Assimilation of satellite data for the environment

Frédéric Chevallier[1] and Richard J. Engelen[2]

1 Laboratoire des Sciences du Climat et de l'Environnement (LSCE), France
2 European Centre for Medium-Range Weather Forecasts (ECMWF), UK
frederic.chevallier@lsce.ipsl.fr

Abstract

Satellite data play an outstanding role in the monitoring of the Earth, not only for numerical weather prediction, but also for the study of atmospheric composition, of the ocean, and of soil properties. This paper illustrates the assimilation of satellite data to monitor the environment with the example of carbon dioxide, for which a space-based observation system is being implemented. Current developments within the GEMS project indicate that even the AIRS instrument, with its limited sensitivity to CO_2 variations in the upper troposphere, brings some information about the CO_2 surface fluxes.

1. Introduction

Orbital remote-sensing provides a powerful but ambiguous way to observe the properties of the Earth system. A photon collected by a satellite radiometer or a lidar contains a mixture of signals. Some of these signals are directly relevant for weather prediction, like those originating from temperature, pressure, clouds or humidity. They are usually convolved with other environment-related information about aerosols, gases, vegetation, ocean and soil. The art of satellite data assimilation relies on the capacity to disentangle these signals. For instance, even though the Television Infrared Observation Satellite Operational Vertical Sounder (TOVS) was designed for NWP, uncertainties in carbon dioxide and aerosols concentrations hamper the retrieval of temperature and humidity from its radiances (e.g., Turner 1993, Pierangelo et al., 2004). The modern NWP data assimilation systems, like four-dimensional variational ones (4D-Var, Courtier et al. 1994) provide a flexible and consistent framework to properly analyse the diverse signals present in the radiances. In addition to potentially reducing the biases of the NWP analyses, such developments make it possible to monitor key components of the global environment within existing NWP systems, i.e. at a reduced computational cost (e.g., Dethof and Holm 2004).

A typical example of the growing interest of the NWP community for the environment is given by the ECMWF-lead Global and regional Earth-system Monitoring using Satellite and in-situ data (GEMS; http://www.ecmwf.int/research/EU_projects/GEMS/index.jsp) project funded by the European Commission. GEMS aims at building a global assimilation/forecasting system for greenhouse gases, reactive gases, and aerosol and will implement a linked programme of regional air quality forecasting. The GEMS consortium consists of 31 members for the duration of four years (started in March 2005).

This paper illustrates the vast topic of the assimilation of satellite data for the environment with the case of carbon dioxide. The monitoring of this gas has become a major scientific challenge over the last few decades. It has progressed mainly through the exploitation of in situ measurements of CO_2 concentrations at a network of ground stations. With the help of atmospheric transport information, inverse methods have been applied to deduce the patterns and amplitude of surface fluxes that generated the observed concentrations (e.g., Bousquet et al. 2000). However, with about 100 stations, the current network of surface stations provides limited information about the global-scale spatial and temporal variations of CO_2 fluxes, most particularly over land (e.g., Gurney et al. 2002). In spite of the known impact of CO_2 variations on some satellite records (e.g., Turner 1993, 1994), CO_2 concentrations have been estimated from radiance measurements only recently. Retrieval methods were first applied to TOVS (Chédin et al. 2003) and to the Atmospheric Infra-Red Sounder (AIRS) (Engelen et al. 2004, Crevoisier et al. 2004) to infer concentrations in the upper troposphere. Within the GEMS project, the ECMWF 4D-Var system has been extended to analyse CO_2 concentrations. An inverse method is applied at LSCE to infer the CO_2 surface fluxes from these analyses. This paper presents the progress achieved. The next section describes the chosen strategy. Results of the CO_2 analysis and the CO_2 flux inversions are shown in section 3 and 4. Conclusions follow in section 5.

