Four-dimensional data assimilation of atmospheric CO 2 using AIRS observations

Richard Engelen, Soumia Serrar, Frederic Chevallier

Four-dimensional data assimilation of atmospheric CO_2 using AIRS observations

Richard J. Engelen,[1] Soumia Serrar,[1] and Frédéric Chevallier[2]

Received 7 July 2008; revised 15 October 2008; accepted 2 December 2008; published 7 February 2009.

[1] The European Global and regional Earth-system (Atmosphere) Monitoring using Satellite and in situ data (GEMS) project has built a system that is capable of assimilating various sources of satellite and in situ observations to monitor the atmospheric concentrations of CO_2 and its surface fluxes. This consists of an atmospheric four-dimensional variational data assimilation system that provides atmospheric fields to a separate variational flux inversion scheme. In this paper, we describe the atmospheric data assimilation system that currently uses radiance observations from the Atmospheric Infrared Sounder (AIRS) to constrain the CO_2 mixing ratios of the data assimilation model. We present the CO_2 transport model, the bias correction of the observation-model mismatch, and the estimation of the background error covariance matrix. Data assimilation results are compared to independent CO_2 observations from NOAA/ESRL aircraft showing a reduction of the mean difference of up to 50% depending on the altitude of the aircraft observations relative to an unconstrained transport model simulation. In the coming years, observations from dedicated CO_2 satellite missions will be added to the system. Together with improved error characterization and bias correction, we hope to show that satellite observations can indeed complement the in situ observation system to get a better estimate of global carbon fluxes.

Citation: Engelen, R. J., S. Serrar, and F. Chevallier (2009), Four-dimensional data assimilation of atmospheric CO_2 using AIRS observations, *J. Geophys. Res.*, *114*, D03303, doi:10.1029/2008JD010739.

1. Introduction

[2] Over the last several years considerable effort has been put in extracting information about atmospheric CO_2 from infrared satellite sounders. The main driving force has been the potential improvement of atmospheric flux inversions, which are still limited by the amount of available accurate observations. Significant progress has been made in expanding the surface based observation networks as well as the airborne CO_2 observations, but most of these observations are still confined to developed countries, which leaves large gaps in for instance the tropics. Satellite data are well suited to fill these gaps, but the first dedicated CO_2 observing instruments will not be launched before the end of 2008. In the meantime, however, there is already a potential wealth of information through the various infrared sounding instruments. These instruments were mainly designed to observe atmospheric temperature, by assuming fixed concentrations of CO_2 in the CO_2 absorption bands, and water vapor. However, the same CO_2 absorption band can also be used to extract information about CO_2 itself. Various efforts have been made to extract this information, so far with mixed results. *Chédin et al.* [2008] showed CO_2 fire emission patterns in the Tropics by looking at the day-night differences of CO_2 estimates from the Television Infrared Observation Satellite (TIROS-N) Operational Vertical Sounder (TOVS), but *Chevallier et al.* [2005b] and *Peylin et al.* [2007] also showed that these TOVS retrievals are not good enough yet for surface flux inversions because of their significant regional biases. *Engelen and McNally* [2005], *Chahine et al.* [2005, 2008], and *Strow and Hannon* [2008] presented CO_2 estimates from the Advanced Infrared Sounder (AIRS) using a one-dimensional variational (1D-Var) data assimilation system, a classic retrieval scheme, and a least squares fit of brightness temperature departures, respectively, with promising results. However, *Chevallier et al.* [2005a] warned that great care should be taken with these CO_2 estimates when used in flux inversions. Both the quality of the satellite estimates and the quality of the transport models used in flux inversions should be watched very carefully to assure accurate flux estimates.

[3] As part of the GEMS (Global and regional Earth-system (Atmosphere) Monitoring using Satellite and in situ data) project [*Hollingsworth et al.*, 2008], a data assimilation system to monitor atmospheric concentrations of CO_2 and CH_4 and their fluxes has been built. The system consists of a four-dimensional variational (4D-Var) atmospheric data assimilation system run at the European Centre for Medium-Range Weather Forecasts (ECMWF), a variational CO_2 flux inversion system run at Laboratoire des Sciences du Climat et de l'Environnement (LSCE), a variational CH_4 flux inversion system run at the Joint Research Centre (JRC),

[1] European Centre for Medium-Range Weather Forecasts, Reading, UK.
[2] Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France.

0148-0227/09/2008JD010739

an independent CO_2 and CH_4 neural network retrieval system run at Laboratoire de Météorologie Dynamique (LMD), and a validation effort with independent models and in situ observations run at the Max Planck Institüt–Jena (MPI). This paper describes the atmospheric CO_2 4D-Var system at ECMWF, which is based on the earlier work described by *Engelen and McNally* [2005], but comprises now a full 4D-Var setup for both the meteorology and the tracers. The aim is to assimilate observations from various satellite instruments, such as AIRS (http://airs.jpl.nasa.gov), the Infrared Atmospheric Sounding Interferometer (IASI, http://smsc.cnes.fr/IASI/), the Orbiting Carbon Observatory (OCO, http://oco.jpl.nasa.gov), and the Greenhouse gases Observing Satellite (GOSAT, http://www.gosat.nies.go.jp), to obtain a consistent estimate of the atmospheric CO_2 concentrations. These atmospheric concentration fields are then subsequently used in an off-line surface flux inversion. This two-step approach was chosen over a direct flux inversion approach, such as currently being used for CO_2 flask inversions [e.g., *Gurney et al.*, 2002] or CH_4 inversions using retrievals from the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) instrument [e.g., *Meirink et al.*, 2008], for the following reasons. Firstly, an atmospheric data assimilation system based on systems developed for numerical weather prediction (NWP) has the capability of assimilating satellite observed radiances instead of retrieval products, which makes it easier to have consistent prior information among all instruments. The radiance assimilation is possible because the observations are processed in small chunks (typically 12 h intervals in our system), which makes the extra computations feasible, and because all the needed meteorological information (such as temperature and humidity) is assimilated at the same time. Secondly, the atmospheric data assimilation system can run at higher resolution than a direct flux inversion system that needs to process data over a much longer time frame (typically 1 year). This higher resolution allows for a better interpretation of the observations by reducing the representation error (difference in scale between a small observational field of view and a large grid box). For instance, synoptic weather systems or orography can create sharp gradients in CO_2 concentrations and therefore the observations. These gradients are better captured by a higher-resolution model. The resulting atmospheric fields can then be gridded to a feasible size for the subsequent flux inversion.

[4] The outline of the paper is as follows: we will first introduce the CO_2 transport model, the used AIRS observations, and the data assimilation system. Then the bias correction method and the background covariance matrix will be described. Finally, we will show results from the data assimilation with extensive comparisons to independent observations and, as noted above, briefly describe initial results from a flux inversion using the output from the CO_2 data assimilation.

