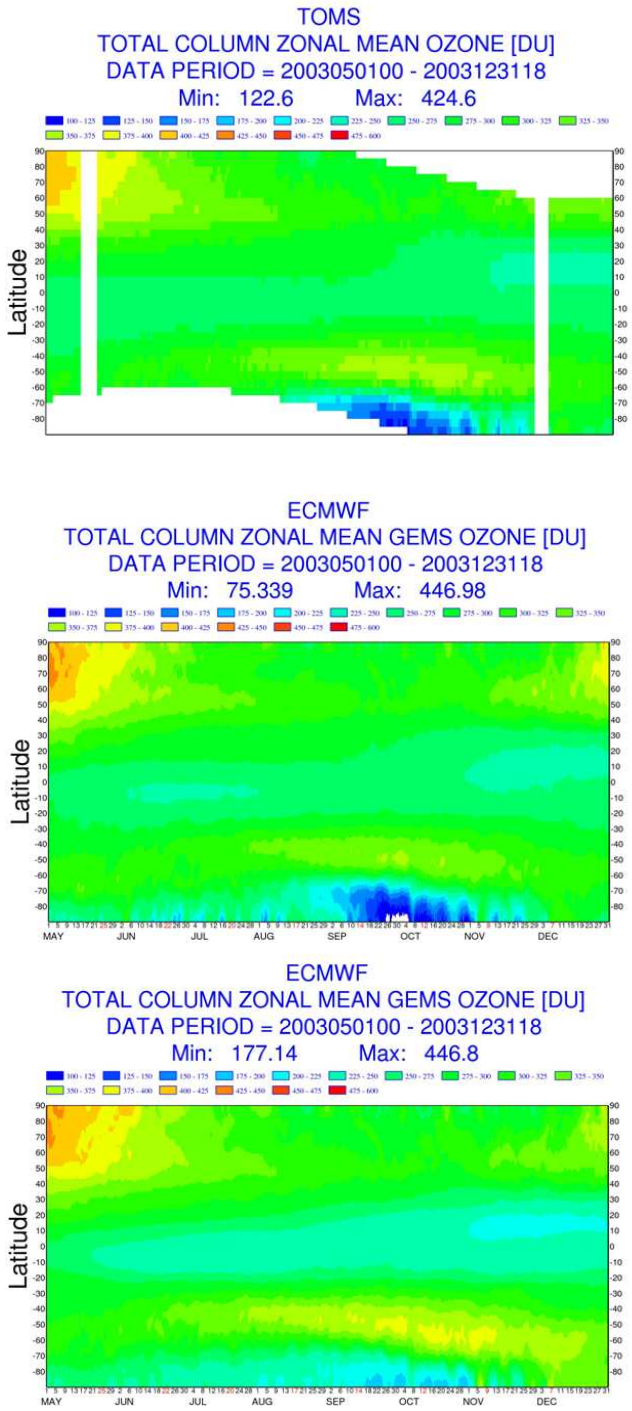

Figure 17: Time series of zonal mean total column ozone in Dobson units. Top panel: TOMS data; middle panel: analysed total column ozone; and bottom panel: total column ozone field from a control run.