2. Method

Even though they involve two distinct systems with different models and different space-time scales, both the CO_2 analysis and the CO_2 flux inversion rely on the same Bayesian statistical framework. In both cases a similar cost function J is minimized to find the optimal concentrations and fluxes. J is defined by:

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x_b})^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x_b}) + (H(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1} (H(\mathbf{x}) - \mathbf{y}) , \tag{1}$$

with $\mathbf{x}$ the vector of variables to be optimized, $\mathbf{x_b}$ its prior values, $\mathbf{y}$ the observation vector, H the observation operator that computes the observation-equivalent from $\mathbf{x}$, and $\mathbf{B}$ and $\mathbf{R}$ the covariance matrices of the prior and of the observation errors respectively.

For the analysis of CO_2 concentrations, J is the full 4D-Var cost function. In this case, $\mathbf{x}$ includes not only the 3D CO_2 concentrations but also all the standard NWP variables like 3D temperature, humidity and pressure. In the same way, the observations in $\mathbf{y}$ are not restricted to the AIRS radiances, but gather all standard NWP observations,such as the various microwave sounders, geostationary satellites and radiosonde reports. H is the ECMWF forecast model in combination with the Radiative Transfer for the TIROS Operational Vertical Sounder (RTTOV, Matricardi et al. 2004 and references therein), upgraded to model the CO_2 transport with prescribed climatological surface fluxes as boundary conditions. The cost function is minimized for consecutive12-hour time windows. The information, including the CO_2 concentrations, is transferred from one segment to the next via the forecast model. AIRS radiances mainly inform about CO_2 in the mid- and upper-troposphere but thanks to the dynamical analysis of the 12-hour 4D-Var, some deeper information in the troposphere is collected. The spatial resolution of this study currently corresponds to a T159 truncation, i.e. about 125 km, and 60 vertical levels. This system is an extension of the work of Engelen et al. (2004) and Engelen and McNally (2005).

The CO_2 flux inversion system is based on the work of Chevallier et al. (2005b, 2007). The control vector $\mathbf{x}$ is made of eight-day surface fluxes at a 3.75^ox2.5^o (longitude-latitude) horizontal resolution throughout the temporal window of the inversion and of the concentrations at the initial time step of the inversion window. The eight-day resolution is motivated by the large auto-correlations of the prior surface flux errors over days (Chevallier et al. 2006). As observations $\mathbf{y}$, the inversion system currently uses the CO_2 concentrations at 200hPa analysed by the first system at 00, 06, 12 and 18 UTC. In order to reduce the amount of data and to limit the observation error spatial correlations, only 10% of the data (randomly selected) are actually processed. The CO_2 flux inversion implies modelling the CO_2 transport from the surface to the upper troposphere, where the appropriate AIRS channels are the most sensitive to CO_2. This second inversion system therefore involves timescales of weeks or months, much beyond the 12-hour window of the ECMWF 4D-Var. In order to keep the computational burden of a single inversion to a reasonable amount, the spatial resolution has been reduced for this second study: we currently use 19 vertical levels, 3.75^o in longitude and

2.5° in latitude. The transport is also simplified. $H(\mathbf{x})$ in equation (1) is actually linearized for the flux inversion around the values of the ECMWF prior concentrations, so that:

$$H(\mathbf{x_b}+\delta\mathbf{x}) = H(\mathbf{x_b}) + \mathbf{H}\,\delta\mathbf{x} \quad (2)$$

In Eq. (2), $\delta\mathbf{x}$ represents any flux increment in the inversion system. **H** is a Jacobian matrix obtained from the transport model of the Laboratoire de Météorologie Dynamique (LMDZ, Hourdin et al., 2006), nudged to ECMWF winds and in an off-line mode (transport mass fluxes are read from a frozen archive rather than computed on-line). This linearization strategy is inspired from the ECMWF incremental 4D-Var where the prior state of the model is computed at the highest affordable resolution with the full forecast model and increments are computed at lower resolution with a simplified model (e.g., Trémolet, 2005). Note that the computations involving the Jacobian matrix **H** are performed by the tangent-linear and the adjoint codes of the LMDZ model.