2. CO_2 Transport Model

[5] A 4D-Var data assimilation system requires a prognostic transport model that will forecast the atmospheric state from specified initial conditions. At ECMWF, tracer transport has been introduced in the Integrated Forecasting System (IFS) using the existing modeling framework for advection, convection and vertical diffusion (ECMWF, IFS documentation CY31r1, 2007, http://www.ecmwf.int/research/ifsdocs/CY31r1/). For tracers, the conservation of monotonicity and positive definiteness in the semi-Lagrangian advection scheme is assured by using the so-called "quasi-monotone" interpolations [*Hortal*, 1994]. Although this advection scheme is not fully mass conserving, the gain of mass is lessened when the surface pressure is constrained by the analyses at the beginning of each forecast. A tracer mass fixer has also been coded in the IFS (to be used in long forecast integrations) but has not been used here as the gain of mass is not significant within the framework of a data assimilation system in which the surface pressure is constrained by observations.

[6] Various climatologies were used to prescribe the CO_2 fluxes at the surface. These are based on what was available at the start of the described experimentation, but are subject to updating in case better surface fluxes become available. The exchange with the terrestrial biosphere is described by a monthly mean climatology on which the short-term variability (diurnal cycle) has been superimposed. The monthly means derive from a climatological run of the CASA biosphere model [*Randerson et al.*, 1997] while short-term variability is generated by using time-specific information on Incoming Solar Radiation and 2 m Temperature provided by ECMWF 12 h forecasts, in a similar way as in the work by *Olsen and Randerson* [2004]. CASA fluxes are used at a 3-hourly resolution and have a zero annual mean everywhere. A linear interpolation was used to provide the fluxes at the transport model time step. More recently, we have also implemented natural biosphere fluxes from the ORCHIDEE model [*Krinner et al.*, 2005] driven by year specific meteorology. These ORCHIDEE fluxes are not annually balanced. Data assimilation results for both models will be presented in section 7 for 2003 to illustrate the dependence of the results on the choice of natural biosphere fluxes. However, the results in the current section are based on the use of CASA fluxes. The air-sea CO_2 exchange is described by a monthly mean climatology and is based on the revised version of *Takahashi et al.* [2002]. Anthropogenic emissions are based on the EDGAR 3.0 $1^\circ \times 1^\circ$ global map for 1990 [*Olivier and Berdowski*, 2001] rescaled to the Carbon Dioxide Information Analysis Centre (CDIAC) country level estimates for 1998. These emissions are kept constant throughout the years. Finally, wildfire emissions are from the Global Fire Emission Database version 2 (GFED2) [*van der Werf et al.*, 2006] and are provided at a 8-day resolution using MODIS fire hot spots [*Giglio et al.*, 2003]. These emissions are currently injected at the surface. Since the start of our analysis experiments some studies have been published to estimate more realistic injection heights [e.g., *Dentener et al.*, 2006; *Freitas et al.*, 2007], although there is still some debate about the utility of these injection heights [e.g., *Labonne et al.*, 2007]. The proposed methods will be compared and taken into account in the next version of our data assimilation and modeling system. All data sets were interpolated to the various horizontal resolutions used in the assimilation system. The IFS transport model has the capability to run at horizontal resolutions ranging from about 210 km by 210 km to about 25 km by 25 km. For our CO_2 data assimilation runs we use a reduced Gaussian grid with a

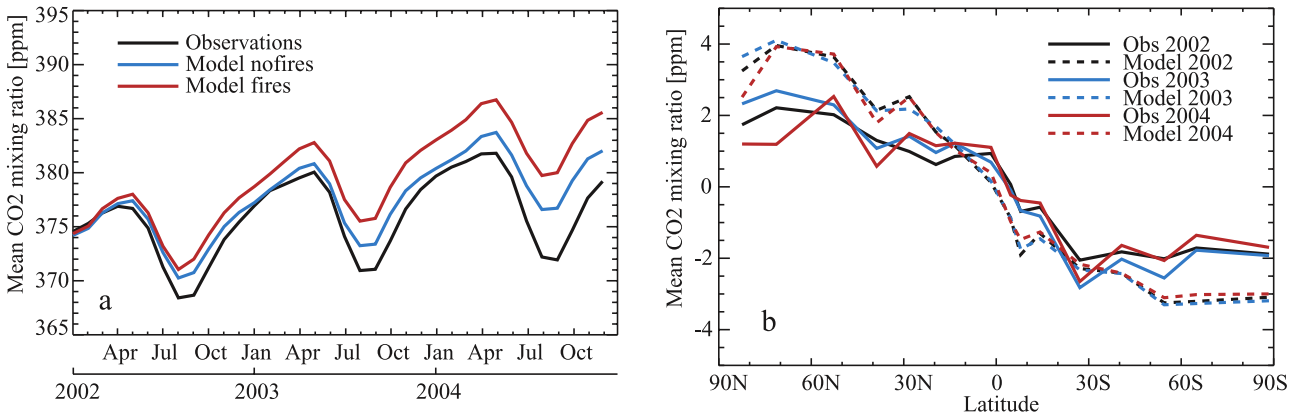

Figure 1. (a) Mean seasonal cycle of atmospheric CO_2 mole fraction as calculated from the average of eight surface locations in the Northern Hemisphere, observed in solid and modeled in dashed. The eight locations correspond to the eight flask stations Alert (ALT), Barrow (BRW), Shemya Island (SHM), Terceira Island (AZR), Sand Island (MID), Mauna Loa (MLO), Guam (GMI), and Christmas Island (CHR). (b) Mean north-south gradient of atmospheric CO_2 mole fraction constructed from 16 surface locations, observed in solid and modeled in dashed. The model is sampled at the 16 flask stations, from north to south: Alert (ALT), Barrow (BRW), Shemya Island (SHM), Terceira Island (AZR), Sand Island (MID), Mauna Loa (MLO), Guam (GMI) and Christmas Island (CHR), Mahe Island (SEY), Ascension Island (ASC), Tutuila (SMO), Easter Island (EIA), Cape Grim (CGO), Macquarie Island (MQA), Palmer Station (PSA), and South Pole (SPO).

resolution of approximately 125 km by 125 km on 60 sigma hybrid levels.