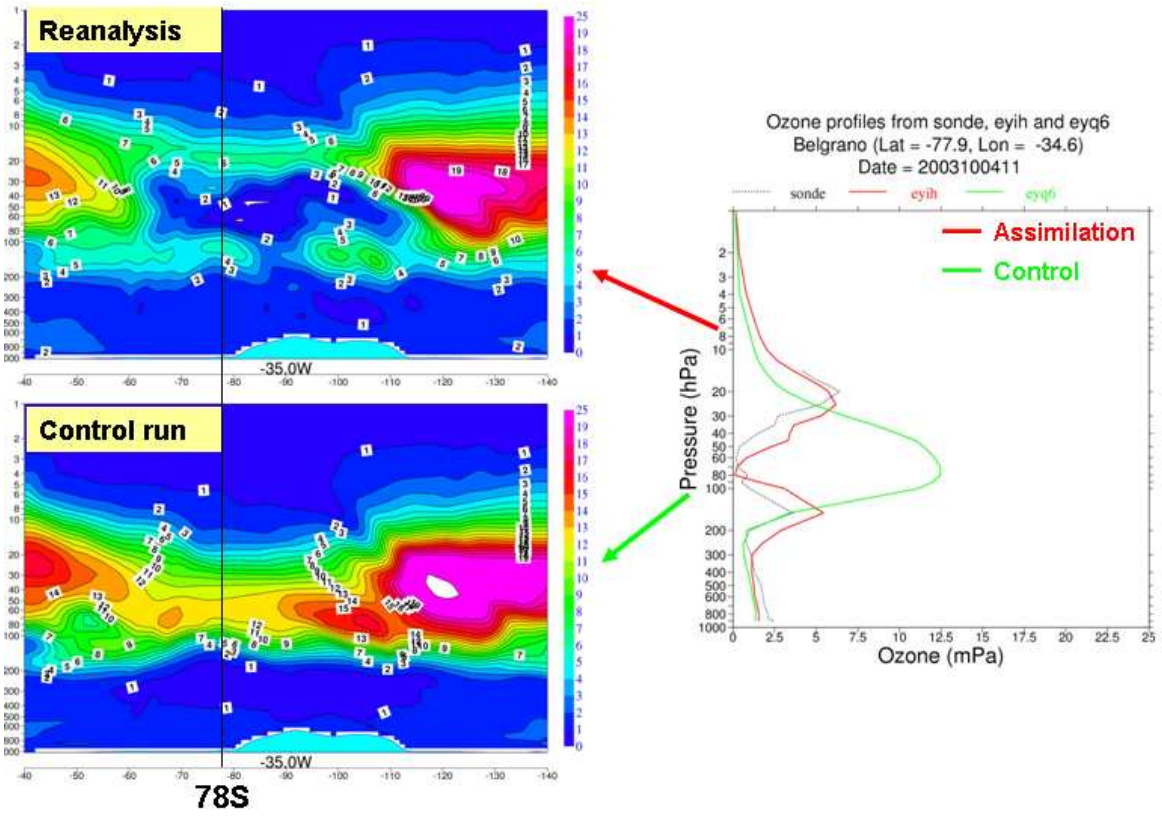

Figure 18: Ozone cross section from 40S across the South Pole back to 40S along 35E from assimilation run (top) and control run (bottom) on 4 October 2003, 12z. Ozone profiles at the Belgrano station (78S, 35E) from an ozone sonde launched on 4 October 2003 (black), assimilation run (red) and control (green). The Unit is mPa.