Despite their differences, effort has been made to homogenize the two systems. This is all the more important since the information content of AIRS about surface fluxes is small (Chevallier et al. 2005a) and artificial patterns can easily be introduced in the analysed fluxes. The linearization strategy of Eq. (2) illustrates this effort. Further, the same prior surface fluxes (anthropogenic from http://cdiac.esd.ornl.gov/ndps/ndp058a.html , air-sea exchange from http://www.ldeo.columbia.edu/res/pi/CO2/carbondioxide/air_sea_flux/fluxdata.txt, terrestrial ecosystem exchange from the Carnegie Ames Stanford Approach model - CASA, Randerson et al. 1997 - and biomass burning from http://www.daac.ornl.gov/VEGETATION/guides/global_fire_emissions_v2.1.html) are used in both systems. Last, the concentrations at the initial time step of the inversion window come from the ECMWF system.

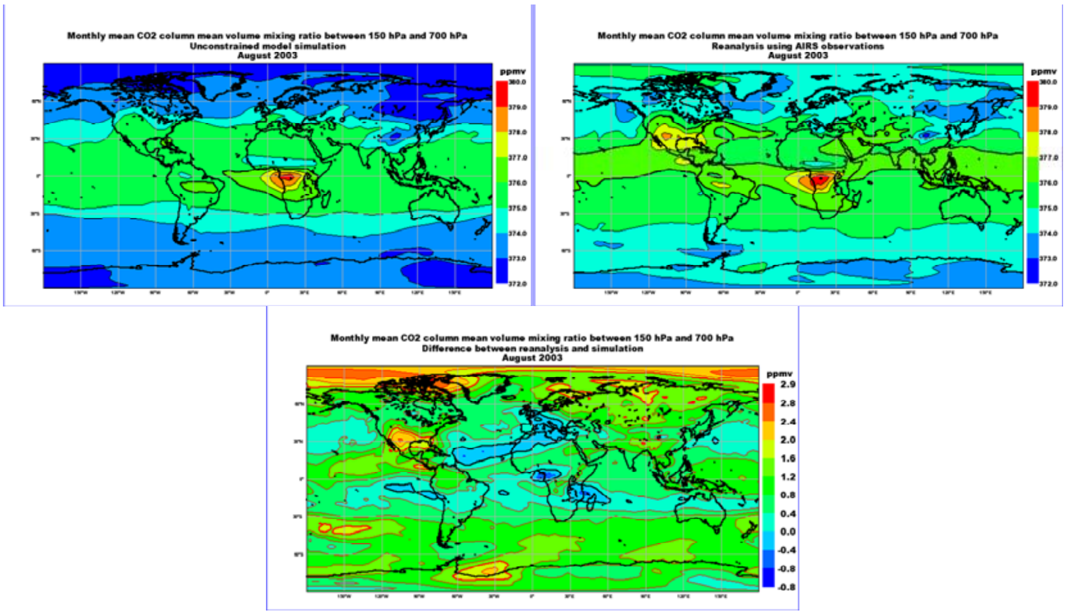

Figure 1: Monthly mean partial column CO_2 for August 2003 after eight months of assimilation using AIRS information on CO_2 (top right) or with unconstrained CO_2 (top left). The bottom figure displays the difference between the upper maps. All values are in ppm.

3. Analysis of the CO_2 concentrations

Figure 1 illustrates the behaviour of the CO2 concentration analysis system: the monthly mean column-averaged (between 150 hPa and 700 hPa) CO_2 volume mixing ratio after eight months of assimilation using AIRS shows small (about a few tens of a ppm) but significant changes to a simulation with unconstrained CO_2.

Direct validation of the analysed concentrations can only be performed by comparison with the numerous but sparse in situ measurements. Those data mainly consist of point-wise measurements at several tens of surface stations, but some aircraft measurements are also available that allow validating the concentration profiles. Figure 2 gives the example of two flight profiles between the surface and 7 km at Molokai Island in the Central Pacific. In both cases, the analysis improves on the free-running model at all altitudes. Due to the AIRS limited sensitivity, the analysis increments are less than a ppm and the improvement is more spectacular when the background error is not any larger.