[7] The CO_2 transport model has been extensively compared to in situ observations both in terms of large-scale and short-scale atmospheric CO_2 variability. The baseline-air seasonal cycle at the surface is well described by the transport model, although its amplitude is underestimated in the Northern Hemisphere (Figure 1a). The difference to observations is larger during the NH summertime and is due to the annual carbon balance in CASA. This generates a larger annual trend in the modeled CO_2 than in the observations. Fire emissions were included in the IFS model to better describe the spatiotemporal variation of the atmospheric CO_2 concentration. Their inclusion, however, increases the trend unrealistically, because part of fire emissions is normally reincorporated into the biomass, likely in the tropics during the growth of the vegetation. This effect is not taken into account in the estimation of the net ecosystem flux of this version of the CASA model. Note that the global trend in atmospheric CO_2 is usually the easiest term for a flux inversion to correct.

[8] The modeled north-south (NS) gradient is overestimated by 3 to 3.5 ppm (Figure 1b). Again, deficiencies in the prescribed fluxes are likely to explain this result, although one cannot exclude a contribution from the transport model. The interannual variability shown by the observed NS gradient north of 50°N is not well reproduced by the transport model either. Since most of the CO_2 surface fluxes are prescribed as climatologies, any interannual variability in the IFS CO_2 model results can only be explained by the meteorology or fire emissions. These seem insufficient to explain the observed variability.

[9] Deficiencies in the transport modeling hamper the interpretation of the observed gradients in atmospheric CO_2 in terms of surface fluxes. Therefore, independent simulations with SF_6, for which the emissions are better known than for CO_2, have been carried out to investigate model behavior. They show an overestimation of the NS gradient by ~0.15 ppt (not shown). Although this could be an indication of too slow interhemispheric transport in our transport model, recent studies draw attention to the significant uncertainties in the SF_6 emissions database [*Hurst et al.*, 2006]. As most of these tracer sources are in the Northern Hemisphere, uncertainties in the emissions may translate into an incorrect meridional gradient. Therefore, these SF_6 comparisons cannot be conclusive in detecting errors in the transport modeling. We have also used a two-box three-dimensional model using SF6 as a tracer to estimate the interhemispheric time exchange in our model, similar to what is described by *Denning et al.* [1999]. We found a mean value of 1 year, which lies within the upper range of time scales (0.55 to 1.26) found during the TRANSCOM-2 experiment [*Denning et al.*, 1999] and agrees more closely with recent model results [*Lintner et al.*, 2004; *Rind et al.*, 2007; *Patra et al.*, 2008a].

[10] Finally, within the framework of the TransCom continuous data experiment, our transport model has been compared to other transport models and to high-frequency observations for the simulation of atmospheric CO_2 short-term variability. It has shown favorable results both for the simulation of the CO_2 diurnal cycle [*Law et al.*, 2008] and for describing the synoptic CO_2 variability [*Patra et al.*, 2008b]. The use of ECMWF meteorology at a relatively fine horizontal and vertical resolution is certainly an asset, although the relatively shallow surface layer makes the IFS more sensitive to potential errors in prescribed nocturnal fluxes. Also, the difficulties in resolving the nocturnal boundary layer itself create errors in the representation of boundary layer concentrations. These factors probably explain our model's tendency to overestimate the diurnal cycle

amplitude at a number of continental stations as was shown by *Law et al.* [2008].

3. AIRS Data

[11] The Atmospheric Infrared Sounder (AIRS) [*Aumann et al.*, 2003] was launched on board the NASA AQUA satellite in May 2002. After an initial period of testing, data were received operationally at ECMWF from October 2002 onward. AIRS is a grating spectrometer covering the 650–2675 cm^{-1} infrared spectral domain at a resolution of $\lambda/\Delta\lambda = 1200$, giving 2378 channels. The instrument flies onboard the Aqua satellite with equator crossing times of 0130 and 1330 local time. The AIRS field of view (fov) is 13 km at nadir with a 3×3 array of AIRS footprints falling into one AMSU-A fov. Because of bandwidth limits in the transatlantic line and other operational constraints ECMWF receives only 324 of the total 2378 channels in near real time and only 1 out of every 9 AIRS fovs within a AMSU-A fov.

[12] The channel selection is based on an original selection of 281 channels by NOAA/NESDIS appended with 43 extra channels in the two main CO_2 absorption bands based on the work by *Crevoisier et al.* [2003]. In our AIRS CO_2 reanalysis the number of channels was further reduced to avoid problems specific to certain spectral bands: (1) channels in the short-wave band were excluded from the analysis during local day time, because our radiative transfer model currently does not model solar radiation and the effects of nonlocal thermodynamic equilibrium; (2) channels in the main water vapor and ozone bands as well as channels sensitive to the surface over land and sea ice are excluded to minimize the effect of IFS model errors in water vapor, ozone, and the surface skin temperature on the CO_2 analysis; and (3) channels sensitive to the upper stratosphere were also excluded from the assimilation, because the IFS model has large temperature biases in the mesosphere of the polar winters. This way we have attempted to remove the known biases as much as possible. However, the cutoffs are somewhat arbitrary, because it is difficult to estimate the exact amplitude of these biases. The largest potentially remaining bias is caused by the mesospheric temperature bias in the IFS model and can reach values as high as 0.5 ppm. However, this bias only applies to the polar winter situation.

4. Four-Dimensional Variational Data Assimilation

[13] Our atmospheric 4D-Var data assimilation system is a practical formulation of Bayesian estimation theory for the particular case of a (near-)linear problem with unbiased Gaussian errors for time-evolving three-dimensional fields like temperature or CO_2 [*Lorenc*, 1986]. It seeks an atmospheric model trajectory that is statistically consistent with the information provided by the observations $\mathbf{y}^o$ available for the analysis time window $[t_0, t_n]$ and the information provided by an a priori atmospheric model state $\mathbf{x}^b$ called the background state. This background state is usually taken from a short-range forecast valid for time t_0. The atmospheric model trajectory within the assimilation window (the reference state $\mathbf{x}^r$) is then completely defined by the initial state $\mathbf{x}_0$ at time t_0, which is the same as the background state at the start of the minimization, and the boundary conditions (e.g., the prescribed CO_2 surface fluxes) through the use of the dynamical and physical forecast model.

