Satellite Observations of Greenhouse Gases

Richard J. Engelen

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

1 Introduction

Within the EU-funded COCO and GEMS projects ECMWF has been building a greenhouse gas data assimilation system (*Engelen et al.*, 2004; *Engelen and McNally*, 2005). The idea is to monitor atmospheric concentrations of CO_2, CH_4, and N_2O by using observations from various satellites using the state-of-the-art data assimilation system at ECMWF. These consistent fields of atmospheric trace gases can then be used in surface flux inversions that are currently based on in-situ surface flask data only. Hopefully, the satellite observations will be able to fill the large spatial and temporal gaps in the surface flask network.

2 Sink variable

Within the COCO project, CO_2 was introduced in the 4-dimensional variational (4D-Var) data assimilation system as a special column variable. In a normal 4D-Var configuration variables are added to the state vector (X) at time t_0 and then adjusted to fit the observations within the assimilation time window as best as possible (minimizing root-mean-square (RMS) differences), using the model dynamics and physiscs as a hard constraint (see Figure 1). However, CO_2 was added to the minimization as tropospheric column amounts for each observation of the Atmospheric Infrared Sounder (AIRS) instrument (*Aumann et al.*, 2003). These are then adjusted individually as part of the total cost minimization. The output of one analysis cycle consists then of all the relevant atmospheric fields (temperature, winds, humidity, etc.) together with CO_2 tropospheric column estimates at all the AIRS observation locations that go into the assimilation. Because CO_2 is not part of the assimilation transport model, there are no forecasts for CO_2 and CO_2 information from one analysis cycle cannot be used in the next analysis cycle. Also, to generate global fields, individual estimates have to be gridded in for instance 5° by 5° boxes for a certain time period.

A set of only 18 AIRS spectral channels (out of 324 available channels) sensitive to tropospheric CO_2 was used to estimate the tropospheric CO_2 columns. These channels were chosen to minimize the effect of water vapour and ozone absorption. Because the signal of CO_2 in the observed radiances is so small, it is easily obscured by uncertainties in the water vapour and ozone distributions. For the background constraint a global mean value of 376 ppmv was chosen with a background error standard deviation of 30 ppmv. The analysis error was estimated based on the background error, the observation error, and the sensitivity of the observations to atmospheric CO_2. This sensitivity largely depends on the temperature lapse rate and the depth of the tropospheric layer (i.e., the height of the tropopause).

More than one year of AIRS data has been processed and Figure 2 shows monthly mean results for March 2003, September 2003, and March 2004. The fourth panel shows the monthly mean analysis error for March 2003. White areas represent areas with extensive cloud cover throughout the month. The largest signal in atmospheric

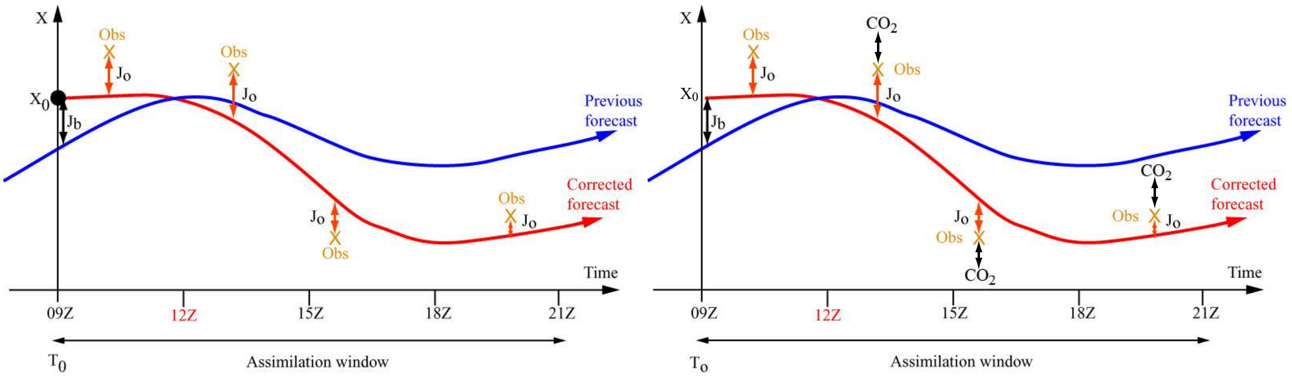

Figure 1: Schematic diagram of the principles of 4D-Var in a normal configuration (left panel) and with the extra CO_2 column variable (right panel). The initial state X_0 and optionally the CO_2 column amounts are adjusted such that the root-mean-square difference with the observations is minimized. A background term (J_b) constrains these adjustments.