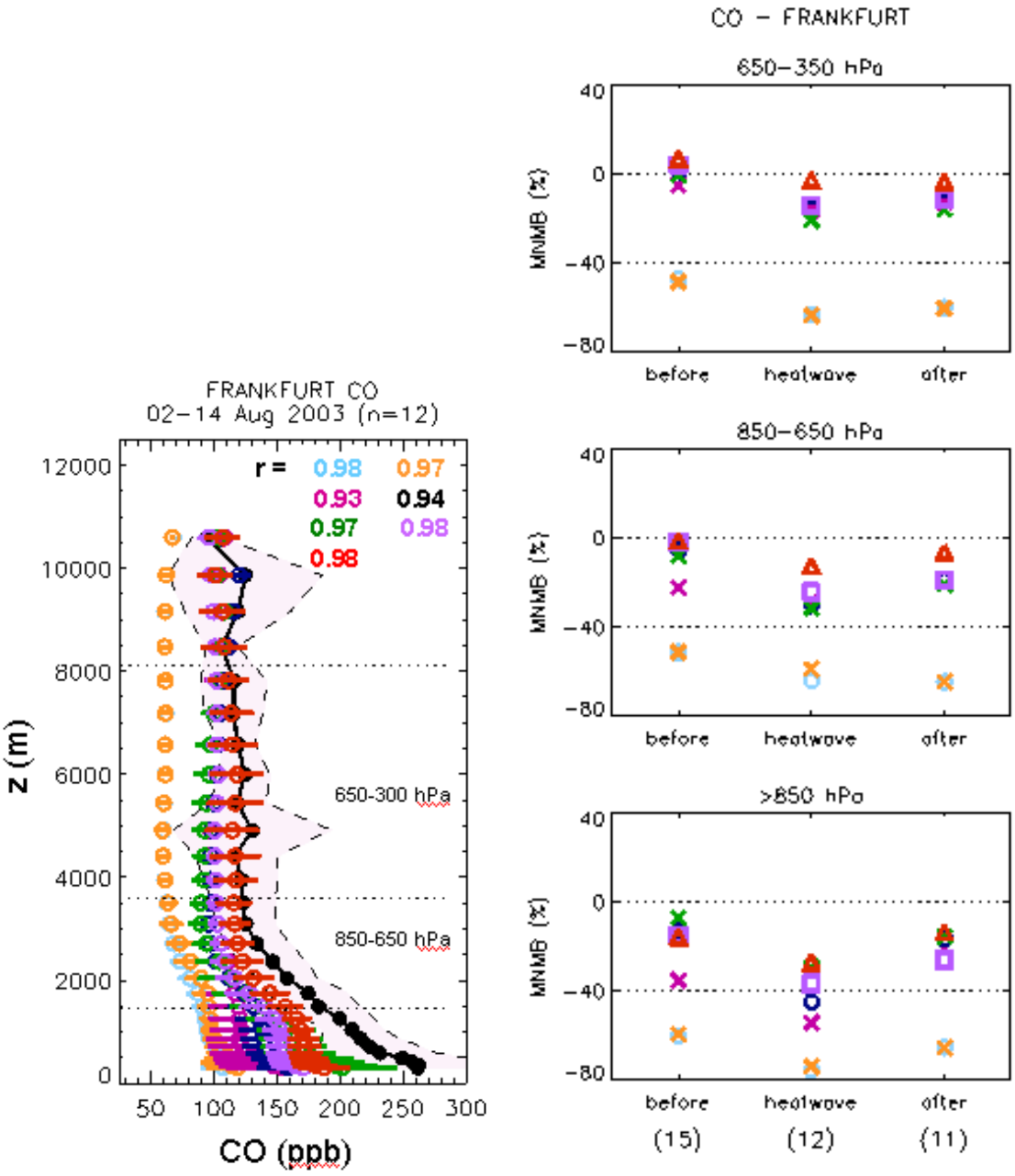

Figure 19: Left: profiles of CO over Frankfurt (02-14 August 2003) in ppb. MOZAIC observations are shown in black. Right: Modified Normalized Mean Bias in percentage before (15 day average), during (12 day average) and after (11 day average) the heat wave. Top panel shows the upper troposphere (650-350 hPa); middle panel shows the middle troposhere (650-850 hPa); and bottom panel shows the lower troposphere/surface (>850 hPa). Models are differentiated as follows: MOZART+IFS+assim (red); MOZART+IFS coupled (purple); MOZART high res (1.125^o X 1.125^o, green); MOZART med res (1.875^o X 1.895^o, blue); MOCAGE (2^o X 2^o, magenta); TM5 with zoom of 1^o X 1^o over the European domain (dark yellow); and TM5 3^o X 2^o (cyan).