[14] The analysis correction ($\delta\mathbf{x}(t_0)$) to the atmospheric model initial state is sought as a combination of the information from the observations and from the background using an objective cost function with two terms [e.g., *Courtier et al.*, 1994]:

$$J(\delta\mathbf{x}(t_0)) = \frac{1}{2}\delta\mathbf{x}(t_0)^T\mathbf{B}^{-1}\delta\mathbf{x}(t_0) + \frac{1}{2}\sum_{i=0}^{n}[\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i]^T \cdot \mathbf{R}^{-1}[\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i] \quad (1)$$

the background term and the observation term. The observation departures ($\mathbf{d}_i$) are the differences between the observed radiances and the atmospheric model simulated equivalent radiances, as in

$$\mathbf{d}_i = \mathbf{y}_i^o - \mathcal{H}_i[\mathbf{x}^r(t_i)] \quad (2)$$

where $\mathcal{H}_i$ is the full nonlinear observation operator in the form of the Radiative Transfer for the TIROS Operational Vertical Sounder (RTTOV) radiative transfer model. RTTOV [*Matricardi et al.*, 2004] is a fast radiative transfer model using profile-dependent predictors to parameterize the atmospheric optical depths. For the CO_2 assimilation experiments we applied the methods developed for RTIASI [*Matricardi*, 2003] to include CO_2 as a profile variable in RTTOV. $\mathbf{H}_i$, which appears in (1), is the tangent linear observation operator that is part of the RTTOV model. The reference state values ($\mathbf{x}^r$) at time t_i, needed for the calculation of the observation departures $\mathbf{d}_i$, are evolved according to the full nonlinear forecast model $\mathcal{M}$:

$$\mathbf{x}^r(t_i) = \mathcal{M}[\mathbf{x}^r(t_0)] \quad (3)$$

The increments themselves are evolved through time according to the tangent linear model $\mathbf{M}$:

$$\delta\mathbf{x}(t_i) = \mathbf{M}_i\delta\mathbf{x}(t_0) \quad (4)$$

Finally, $\mathbf{B}$ and $\mathbf{R}$ are the background error covariance matrix and the observation error covariance matrix, respectively. The estimation of the background matrix is described in section 6, while the estimation of the observation error matrix is described by *McNally et al.* [2006].

[15] The cost function is then minimized with respect to the increments of the initial state ($\delta\mathbf{x}(t_0)$). These increments are added to the background state to obtain the analysis $\mathbf{x}(t_0)$:

$$\mathbf{x}(t_0) = \mathbf{x}^b + \delta\mathbf{x}(t_0) \quad (5)$$

The CO_2 increments are not constrained directly by mass conservation, although the total atmospheric mass is constrained by the assimilation of surface pressure data.

[16] The advantage of a full data assimilation system is that it seeks to combine all available observations in an optimal way. At ECMWF, ground based and satellite based data are used to constrain the relevant fields in the forecast model, i.e., atmospheric temperature, vorticity, horizontal

wind divergence, surface pressure, normalized relative humidity, ozone, and in the configuration described here carbon dioxide as well. In addition to AIRS, which is the only instrument used to constrain CO_2, satellite data from various sensors are assimilated, such as the High Resolution Infrared Radiation Sounder (HIRS), the Advanced Microwave Sounding Unit (AMSU-A and AMSU-B), the Special Sensor Microwave/Imager (SSM/I), the Geostationary Operational Environmental Satellites (GOES), and the Meteosat instruments. in addition, many in situ observations are used, such as radiosondes, buoys, and surface station observations.

[17] As mentioned in the Introduction, a NWP-like data assimilation setup was chosen to enable the best use of the various CO_2 related observations. The short assimilation time window enables the use of radiances instead of retrieval products, which is not feasible in long-term flux inversions. This is especially beneficial for the interpretation of thermal infrared data (e.g., AIRS and IASI) that strongly depend on atmospheric temperature, but will also play a role in the interpretation of near-infrared data (e.g., OCO and GOSAT) in the future. Another benefit is that the data assimilation can run at higher resolution than a long-term flux inversion, which reduces the representation errors in the presence of strong horizontal gradients in the CO_2 concentrations. The actual length of the data assimilation time window is a compromise between assimilating as much data as possible on the one hand, and computer cost and the effect of model errors on the other hand. At the moment we use the same length (12 h) as is used in the operational NWP system, which is based on the above mentioned compromise. Work is ongoing to evaluate the potential of weak-constraint 4D-Var, in which the analysis time window is much longer (in the order of 10 days) and an estimate for the model error covariance matrix is taken into account [*Fisher et al.*, 2005].

[18] The satellite data are thinned to reduce spatial correlations of the measurement errors and they also undergo a bias correction, which is described in the next section. The 4D-Var data assimilation system currently only uses radiance data that are not affected by clouds. The AIRS cloud detection is described by *McNally and Watts* [2003]. The scheme detects which AIRS channels are affected by clouds and removes those channels from the assimilation, while retaining the channels that are not affected by clouds. This allows use of AIRS data even where the field of view is cloudy. If there is high cloud, only stratospheric information will be assimilated, but, if there are low clouds only, a significant amount of tropospheric information can be used. Finally, some channels were removed from the analysis either because of instrumental problems or because of unaccounted errors in the observation operator. Main example of the latter is the removal during local day time of the short-wave 4.2 μm band, which is affected by solar radiation not modeled in the current version of RTTOV as well as nonlocal thermodynamic equilibrium effects.

5. Bias Correction

[19] An important part of any data assimilation system is the bias correction. Various sources of bias exist, such as systematic errors in the IFS transport model, systematic errors in the spectroscopy, and systematic instrument errors. Generally, an attempt is made to separate transport model biases from observation related biases. The latter category not only includes instrument biases, but also biases in the observation operator, such as the radiative transfer model. This separation is usually achieved by bias correcting to a certain baseline that is defined by the most accurate observations, such as temperature radiosondes. However, this is not always straightforward, either by the lack of sufficient accuracy in the baseline observations or by the lack of enough baseline observations. The former is for instance the case for atmospheric humidity, while the latter is the case for CO_2. The essential part of the bias correction is the bias correction model. This is a regression model that should explain the main components of the bias. Time-averaged first-guess departures (observation minus model forecast) are used to calculate the coefficients of the bias model either through an off-line least squares method or through an online method, such as variational bias correction (VarBC) [*Auligne et al.*, 2007]. These coefficients are then used for the bias correction of individual satellite observations with the chosen bias model.

[20] For CO_2 we have chosen the so-called gamma-delta model, which is described by *Watts and McNally* [2004]. It is based on the assumption that the main bias consists of systematic instrument errors that can be represented by a global mean offset in the observed brightness temperatures (δ) and of systematic errors in the radiative transfer modeling that can be modeled by a multiplier (γ) of the total optical depth as illustrated below in the radiative transfer equation for the atmospheric contribution to the observed radiance in channel i of a passive infrared sounding instrument:

$$R_a^i = \int_{p_s}^{0} B^i(T(p)) d\mathcal{T}^i(p) \tag{6}$$

where the channel transmission is defined as

$$\mathcal{T}^i(p) = \exp[-\gamma \int_{p}^{0} \kappa^i(p)\rho(p)dp] \tag{7}$$

and $B^i(T(p))$ is the Planck function of temperature T at pressure p. Both the absorption coefficient κ and the absorber amount ρ are a function of pressure as well. We estimated the γ and δ values for each AIRS channel using the VarBC method, although the same results could be obtained with an off-line least squares method. Table 1 lists the values of the estimated γ and δ values. The listed values were estimated from the bias in observation departures for the months of April and May. On the basis of comparisons with in situ data, these months are considered to have the smallest contribution of the transport model bias to the total bias in a global mean sense. Although the used bias model is designed to remove the bias components of the instrument and the radiative transfer model, there is always the possibility of correcting the CO_2 values themselves. However, the spectral pattern of the bias correction is significantly different from the spectral pattern of a CO_2