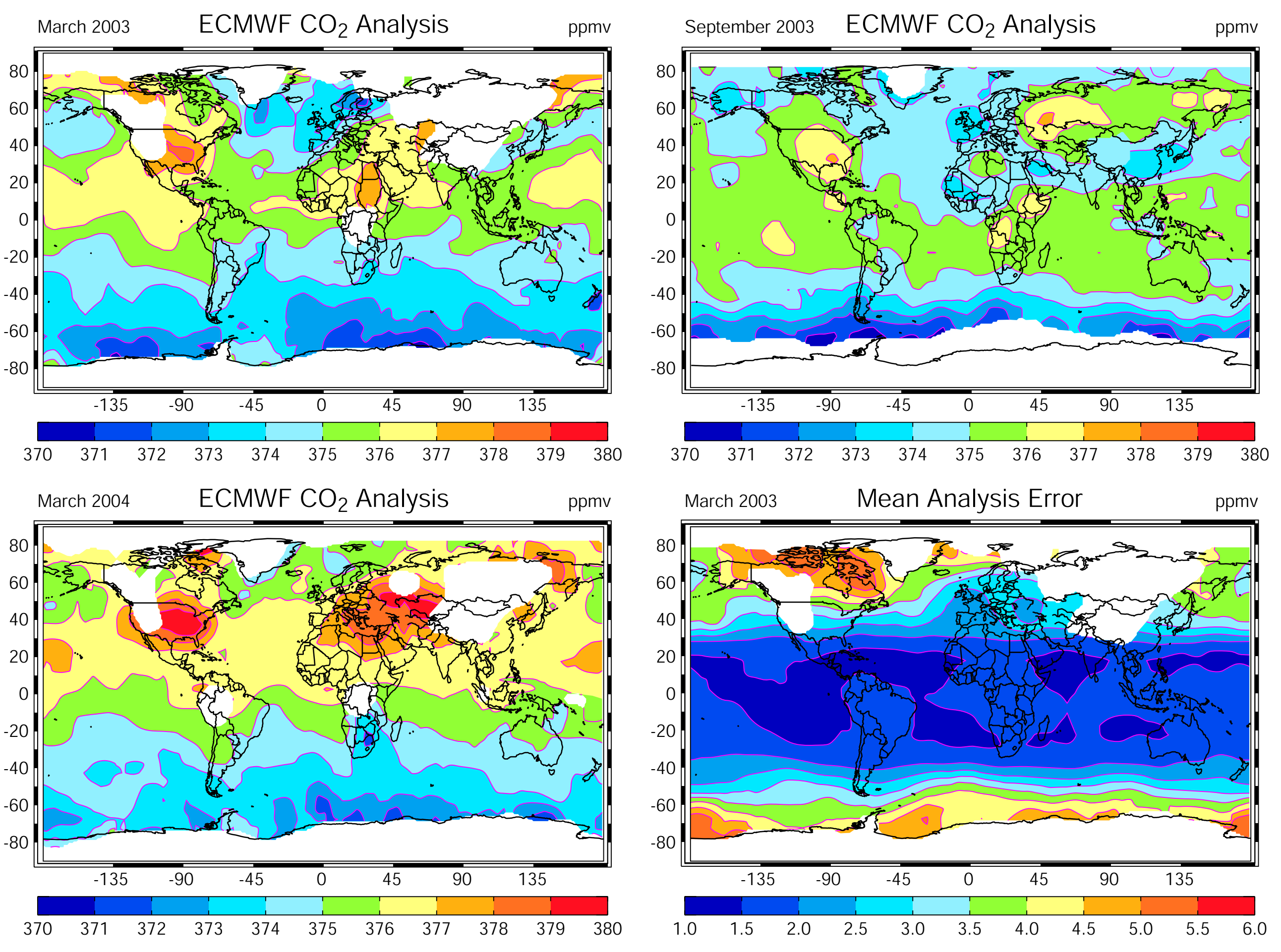

Figure 2: Monthly mean analysis results for March 2003, September 2003, and March 2004 as well as the monthly mean analysis error for March 2003.