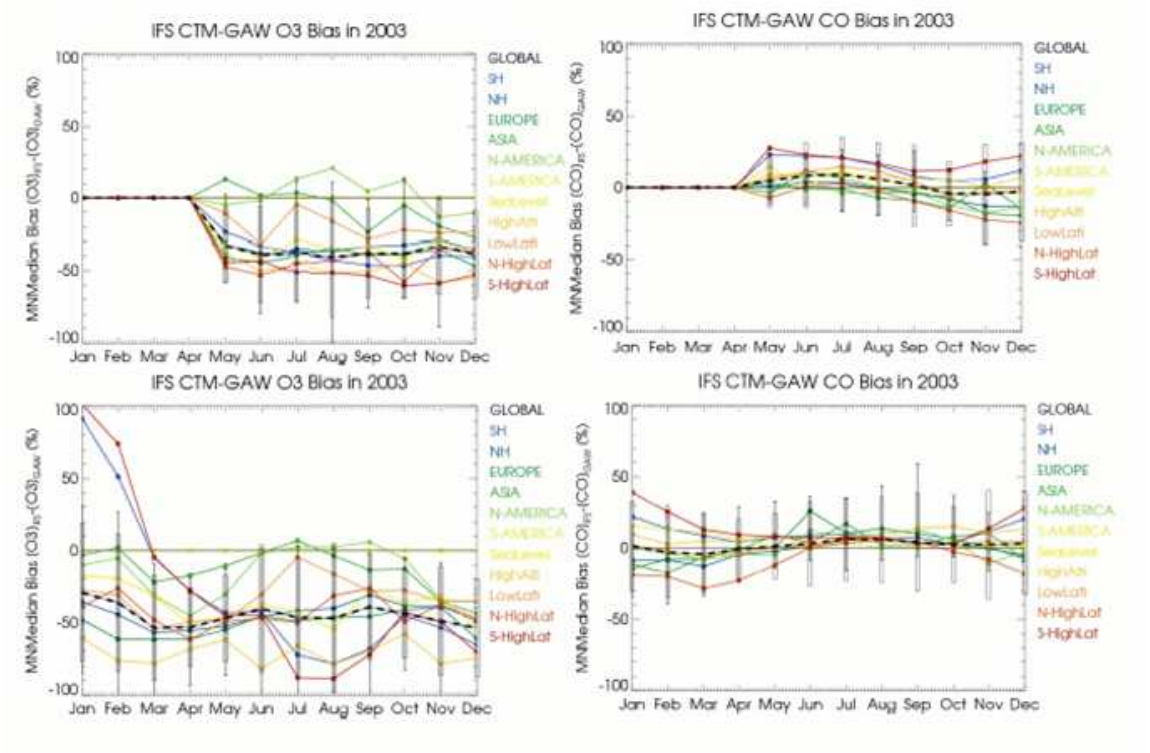

Figure 20: Normalized median bias in percentage for surface ozone (left) and CO (right). Top panels show results from the simulation without assimilation and bottom panels show results from the analysis. Regions are color-coded as follows: Global (black); Southern Hemisphere (blue); Northern Hemisphere (green/blue); Europe (dark green); Asia (green); North America (light green); South Ameirca (yellowish green); SeaLevel (yellow); High Altitude (dark yellow); Low Latitude (orange); High Latitude NH (red); High Latitude SH (purple).

6.4 Issues with representativity of mountain sites

GAW stations are supposed to be horizontally representative for a grid box size of 120 km but what is their vertical representativeness, i.e. which model level to compare with if the observation came from a mountain site? Modelled CO (O3, Aerosol) concentrations have often large vertical gradients because of surface emissions. Choosing the wrong level may lead to biases. There are several methods to choose the model levels:

- Ignore mountain stations
- Consider the difference between stations height and model orography
- Consider the level according to the fit of simulated and observed meteorological parameters such as Temperature or Relative Humidity.

6.4.1 Example: Hohenpeissenberg, 980 m

Hohenpeissenberg (HPB) is a singular mountain close to the Alps. The vertical modelled gradient in the Planetary Boundary Layer for this site can be as large as 70% for CO and -64 % for ozone. There is a large difference between station height and model orography. For example, with the 125 km orography (GEMS), the HPB peak cannot be resolved and the model only see flat high terrain. With the 16 km orography, the shape of the peak starts to being resolved and it is possible to identify a model level which could be representative of the air which is sampled at HPB. In this example, level 54 and level 50 are tested. Choosing a small-scale orography seems to better indicate to what extent the observed air was influenced by surface processes or could be considered instead free tropospheric air.

Figure 21 shows CO, temperature and relative humidity at different model levels. A few key features are noted for CO (shown in panel 21a): there are large differences between levels 60 and 54; the modelled surface diurnal cycle is very strong and level 54 and 50 are very similar. In constrast, panel 21b shows that for temperature there are small differences for level 54 and level 60. The values of temperature at level 50 appear very different from level 54. Moreover, it appear that for the 1st half of September level 60 provides a better fit to the observations while for the 2nd half of September it is level 54 that provides a better fit. For relative humidity it is difficult to tell which level is most appropriate and it appears that sub-scale influence is of paramount importance.