Table 1. Bias Correction Coefficients Used in the δ-γ Model

Central Wave Number (cm^{-1})	δ (K)	γ
699.4	0.13	1.000
699.7	0.12	1.008
701.1	0.12	1.002
701.3	0.05	1.006
702.7	0.12	1.005
703.0	0.07	1.008
704.2	0.10	0.993
704.7	0.17	0.995
705.0	0.13	1.000
706.4	0.09	0.989
707.3	0.04	0.988
708.1	0.06	0.994
709.0	0.23	0.996
709.9	0.43	0.983
710.7	0.31	0.985
711.3	0.20	0.991
711.6	0.13	0.985
712.2	0.29	0.982
2253.5	0.21	0.998
2258.3	0.09	1.010
2259.3	0.10	1.013
2260.2	0.11	1.017
2270.0	0.03	1.040
2282.8	0.12	1.026
2284.7	0.09	1.019
2387.2	0.91	0.956
2388.2	1.33	0.952
2389.1	1.83	0.984

perturbation, which lessens the effect. A simple first estimate shows that the maximum effect is less than 3 ppm variation in CO_2 errors between different air masses. This value is substantially reduced by the constraint of the background field, but this is difficult to estimate precisely. Therefore, validation of the analysis results will be crucial to ensure there are no artificial spatial gradients introduced by the bias correction. The problem is central to the use of any (CO_2) satellite data and will require substantial research when new satellite sensors come along.

6. Background Constraint

[21] The background constraint is used to limit the amount of possible solutions of the 4D-Var minimization and by doing so to stabilize the inversion. It also plays an important role in distributing the information from the observations. For the CO_2 problem we use measurements from the AIRS instrument, which observes in the infrared. The weighting functions of these measurements are generally quite broad and therefore contain limited information at high vertical resolution. The vertical error correlations specified in the background covariance matrix distribute the broad information of the observations in the vertical under the assumption that an error at a specific height is correlated with errors at other heights. If these error correlations are known exactly, it is accurate to use them for the vertical distribution of the satellite corrections. However, in reality the error correlations are only approximations. A similar rationale also applies to the horizontal distribution of the CO_2 information from the observations. The background covariance matrix defines how errors in the short-term forecast are correlated between grid boxes and this is used to make small CO_2 adjustments in the grid boxes surrounding the grid box containing the observation. In practice, this works as a smoothing operator on the relatively noisy CO_2 increments from individual observations. To illustrate the combined effect of the broad AIRS weighting functions and the specified background error correlations, Figure 2 shows the CO_2 increments (analysis values minus background values) for a single AIRS footprint using 24 of the available CO_2 sensitive channels for two different locations, tropical Africa on the left and midlatitude Asia on the right. The patterns are broad, both in the horizontal and the vertical. The increments show a much deeper vertical structure over Africa (Figure 2, left), which is caused by the

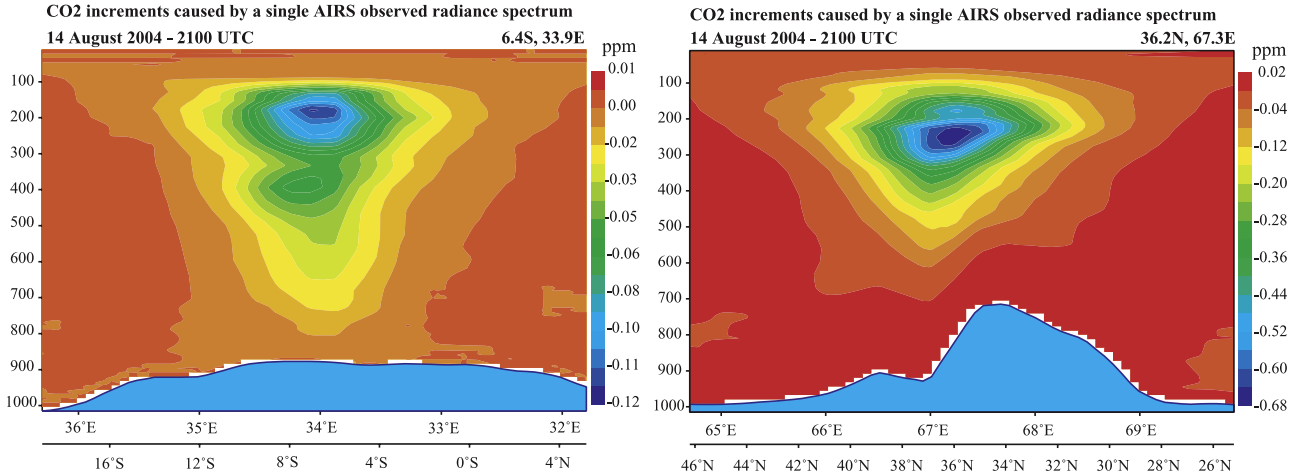

Figure 2. Vertical cross sections of CO_2 increments caused by a single AIRS observation (using 24 CO_2 sensitive channels) (left) over tropical Africa and (right) over midlatitude Asia.