CO_2 concentrations comes from the terrestrial biosphere. Vegetation absorbs CO_2 by photosynthesis and emits CO_2 through respiration. Plant litter on and in the soil releases CO_2 as well due to decomposition. A strong seasonal cycle is produced, although the annual net biosphere flux is very close to zero. The terrestrial biosphere also creates a latitudinal gradient in the atmospheric concentrations due to the large amount of land in the northern hemisphere compared to the southern hemisphere. This latitudinal gradient is amplified by the anthropogenic emissions that mainly originate from the northern hemisphere. Both the seasonal cycle and the latitudinal gradient are visible in the results of Figure 2. It is encouraging to see that the assimilation is capable of producing these spatial and temporal variations without having that information in the background. March 2004 shows generally higher CO_2 concentrations than March 2003, representing the upward trend in global atmospheric CO_2. The difference between March 2003 and March 2004 at the location of Hawaii is 1.6 ppmv compared to the 1.4 ppmv observed at the Mauna Loa flask station. The monthly mean error shows the clear dependence of the analysis error on the temperature lapse rate as well as the thickness of the tropospheric layer. Errors are smallest in the tropics were the tropopause is high and the temperature lapse rate is large, while they increase at higher latitudes where the tropopause is lower. The relatively low errors over Europe are caused by a higher tropopause (deeper tropospheric layer) in the sub-tropical air mass.

The presentation of monthly mean results is interesting by itself, but an important check of the validity of our analysis results is by comparing these results to independent observations of atmospheric CO_2. There are only very few data sources for 2003 and we can generally not use the surface flask data, because our estimates represent a layer between about 700 hPa and the tropopause, while the surface flasks are sampled in the boundary layer. Only if we are sure that the full tropospheric CO_2 profile is well-mixed, a comparison would be useful. However, Dr Hidekazu Matsueda and colleagues at the Japanese Meteorological Agency have been measuring CO_2 on board commercial flights of the Japanese Airlines (JAL) flying between Japan and Australia (*Matsueda et al.*, 2002). These observations consist of automatic flask samples gathered at altitudes between 8 and 13 km on biweekly commercial flights. For 2003, 21 flights were available for our comparisons. Figure 3 shows the CO_2 annual cycle for both the flight observations and the assimilation estimates. For the full processed period (1 January 2003 - 31 March 2004), CO_2 analysis estimates were sampled in 6° x 6° boxes around the locations and over a period of 5 days around the date of the flight observations. We then generated three plots that represent the northern hemisphere region, the equatorial region, and the southern hemisphere region, by averaging the respective box averages for each region together. The figure shows that the analysis estimates follow the JAL observed annual cycle quite well. All differences fall within the 1-σ error bars and are of the order of 1 ppmv in most cases. There is a clear improvement compared to the used background, which is 376 ppmv throughout the year. The main anomaly can be seen in both the northern hemisphere and the southern hemisphere in January and February, in which period the analysis estimates are consistently higher than the JAL observations.

3 4D-Var system

Within the GEMS project the above described system is being extended to a full 4D-Var greenhouse gas data assimilation system. This requires implementation of the greenhouse gases in the forecast model as well as a proper specification of the 3-dimensional background constraint. We have outlined the main ingredients for the system in the subsections below.

3.1 Observations

The main focus initially will be on the assimilation of satellite data. In-situ data (e.g., surface flasks, tall tower continuous measurements, airborne flasks) will be used as validation data, but can be added to the data analysis