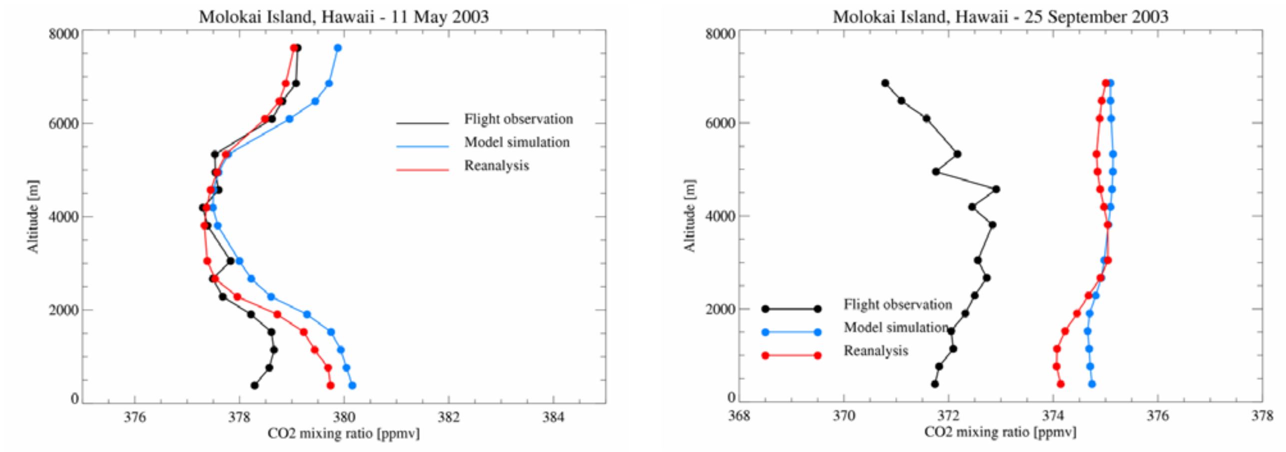

Figure 2. Comparison with flight data over Hawaii (courtesy of Pieter Tans, NOAA/ESRL).

Currently, a number of aspects of the analysis system are being improved. They concern the AIRS channel selection and the error statistics of the prior (variances and correlations) and of the observations (biases).

4. Inversion of the surface fluxes

Eleven months of CO_2 analysis, from January to November 2003, were available for the surface flux inversion. The fluxes were inferred in a single inversion. The resulting cumulated flux increments are shown in Figure 3. As expected with observations high in the atmosphere (e.g., Chevallier et al. 2005b), the increment patterns are not primarily driven by gradients in the observation-minus-background departures. They are mainly located in regions of large vertical transport, such as the equatorial oceans, and in regions of large uncertainty about the surface fluxes, i.e. in the continental vegetated areas. These expected shapes are modulated by a zonal pattern: positive in the Tropics and negative elsewhere.

Flux measurements are only available at small scales of the order of 1 km2. Given the flux spatial heterogeneity, they can hardly be used to validate the inversion. The in situ concentration measurements, already used in the previous section, allow us to indirectly evaluate the quality of the analysed fluxes. In this case we use the LMDZ transport model with prior and inverted fluxes successively as boundary conditions and we evaluate the two simulations against the in situ measurements gathered in the GLOBALVIEW-CO2 (2006) dataset. The concentrations at the initial time step of the 11-month window are also adjusted in the inversion and are evaluated in the same way. The impact of the inversion is shown in Figure 4 with the difference between the root mean square (RMS) prior departures and the posterior RMS. One can see spatially consistent patterns of improvements in Europe, around Africa and around Central America, and of degradation in the northern equatorial Pacific.