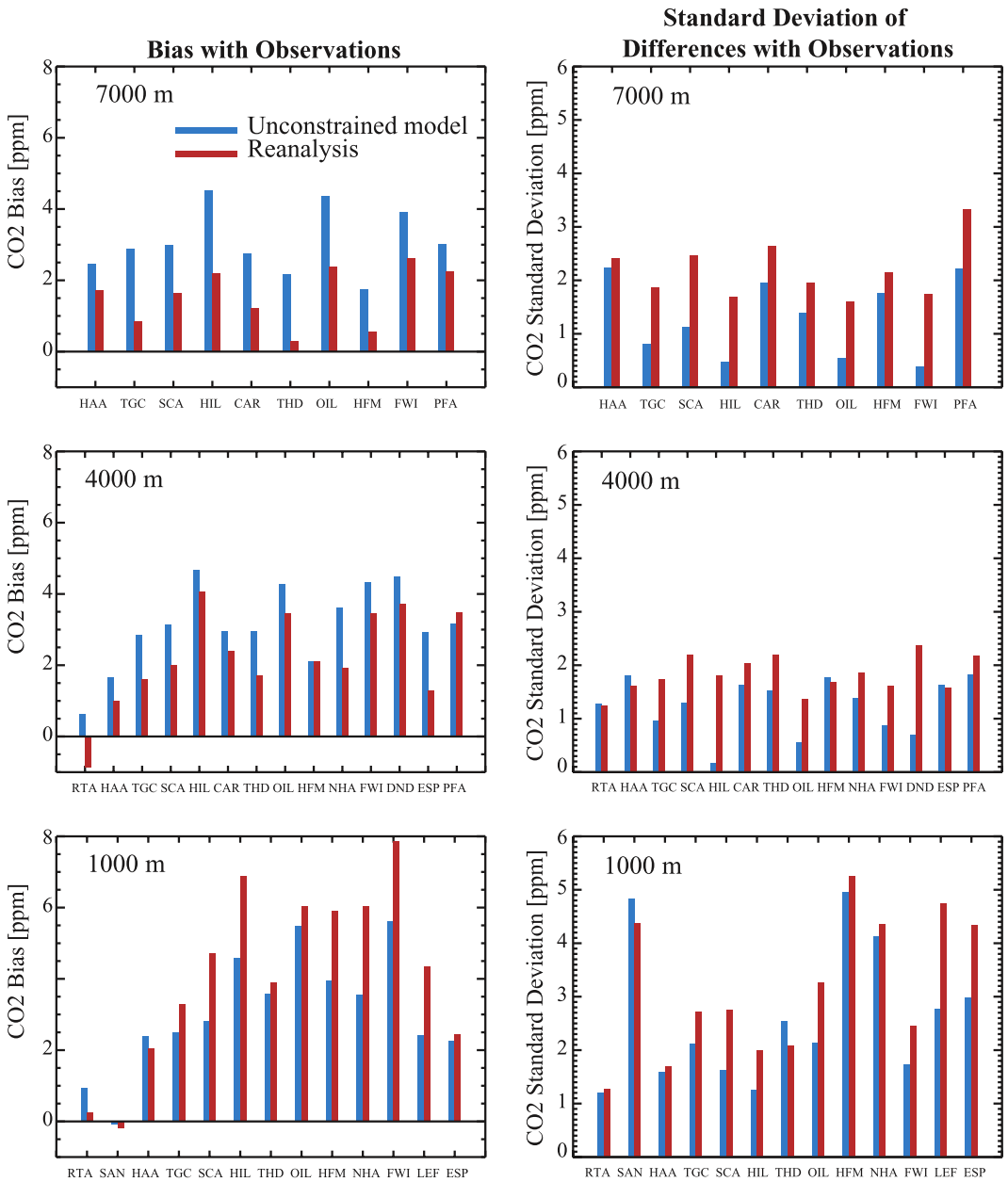

Figure 3. (left) Bias and (right) standard deviation of the difference of the unconstrained model run (blue) and the AIRS reanalysis (red) relative to independent flight observations from the NOAA/ESRL network.

larger temperature lapse rate. This means that information from the observations is used closer to the surface.

[22] To estimate the CO_2 background covariance matrix we used the method developed by *Parrish and Derber* [1992] (also known as the NMC method). This method consists of taking the differences between 24 h forecasts and 12 h forecasts valid at the same time (e.g., 1200 UTC) over a certain period (typically one month), and assumes that these differences represent a good sample for the background error. An extensive description of the application of this method to the new tracer variables in the ECMWF data assimilation system is given by *Benedetti and Fisher* [2007]. We used the horizontal and vertical correlations directly from this method, but inflated the standard deviations of the errors by a factor of 8 with a maximum of 15 ppm to account for the uncertainties in the prescribed surface fluxes. The inflation factor was based on comparisons between the CO_2 model concentrations and independent surface and aircraft observations. The mathematical formulation of the background error covariance matrix **B** is based on the work by *Fisher* [2003, 2004, 2006].

7. Reanalysis Results and Validation

[23] Assessing the quality of a complex system like a 4D-Var is critical. Its various components have to be carefully evaluated. As described in section 2, the transport model has been extensively compared to other CO_2 transport models

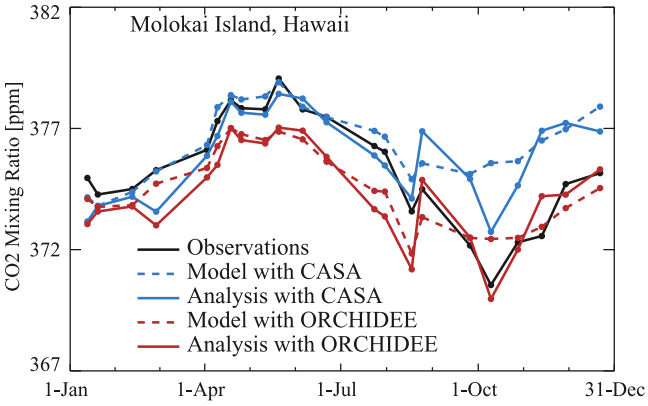

Figure 4. Times series of layer averaged (2–6 km) CO_2 mixing ratio over Molokai Island, Hawaii, for 2003. The black curve shows the aircraft observations from the NOAA/ESRL flights. The blue lines represent the unconstrained model run (dashed) and reanalysis run (solid) using the CASA fluxes, and the red lines represent the unconstrained model run (dashed) and reanalysis run (solid) using the ORCHIDEE fluxes.

and surface flask data [e.g., *Law et al.*, 2008]. However, the assimilation of AIRS radiances has the largest impact in the free troposphere. Therefore, other means of validation had to been found. The National Oceanic and Atmospheric Administration/Global Modeling Division (NOAA/ESRL) has compiled a data set of flask samples collected from profiling aircraft at various locations [*Tans*, 1996; C. Sweeney, NOAA/ESRL, personal communication, 2008; see also http://www.esrl.noaa.gov/gmd/ccgg/aircraft.html]. The profiles usually observe the atmosphere between the surface and about 8 km altitude, which is more appropriate to assess the impact of AIRS on the CO_2 fields than the surface flasks. Because the flight profile data have much higher accuracy than the results from our CO_2 transport model and assimilation system, we assumed them to represent the true atmospheric state.