Conclusions from this study can be summarized as follows:

- Disregarding mountain observations is not good because there are few observations and those sampling tropospheric air are extemely valuable.
- Considering model orography versus station height might be misleading for large-scale model (HPB would be below the T159 surface).
- Considering high-resolution orography helps to better judge the near surroundings of the station.
- Looking at temperature may confirm model level choice but one has to bear in mind that temperature and CO profiles have a very different shape.

7 Preliminary conclusions

Validation has been proven fundamental to assess the current status of the GEMS analysis system. Future improvements of the system, planned in MACC, will address some of the problems highlighted in

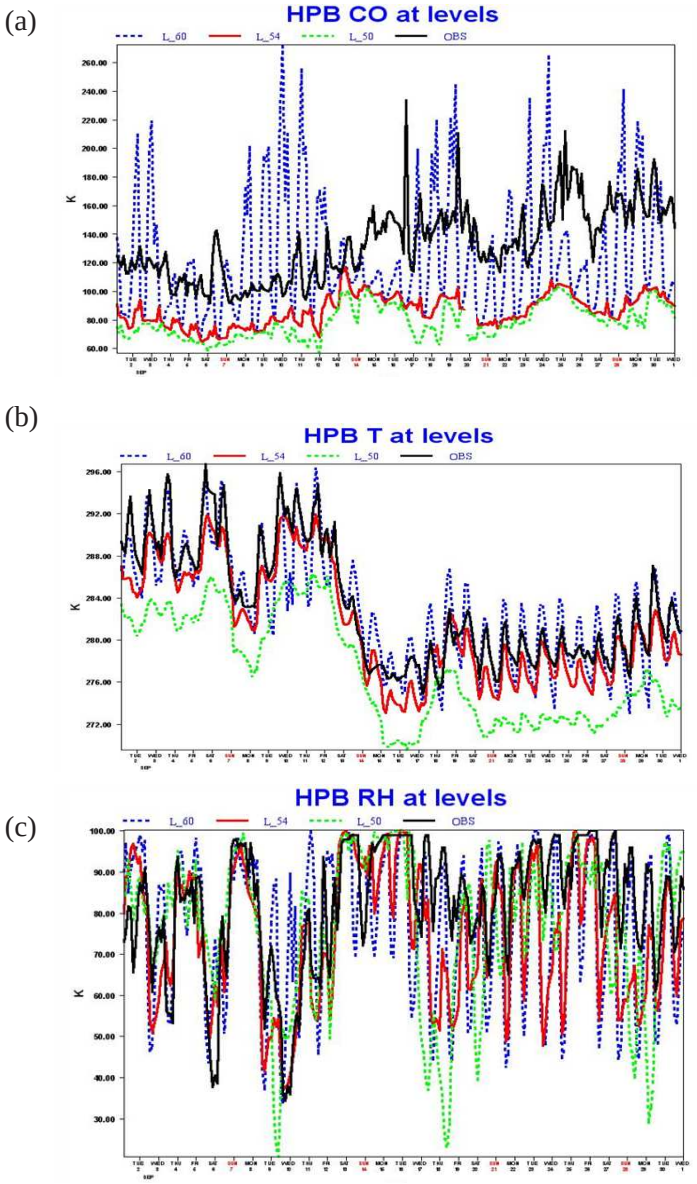

Figure 21: Atmospheric fields at Hohenpeissenberg for the month of September: (a) CO in ppb; (b) Temperature in K; and (c) relative humidity in percentage. Level 60 (8 m) is marked by the dotted blue line, Level 54 (340 m) by the solid red line, Level 50 (950 m) by the dotted green line, and the observations are in black.

the verification. The strategy for the verification has involved the use of available independent satellite, ground and aircraft-based observations of all GEMS tracers. Several metrics to measure the quality of the analyses have been used (bias, RMS, correlation, standard deviation, etc). The validation activity has stressed the need for reliable, readily available, independent verifying data sets to provide a consistent record for the validations of successive versions of the the GEMS/MACC analysis systems. The importance of comparing analysis and observations in the most objective way was also highlighted (see mountain site example of section 6.4.1). It is also important to underline that one should have realistic expectations regarding the performance of the analysis which is limited by the forward model operator, the forecast model, the prescribed background and observation errors and the inherent information content of the observations.