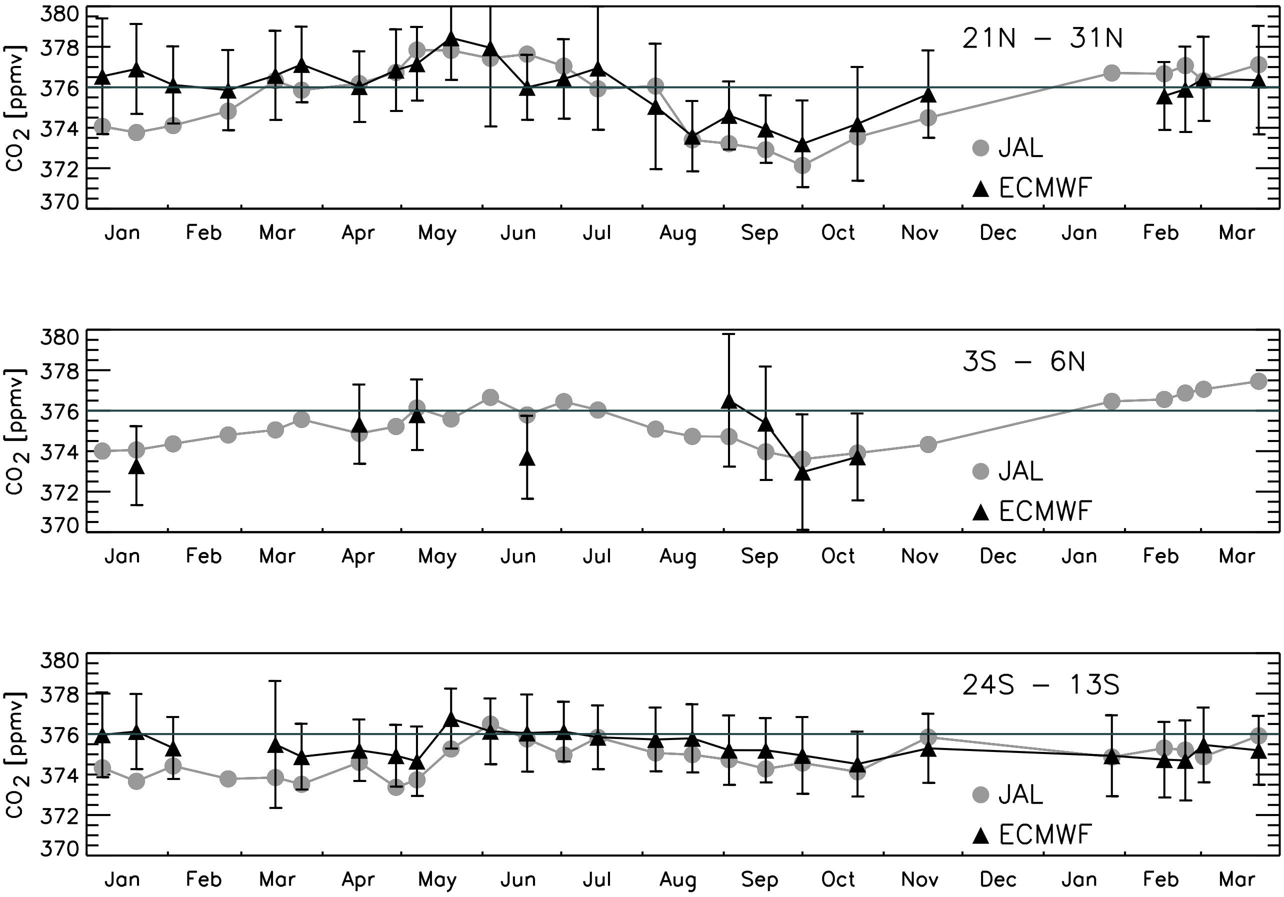

Figure 3: Comparison of CO_2 estimates with JAL observations for three different latitude zones from January 2003 to March 2004. Missing ECMWF data are caused extensive cloud cover in the area.

at a later stage. Currently, ECMWF assimilates AIRS radiances operationally. These observations were already used for the CO_2 column estimates as described above. In the next few years we expect to assimilate infrared radiances from the Infrared Atmospheric Sounding Interferometer (IASI) and Cross-track Infrared Sounder (CrIS) instruments. All three instruments measure the infrared spectrum at high spectral resolution and these observations can be used to constrain atmospheric temperature, water vapour, carbon dioxide, ozone, carbon monoxide, nitrous oxide, and methane. In 2008 two instruments specifically designed to observe CO_2 will be launched: the Orbiting Carbon Observatory (OCO) and the Greenhouse Gases Observing Satellite (GOSAT). Both instruments will measure the solar reflection in the 1.6 μm and 2.0 μm CO_2 absorption bands. While this measurement method will be very sensitive to lower tropospheric CO_2 (in contrast with the infrared methods), it also suffers from aerosol and cirrus cloud scattering. Uncorrected single scattering will result in underestimated CO_2 columns, and uncorrected multiple scattering will result in overestimated CO_2 columns. This scattering correction requires very accurate radiative transfer modelling, which is even more pressed by the necessity to account for polarization effects. The OCO and GOSAT science teams will initially most likely retrieve CO_2 column amounts for clear pixels only, which we will try to assimilate together with the infrared satellite radiances. A good error characterization of these column amounts will be critical.