[24] In the comparisons shown below, we have tried to reduce the amount of averaging to make the validation tests as demanding as possible. For every measured flight profile in the period January 2003 till December 2004 we have extracted profiles from an unconstrained CO_2 model run and the AIRS reanalysis. The unconstrained model run uses the operational ECMWF analyses to transport CO_2 around starting from the same initial field on 1 January 2003 as the AIRS reanalysis. However, there is no observational constraint on CO_2 during this model run. The AIRS reanalysis uses all available observations to both constrain the meteorology and the CO_2 fields. For the spatial interpolation to the profile location we used the nearest grid box and for the temporal interpolation we used a simple linear interpolation between the closest 6-h analysis fields. The vertical interpolation from the model pressure levels to the observation altitudes was done using the hydrostatic equation. Time series were then created at 1000 m intervals for each station. For each time series the mean difference (bias) between the unconstrained model simulation and the observations and between the reanalysis and the observations was calculated as well as the standard deviation of the differences. Figure 3 shows for three altitudes (1000 m, 4000 m, and 7000 m) these bias and standard deviation values for all stations with sufficient data. Figure 3 shows there is no significant change at 1000 m (Figure 3, bottom) between the unconstrained model and the AIRS reanalysis, both in bias and standard deviation. This is not surprising, because the AIRS sensitivity to CO_2 is very low at this level. Therefore, any information from the observations can only change CO_2 concentrations at this level through the transport or through the information spreading of the background covariance matrix. The latter is most likely not optimal and will therefore spread the information incorrectly. At 4000 m there is already a significant improvement in bias visible using the AIRS data and at 7000 m this improvement is very clear. The bias in the unconstrained model and the remaining bias in the reanalysis are probably mainly caused by the incorrectly specified surface fluxes. An insufficiently strong biospheric sink is causing a trend in the CO_2 concentrations that is stronger than observed. Although the AIRS observations do partially correct this anomalous trend, they are not able to fully correct it. This is most likely a result of a continuous incorrect forcing at the surface (the incorrect fluxes) that is not significantly corrected by the observations where it matters most (in the boundary layer). To obtain a better idea of the influence of the prescribed surface fluxes on the result of the reanalysis we ran a reanalysis experiment for 2003 using natural biosphere fluxes from the ORCHIDEE model instead of the CASA model, as was already mentioned in section 2. Figure 4 shows time series for Hawaii of the mean CO_2 mixing ratio of the layer between 2000 m and 6000 m. The black line represents the aircraft observations, the blue lines represent the unconstrained model run (dashed) and reanalysis run (solid) using the CASA fluxes, and the red lines represent the unconstrained model run (dashed) and reanalysis run (solid) using the ORCHIDEE fluxes. Figure 4 illustrates the anomalous trend in CO_2 mixing ratios when using the CASA fluxes, which is not being fully corrected by the observations. This anomalous trend is not visible in the runs with the ORCHIDEE fluxes that are not annually balanced. However, the ORCHIDEE fluxes show a summer uptake of CO_2 that is too large. The comparison shows that the prescribed fluxes have a significant impact on the results, but each database has its own advantages and disadvantages. Another option would be to use optimized fluxes from surface flask inversions. However, current comparisons of these flux distributions within the TRANSCOM community show that the differences between these flux distributions are still very large (see for instance http://inversions.lsce.ipsl.fr). Using the flask inversions as input would also make it more difficult to test our system against independent observations. On the long term, however, the assimilation system will certainly be used to combine the information from both satellite and in situ observations.

[25] Although the reduction of the bias by analyzing AIRS observations is significant, it seems to come at a cost. While the bias is reduced, the standard deviation of the differences with the profile observations is increased. The exact reason for this increase is difficult to pin down, but several potential causes can be identified. Firstly, the observations have limited capability in adjusting details of

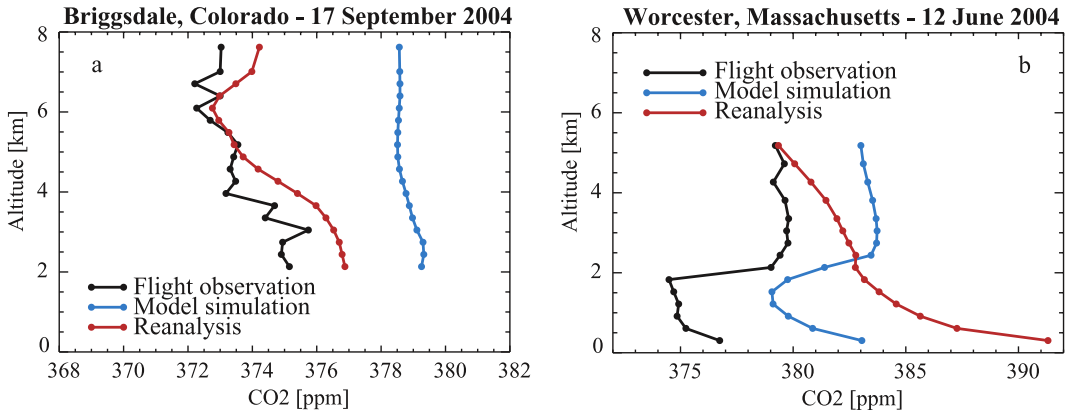

Figure 5. Observed (black), modeled (blue), and reanalyzed (red) profiles of CO_2 (a) for 17 September over Briggsdale, Colorado, and (b) for 12 June 2004 over Worcester, Massachusetts. Flight data were provided by NOAA/ESRL.

a CO_2 profile. The vertical sensitivity functions of the radiance observations are already quite broad and the overlap among the channels is large as well [see, e.g., *Engelen and McNally*, 2005, Figure 1]. Furthermore, this instrument sensitivity function is convolved with the specified model background correlation structures as was described in the previous section. This broad convoluted increment pattern will often adjust the CO_2 values correctly at the level of the strongest satellite observation sensitivity, but at the same time adjust the CO_2 values above and below this level incorrectly. This will then create more variability which is transported around. Figure 5 shows two examples of this behavior by comparing observed flight profiles (black) with profiles extracted from the unconstrained model run (blue) and the AIRS reanalysis (red). Figure 5a shows profiles for 17 September 2004 over Briggsdale, Colorado. The reanalysis is clearly able to reduce the bias with the observations compared to the unconstrained model run. But at the same time the profile has more variability than either the model run or the observations. Figure 5b shows profiles for 12 June 2004 over Worcester, Massachusetts. In this case the reanalysis again tries to reduce the bias with the observations, but is only able to do so at higher altitudes. In the lower 2000 m it stays close to the unconstrained model profile. The reanalysis profile therefore increases the standard deviation of the difference with the observation.