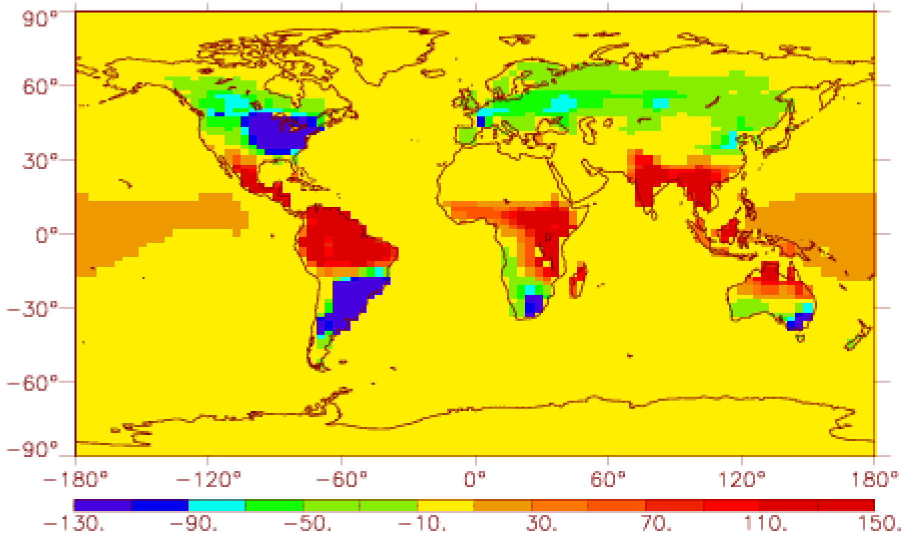

Figure 3. Cumulated surface flux increments, in gC.m-2, generated from the 11-month-inversion (January-November 2003). The increments are defined as the inverted fluxes minus the prior. The fluxes are counted positive when the surface releases carbon in the atmosphere.

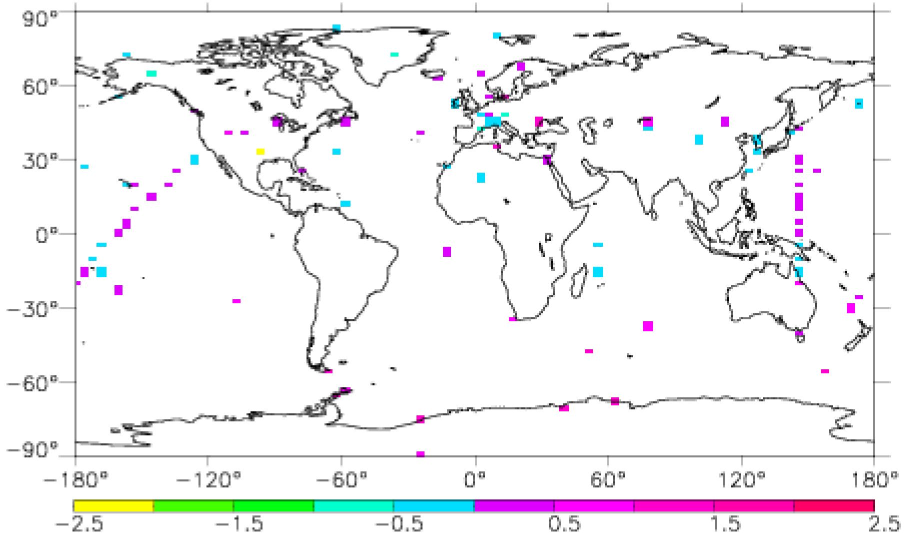

Figure 4. RMS of the posterior departures minus RMS of the prior departures, in ppm. Improvements are characterized by negative (yellow, green and blue) values. The departures are defined as the modelled minus the observations concentrations.

Figure 5 displays more statistics as a function of latitude. Distinction is made between the impact of the inverted fluxes and the impact of the adjusted concentrations at the initial time step of the inversion window. Changes in the RMS appear to be mostly caused by changes in the biases, indicating that consistent increments are generated throughout the 11-month inversion window despite the relatively high temporal resolution (eight days) of the flux vector. The adjusted concentrations at the initial time step of the inversion window seem to have a detrimental impact in the southern hemisphere. The inverted fluxes have a negligible impact south of 20S, a mixed effect in the tropical band and a mostly positive impact in the high northern latitudes. North of 40N, improvements of a few tens of a ppm are seen at many stations, which is significant in comparison to the prior departures (also shown in Figure 5). The bias departures in these latitudes indicate that the model has too high concentrations and manifest the absence of terrestrial carbon sink in the prior fluxes, which are annually balanced over land. The inversion compensates by introducing a 1 Gt C sink spread between Northern America and Europe.