3.2 Forecast model

A vital requirement of any data assimilation system is a forecast model that is able to match the observations within the specified error margins. Fitting the observations by only adjusting the initial state assumes a hard constraint from the model dynamics and physics. This can lead to significant errors in the analysis, if the model is not accurate enough. The greenhouse gases have been implemented as tracers in the Integrated Forecasting System (IFS) forecast model. The tracer transport (both advection and vertical mass fluxes) is currently being tested by using Radon and SF_6 as tracers. For CO_2 we also implemented climatological surface fluxes. For the ocean we use fluxes based on *Takahashi et al.* (1999), the fossil fuel emissions are based on *Andres et al.* (1996), and the natural biosphere fluxes are based on the CASA model (*Randerson et al.*, 1997). Figure 4 shows examples of these fluxes with natural biosphere fluxes for December and July in the top two panels, and Jult ocean fluxes and annual mean anthropogenic emissions in the bottom panels. The top panels show the very distinct seasonal cycle of the northern hemisphere vegetation, as well as the effetc of dry and wet seasons on the tropical vegetation. The ocean fluxes show the release of CO_2 into the atmosphere from the warm tropical water and the uptake of CO_2 from the atmosphere in the cold sinking water around Greenland. At mid-latitudes, phytoplankton generally takes CO_2 from the atmosphere into the ocean.

Monthly mean simulation results are shown in Figure 5 for March 2003. The simulation was started at 1 January 2002 and ran 12 hour forecasts every 12 hours starting from operational analysis fields. Surface fluxes are interpolated in time from monthly means for the biosphere and the ocean, while the anthropogenic emssions are an annual mean. These surface fluxes will be improved to contain day-to-day variability and even diurnal variability in case of the biosphere. The figure shows the clear zonal gradient between northern and southern hemisphere in the northern hemisphere spring, because of the stalled photosynthesis during the winter in combination with CO_2 release due to respiration and due to anthropogenic emissions. The effect of tropical convection is also visible.

3.3 Bias correction

4D-Var data assimilation is based on the general assumption that errors are random. Therefore, any significant systematic errors in the observations and/or the radiative transfer model need to be corrected before proper assimilation can be done. Model bias should be corrected as well, but is difficult to estimate. Ideally, model bias should be corrected by improving the model itself. Apart from biasing the background state and therefore