References

Benedetti, A., J.-J. Morcrette, O. Boucher, A. Dethof, R. J. Engelen, M. Fisher, H. Flentje, N. Huneeus, L. Jones, J. W. Kaiser, S. Kinne, A. Mangold, M. Razinger, A. J. Simmons, and M. Suttie (2009), Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, *J. Geophys. Res.*, **114**, D13205, doi:10.1029/2008JD011115.

Engelen R. J., Serrar, S. and F. Chevallier, 2009: Four-dimensional data assimilation of atmospheric CO2 using AIRS observations, *J. Geophys. Res.*, **114**, D03303, doi:10.1029/2008JD010739.

Flemming,J., A. Inness, H. Flentje, V. Huijnen, P. Moinat, M. G. Schultz, and O. Stein, 2009: Coupling global chemistry transport models to ECMWFs integrated forecast system. Technical Memorandum 590, ECMWF, 27 pp.

Hollingsworth, A., R. J. Engelen, C. Textor, A. Benedetti, O. Boucher, F. Chevallier, A. Dethof, H. Elbern, H. Eskes, J. Flemming, C. Granier, J. W. Kaiser, J.-J. Morcrette, P. Rayner, V.-H. Peuch, L. Rouil, M. G. Schultz, A. J. Simmons, and the GEMS consortium, 2008: The Global Earth–system Monitoring using Satellite and in–situ data (GEMS) Project: Towards a monitoring and forecasting system for atmospheric composition, *Bull. Amer. Meteor. Soc.*, **89**, 1147–1164, doi: 10.1175/2008BAMS2355.1.

Inness, A., J. Flemming, M. Suttie, and L. Jones, 2009: GEMS data assimilation system for chemically active reactive gases. Technical Memorandum 587, ECMWF, 26 pp.

Morcrette, J.-J., O. Boucher, D. Salmond, L. Jones, P. Bechtold, A. Beljaars, A. Benedetti, A. Bonet, A. Hollingsworth, J. W. Kaiser, M. Razinger, S. Serrar, A. J. Simmons, M. Suttie, A. Tompkins, A. Untch, and the GEMS–AER team, 2008: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecast Integrated Forecast System: 1. Forward modelling, *J. Geophys. Res.*, **114**, D06206, doi:10.1029/2008JD011235.

Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Center's spectral statistical–interpolation analysis system. *Mon. Weather Rev.*, **120**, 1747–1763.

Acronyms

AEROCE = Atmosphere/Ocean Chemistry Experiment

AERONET = AErosol Robotic NETwork

AIRS = Atmospheric Infrared Sounder

CALIPSO = Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation

CTM = Chemical Transport Model

GAW= Global Atmospheric Watch

GPS= Global Positioning System

GOME= Global Ozone Monitoring Experiment

IASI = Infrared Atmospheric Sounding Interferometer

IFS = Integrated Forecasting System

MIPAS = Michelson Interferometer for Passive Atmospheric Sounding

MLS = Microwave Limb Sounder

MOCAGE = Model of atmospheric chemistry at large scale

MODIS = Moderate Resolution Imaging Spectroradiometer

MOPITT = Measurements Of Pollution In The Troposphere

MOZAIC = Measurements of OZone, water vapour , carbon monoxide and nitrogen oxides by in-service AIrbus airCraft

MOZART = Model for OZone And Related chemical Tracers

OMI = Ozone Monitoring Instrument

SBUV = Solar Backscatter Ultraviolet

SCHIAMACHY = SCanning Imaging Absorption spectroMeter for Atmospheric ChartograpHY

TM5 = Test Model 5