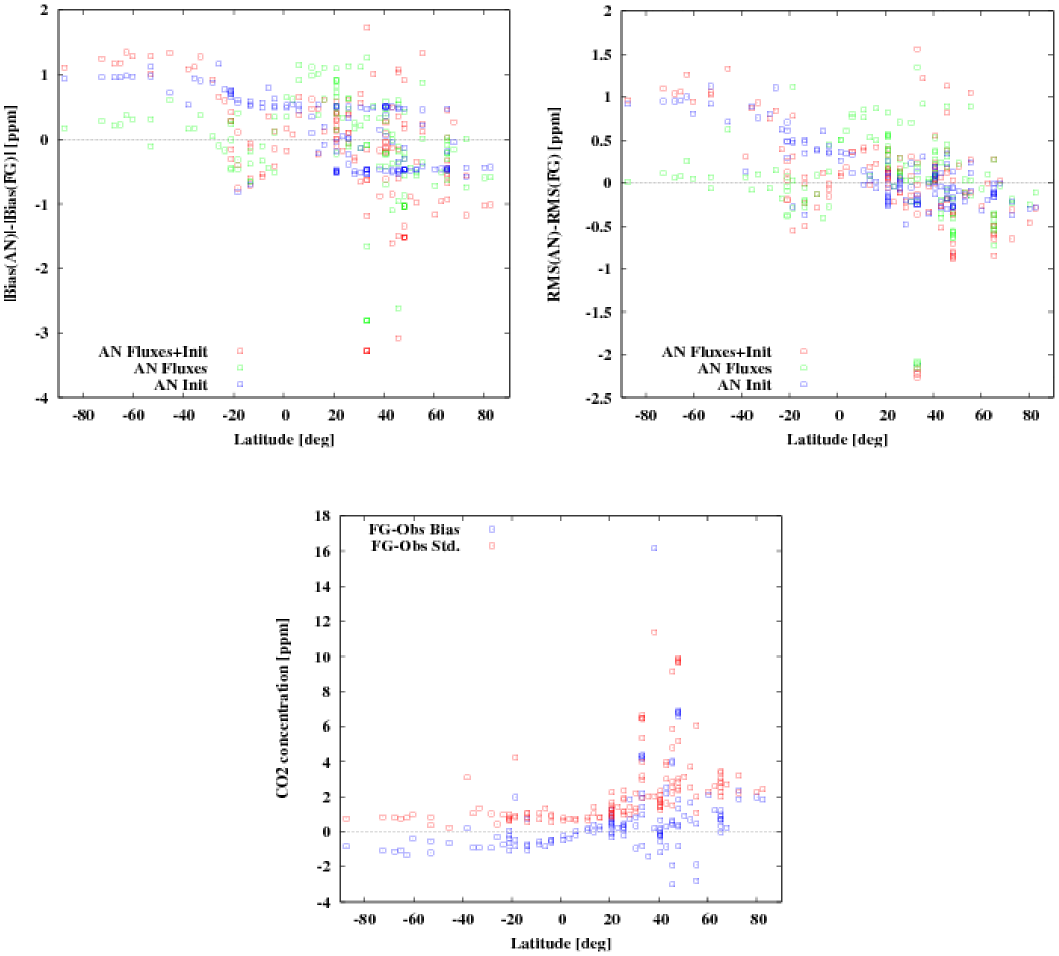

Figure 5. The bottom panel shows the prior departure statistics as a function of latitude, comparing the LMDZ simulation to the GLOBALVIEW dataset. The change in the absolute value of the biases and in RMS brought by the inverted fluxes (AN Fluxes), the inverted concentrations at the initial time step of the inversion window (AN Init) and by both together (AN Fluxes+Init) is shown in the top panels.