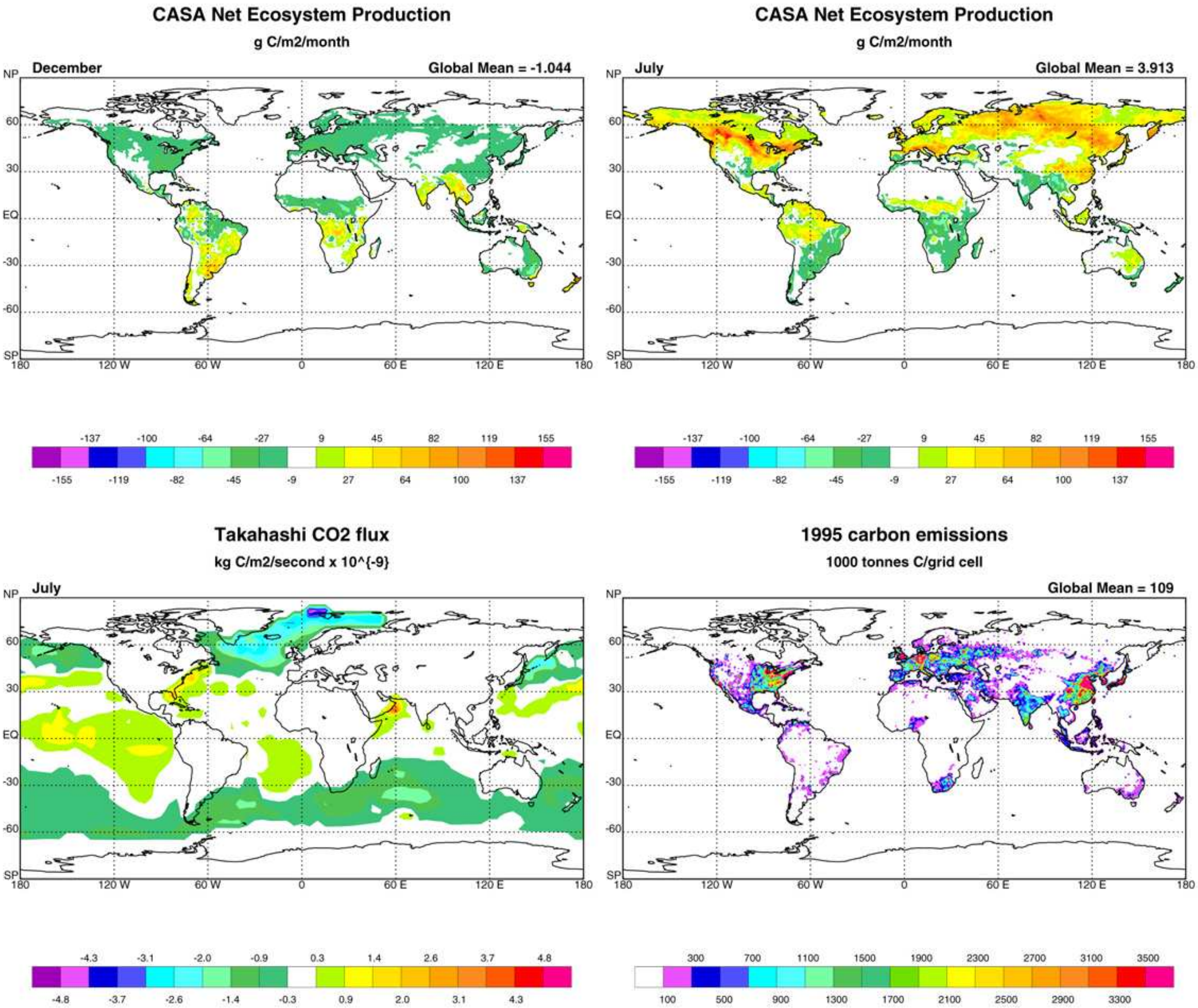

Figure 4: Monthly mean biosphere fluxes from the CASA model for December (top left) and July (top right); monthly mean ocean fluxes from Takahashi et al. (1999) (bottom left); annual mean anthropogenic fluxes from Andres et al. (1996) (bottom right). Biosphere fluxes are positive into the vegetation, while ocean and anthropogenic fluxes are positive into the atmosphere.

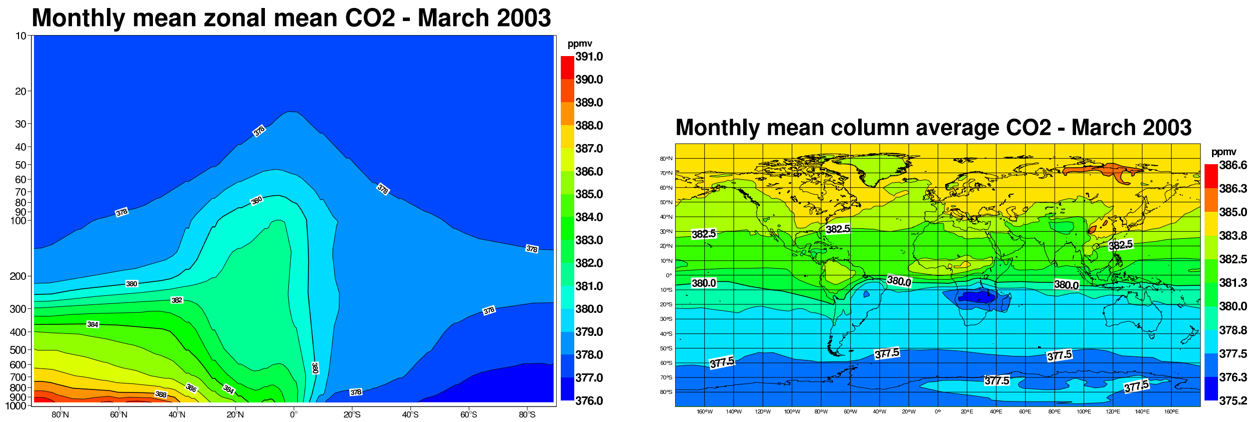

Figure 5: Simulated monthly mean zonal mean CO_2 mixing ratios (left) and monthly mean column averaged CO_2 mixing ratios (right) for March 2003 after 15 months of spin-up.

the analysis state, the forecast model is also used as a hard constraint within the assimilation time window. Any bias within this time period (usually 12 hours) will bias the information from the observations. Biases are generaly detected by monitoring the so-called **O** - **B** departures (differences between the observations and the model simulated observations). These departures show all the random variability as well as systematic differences. Systematic differences on time scales of 2 - 4 weeks are then denoted as bias. However, by using this method, model bias might end up in the observation bias correction, because there is no straightforward method to distinguish between model bias and observation bias. Therefore, any bias correction method is in theory capable of removing some of the CO_2 (CO, CH_4, N_2O) signal. The main problem here is that we do not have many accurate CO_2 profile observations to check for model bias, so that we can correct the satellite observations properly for the observation bias. Especially, at the start of the first GEMS reanalysis, we have no clear idea of the errors in the starting analysis. These errors will be partly corrected by assimilating the satellite observations, but systematic differences between the model forecast and the observations will probably remain for some time. By using independent data we will try to get a feeling for these systematic errors, but this will likely be problem area.

3.4 Background constraint

The background constraint is a crucial part of the data assimilation system. The background covariance matrix describes the horizontal and vertical correlations of the errors in the background state. Therefore, any correction of the background state by an observation will be distributed accordingly as can be seen in Figure 6. The figure shows the incremental effect of a single ozone observation on the ozone field, both in the horizontal (left panel) as in the vertical (right panel). The observation not only corrects the initial state at the observation location, but also in a 3-dimensional area around the observation. Therefore, if the background error correlations are wrongly specified, incorrect increments will result.

When trying to specify the background covariance matrix we encounter generally two problems: i) we want to describe the statistics of the errors in the background, but we do not know what the true state is; ii) the background covariance matrix is enormous ($\sim 10^7 \times 10^7$), so we are forced to simplify it. Differences between 48 and 24 hour forecasts (NMC method, *Parrish and Derber*, 1992) or an analysis-ensemble method (*Fisher*, 2003) are usually used to estimate the background error covariance matrix. Both methods, however, rely on the availability of enough observations constraining the relevant atmospheric parameters.

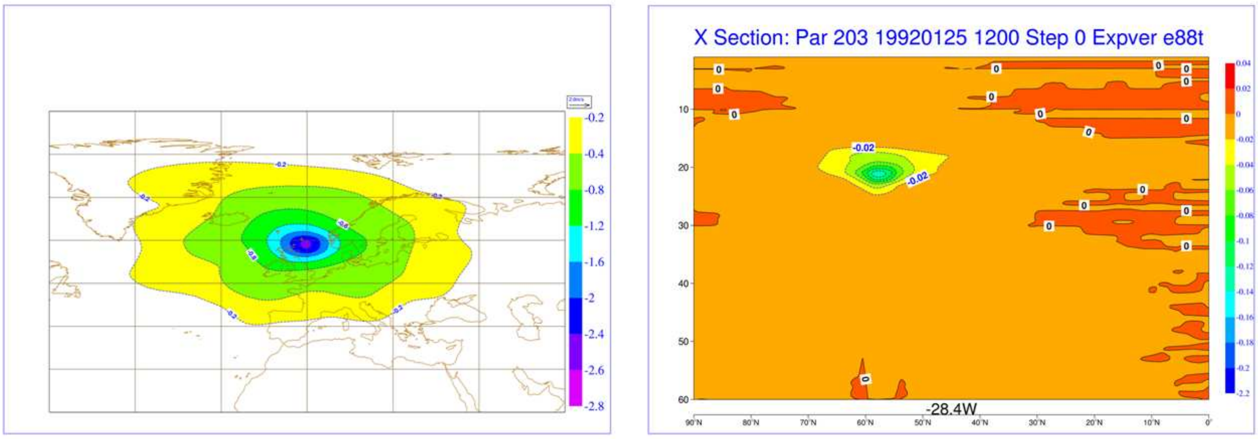

Figure 6: Example of the effect of the specified background error covariances on the ozone increments when a single observation is assimilated.