5. Conclusions

The carbon cycle of the Earth system consists of the exchange of huge amounts of carbon compounds between the atmosphere, the ocean, the biosphere, and the fossil reservoirs (several tens of Gt C per year). Since the cycle is nearly stationary on a yearly timescale, the annual global net flux at the interface between the atmosphere and the Earth is close to zero, with a relatively small gain for the atmosphere (about 3 Gt C per year, mainly of CO_2). This slight imbalance feeds back on another near-balanced budget, that of the energy exchange between the Earth system and outer space, via radiation processes. The importance of the topic has triggered numerous efforts to better quantify the carbon surface fluxes at all spatial scales.

The preliminary results presented here seem to indicate that some information about the northern mid-latitude carbon sink can be extracted from the AIRS radiances. The GEMS system which combines a 4D-Var analysis and a variational inversion system is given as a practical way to achieve this target. The interest of such systems for carbon studies will dramatically increase with the coming launch of the first CO_2-dedicated satellite instruments: the Orbiting Carbon Observatory (OCO, Crisp et al. 2004) and the Greenhouse gases Observing Satellite (GOSAT, Inoue et al. 2004), both expected for the end of 2008 or early 2009.

Similar inversion strategies are being implemented for aerosols (Dubovik et al. 2007) and reactive gases (Pétron et al. 2003, Meirink et al. 2006).

Adding emission models in the observation operator seems to be the next step in the development of such systems following the pioneering study of Rayner et al. (2005). Indeed such an approach allows one to extend the range of useful observations beyond the only measurements of atmospheric concentrations, for instance to vegetation observations (e.g., Demarty et al. 2007) or local flux measurements (e.g., Santaren et al. 2007).

Acknowledgements

We are indebted to Tony Hollingsworth (ECMWF) whose scientific vision made this study possible and whose pioneering spirit will be dearly missed. We would also like to thank the partners of the GEMS project, in particular Peter Rayner (LSCE) who coordinates the carbon activities. GEMS is funded by the European Commission under the EU Sixth Research Framework Programme, contract number SIP4-CT-2004-516099.

References

Bousquet, P., P. Peylin, P. Ciais, C. Quere, P. Friedlingstein, and P. Tans (2000), Regional changes in carbon dixiode fluxes of land and oceans since 1980, *Science*, **290**, 1342-1346.

Chédin, A., S. Serrar, N. A. Scott, C. Crevoisier, and R. Armante (2003), First global measurement of midtropospheric CO_2 from NOAA polar satellites: Tropical zone, *J. Geophys. Res.*, **108**, doi:10.1029/2003JD003439.

Chevallier, F., R. J. Engelen, and P. Peylin (2005a), The contribution of AIRS data to the estimation of CO_2 sources and sinks. *Geophys. Res. Lett.*, **32**, L23801, doi:10.1029/2005GL024229.

Chevallier, F., M. Fisher, P. Peylin, S. Serrar, P. Bousquet, F.-M. Bréon, A. Chédin, et P. Ciais (2005b), Inferring CO2 sources and sinks from satellite observations: method and application to TOVS data. *J. Geophys. Res.*, **110**, D24309, doi:10.1029/2005JD006390.

Chevallier, F., N. Viovy, M. Reichstein, and P. Ciais (2006), On the assignment of prior errors in Bayesian inversions of CO2 surface fluxes, *Geophys. Res. Lett.*, **33**, L13802, doi:10.1029/2006GL026496.

Courtier, P., J.-N. Thépaut, and A. Hollingworth (1994), A strategy for operational implementation of 4D-Var using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387

Crevoisier C., S. Heilliette, A. Chédin, S. Serrar, R. Armante, and N. A. Scott (2004), Midtropospheric CO_2 concentration retrieval from AIRS observations in the tropics, *Geophys. Res. Lett.*, **31**, doi:10.1029/2004GL020141.

Demarty, J, F. Chevallier, A. D. Friend, N. Viovy, S. Piao and P. Ciais (2007), Assimilation of global MODIS leaf area index retrievals within a terrestrial biosphere model. *Geophys. Res. Lett.*, **34**, L15402, doi:10.1029/2007GL030014.