For the greenhouse gases we do not have enough observations to generate proper statistics for the covariance matrix of the full 3-dimensional field. AIRS observations only constrain the mid- and upper troposphere and the stratosphere in a crude sense, and surface flasks are very sparsely distributed over the globe. We will therefore try to define a covariance model in which a few parameters can be estimated from observations. This method, also known as maximum likelihood estimation of covariance parameters, is more extensively described in *Dee and da Silva* (1999a,b) and *Michalak et al.* (2005). Likely covariance parameters to estimate from observations are the variances themselves and horizontal and vertical correlation lenght scales.

4 Summary

Within the COCO project CO_2 has been build into the 4D-Var data assimilation system as a simple column variable using AIRS observations to estimate the mean mixing ratios. Results are very encouraging, especially in the tropics. This system is now being extended into a full 4D-Var greenhouse gas data assimilation system as part of the GEMS project. This is applied research into new territory with its own problem areas that have been described above. When a working system has been build, it will be able to process observations from several satellite instruments to provide consistent fields of atmospheric CO_2. These 3-dimensional fields can then be used to improve off-line flux inversions that are currently based on surface flasks only.

Acknowlegdements

The author would like to thank Soumia Serrar, Antje Dethof, Agathe Untch, Tony McNally, and the rest of the GEMS team for their contributions to this work. The work was funded by the EU FP5 project COCO and the EU FP6 project GEMS.

References

Andres, R. J., G. Marland, I. Fung, and E. Matthews (1996), Distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, *Global Biogeochem. Cycles*, *10*, 419–429.

Aumann, H. H., et al. (2003), AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems, *IEEE Trans. Geosci. Remote Sensing*, *41*, 253–264.

Dee, D. P., and A. da Silva (1999a), Maximum-likelihood estimation of forecast and observation error covariance parameters. part I: Methodology, *Mon. Wea. Rev.*, *127*, 1822–1834.

Dee, D. P., and A. da Silva (1999b), Maximum-likelihood estimation of forecast and observation error covariance parameters. part II: Applications, *Mon. Wea. Rev.*, *127*, 1835–1849.

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.

Engelen, R. J., E. Andersson, F. Chevallier, A. Hollingsworth, M. Matricardi, A. P. McNally, J.-N. Thépaut, and P. D. Watts (2004), Estimating atmospheric CO_2 from advanced infrared satellite radiances within an operational four-dimensional variational (4D-Var) data assimilation system: Methodology and first results, *J. Geophys. Res.*, *109*, D19309, doi:10.1029/2004JD004777.

Fisher, M. (2003), Background error covariance modelling, in *Proceedings of the ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean*.

Matsueda, H., H. Y. Inoue, and M. Ishii (2002), Aircraft observation of carbon dioxide at 8-13 km altitude over the western Pacific from 1993 to 1999, *Tellus*, *54B*, 1–21.

Michalak, A. M., A. Hirsch, L. Bruhwiler, K. R. Gurney, W. Peters, and P. P. Tans (2005), Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, *J. Geophys. Res.*, *in press*.

Parrish, D. F., and J. C. Derber (1992), The National Meteorological Center's spectral statistical-interpolation analysis system, *Mon. Wea. Rev.*, *120*(8), 1747–1763.

Randerson, J. T., M. V. Thompson, T. J. Conway, I. Y. Fung, and C. B. Field (1997), The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, *Global Biogeochem. Cycles*, *11*, 535–560.

Takahashi, T., R. H. Wanninkhof, R. A. Feely, R. F. Weiss, D. W. Chipman, and coauthors (1999), Net sea-air CO_2 flux over the global oceans: An improved estimate based on the sea-air $p\mathrm{CO}_2$ difference, in *Proceedings of the 2nd CO_2 in Oceans Symposium, Tsukuba, Japan*.