Dethof, A.; Holm, E.V. (2004), Ozone assimilation in the ERA-40 reanalysis project, *Q. J. R. Meteorol. Soc.*. **130**, 603B, 2851-2872.

Dubovik, O., Lapyonok, T., Kaufman, Y. J., Chin, M., Ginoux, P., and Sinyuk, A.(2007), Retrieving global sources of aerosols from MODIS observations by inverting GOCART model, *Atmos. Chem. Phys*. Discuss., **7**, 3629-3718.

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 4D-Var data assimilation system: methodology and first results. *J. Geophys. Res*., **109**, D19309, doi:10.1029/2004JD004777.

Engelen R. J., 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.

Hourdin, F., I. Musat, S. Bony, P. Braconnot, F. Codron, J.-L. Dufresne, L. Fairhead, M.-A. Filiberti, P. Friedlingstein, J.-Y. Grandpeix, G. Krinner, P. LeVan, Z.-X. Li and F. Lott (2006), The LMDZ4 general circulation model : climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dynamics*, **27**, 787-813.

GLOBALVIEW-CO2 (2006), Cooperative Atmospheric Data Integration Project– Carbon Dioxide [CD-ROM], NOAA Clim. Monit. and Diagn. Lab., Boulder, Colo. (Also available via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW)

Gurney, K. R., R. M. Law, A. S. Denning, P. J. Rayner, D. Baker, P. Bousquet, L. Bruhwiler, Y. H. Chen, P. Ciais, S. Fan, I. Fung, M. Gloor, M. Heimann, K. Higuchi, J. John, T. Maki, S. Maksyutov, K. Masarie, P. Peylin, M. Prather, B. C. Pak, J. Randerson, J. Sarmiento, J., S. Taguchi, T. Takahashi, and C. W. Yuen (2002), Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models, *Nature*, **415**:6872, 626-630.

Matricardi, M., F. Chevallier, G. Kelly, and J.-N. Thépaut (2004), An improved general fast radiative transfer model for the assimilation of radiance observations, *Q. J. R. Meteorol. Soc*., **130**, 153-173, doi:10.1256/qj.02.181.

Meirink, J. F., Eskes, H. J., and Goede, A. P. H. (2006), Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modelling, *Atmos. Chem. Phys*., **6**, 1275-1292.

Pierangelo, C., A. Chédin, P. Chazette (2004), Measurements of stratospheric volcanic aerosol optical depth from NOAA TIROS Observational Vertical Sounder (TOVS) observations, *J. Geophys. Res*., **109**, D03207, doi:10.1029/2003JD003870

Pétron G., C. Granier, B. Khattotov, V. Yudin, J.-F. Lamarque, L. Emmons, J. Gille, D. P. Edwards (2004), Monthly CO surface sources inventory based on the 2000\u20132001 MOPITT satellite data, *Geophys. Res. Lett*., **31**, L21107, doi:10.1029/2004GL020560.

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, *Glob. Biogeochem. Cycles*, **11**(4), 535-560.

P. Rayner, M. Scholze, W. Knorr, T. Kaminski, R. Giering, and H. Widmann, (2005) Two decades of terrestrial Carbon fluxes from a Carbon Cycle Data Assimilation System (CCDAS). *Global Biogeochemical Cycles*, **19**:doi:10.1029/2004GB002254.

Santaren, D., P. Peylin, N. Viovy, and P. Ciais (2007), Optimizing a process-based ecosystem model with eddy-covariance flux measurements: A pine forest in southern France, *Global Biogeochem. Cycles*, **21**, GB2013, doi:10.1029/2006GB002834.

Trémolet, Y. (2005), Diagnostics of linear and incremental approximations in 4D-Var, *Quart. J. Roy. Meteor. Soc.*, **130**, 2233-2251.

Turner, D. S. (1993), The Effect of Increasing CO2 Amounts on TOVS Longwave Sounding Channels, *J. Appl. Meteor.*, **32**, 1760-1766.

Turner, D. S. (1994), HIRS Sensitivity to CO2 Mixing Ratio and a Pragmatic Correction Term, *J. Appl. Meteor.*, **33**, 1155-1162.