[26] Another cause of increased variability could be due to heterogeneous sampling caused by cloud cover. AIRS is able to make relatively large adjustments to the CO_2 field in clear sky areas. All available channels are used in such areas and the model background errors are relatively large, reflecting the errors in the prescribed surface fluxes, leaving sufficient room for increments of up to 10 ppmv. However, in areas with clouds in the upper troposphere, AIRS is not able to make significant adjustments. This combination of clear and cloudy areas could result in localized adjustments that are then transported downwind, while in reality a much larger scale adjustment should have been made. This will introduce larger variability, while in the end the mean is much less affected. This effect is probably significant because the CO_2 model is biased (mainly through the deficiencies in the prescribed surface fluxes), which means the observations keep making relatively large adjustments instead of steering the model-observation merge to a certain equilibrium. An illustration is shown in Figure 6. Figure 6 (left) shows the CO_2 increments (change to the model at time step 0 of the 12 h assimilation window) over the continental United States for model level 32 (~260 hPa). The AIRS observations that triggered these increments are shown on top in the form of departures (observation value minus model simulated value) for a channel peaking around 260 hPa in its CO_2 sensitivity. Positive departure values mean that the observation brightness temperature is larger than the model simulated brightness temperature, which means that CO_2 should be reduced (assuming the atmospheric temperature stays constant). Negative values initiate the opposite reaction, an increase in CO_2. Because the AIRS observations were made close to the end of the assimilation window, the CO_2 increments had to be made upwind from the observations, which is illustrated by the wind vectors plotted on top. Figure 6 (right) shows the change in the model field closer to the actual AIRS observation time. The incremental patterns have been advected downwind and resemble very nicely the observation departure patterns as one could expect from a well-working 4D-Var system. However, at the same time, the assimilation system will not make any changes to the initial model forecast in areas that are not covered by observations because of for instance cloud cover as can for instance been seen in the top right of the plots. In principle, this should cause no problems if the specified background error correlations are correct for this geographical area and the weather pattern that is at hand. However, this can never be completely achieved. Therefore, more variability is being created, especially in a system that is not in full equilibrium between the model forecast and the observations.

[27] The ultimate test for any satellite CO_2 retrieval or data assimilation system is the subsequent flux inversion. As already described in section 1, we have built a variational flux inversion system within the GEMS project that is able to use the output from the atmospheric 4D-Var data assimilation [*Chevallier et al.*, 2005b]. A time window of several years is used for this flux inversion and the same prior fluxes are used as in the atmospheric 4D-Var to be as consistent as possible. Full details of the flux inversion

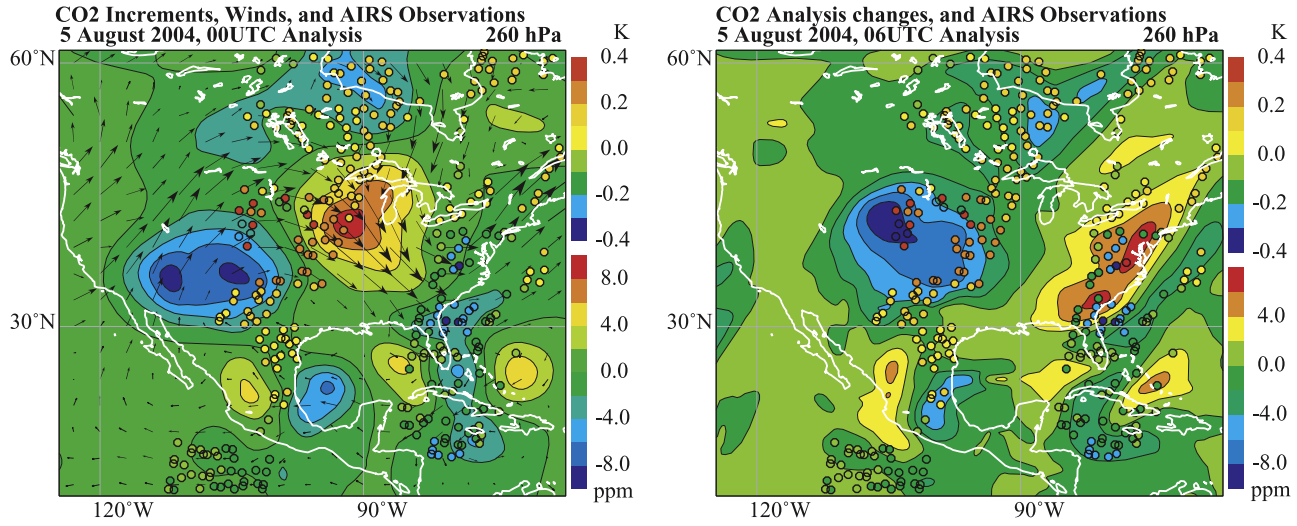

Figure 6. CO_2 increments at (left) 2100 UTC and (right) 0600 UTC at model level 32 (~260 hPa) for 5 August 2004 over North America. The CO_2 increments (analysis values minus forecast values) are shown as filled contours in ppm (bottom color scale). The observation departures (observation values minus equivalent model simulated values in degrees Kelvin) are shown as colored dots for an AIRS spectral channel peaking around 260 hPa for two consecutive AIRS overpasses (top color scale). Wind vectors (black arrows) for the same model level are overlaid in Figure 6 (left) to illustrate the main horizontal advection pattern.

including results will be described in a separate paper (F. Chevallier et al., AIRS-based versus flask-based estimation of carbon surface fluxes, manuscript in preparation, 2009). First results are encouraging, although using AIRS observations only is not sufficient to estimate accurate surface fluxes on regional scales. We think that our two-step system will come to its full potential when more accurate satellite observations will become available from the forthcoming OCO and GOSAT instruments.

8. Discussion and Outlook

[28] Within the European GEMS project we have built a system that is capable of assimilating various sources of satellite and in situ data to monitor the atmospheric concentrations of CO_2 and its surface fluxes and to improve our knowledge of these fluxes. The system consists of an atmospheric 4D-Var data assimilation system that provides atmospheric fields to a variational flux inversion system. The atmospheric assimilation system has been described in this paper and results have been compared with independent observations.

[29] Great care has been taken in setting up the transport model, the bias correction, and in defining the background covariance matrix. Although the current system performs well, work will continue to improve the above components of the system. Especially, the estimates of the error correlations needs improvement to make better use of the observations. Also, the specification of the surface fluxes will need to be improved.

[30] We are entering an exciting decade in which the assimilation system will come to its full potential being able to assimilate not only observations from the AIRS instrument, but also from the IASI, OCO, and GOSAT instruments. The system will provide consistent CO_2 fields that match all the observations within the respective errors, which will be a valuable source of information for flux inversions. However, although the first results from the flux inversions using our system are quite promising, it will take time to address all potential sources for systematic errors as carefully as possible. Only when we are able to remove these systematic errors in the satellite-based estimates will flux inversions really provide us with the needed regional information on carbon surface fluxes. This assumes, though, that systematic errors in tracer transport models, which affect flux inversions from the surface-based networks as well, will be addressed at the same time.

[31] It still remains open how the system will use in situ data. In the current setup, in situ data are used for validation of the atmospheric assimilation system and can therefore in principle still be used in the subsequent flux inversion step. This is currently under investigation. However, one could also envisage using the in situ data directly in the atmospheric assimilation. Especially, continuous surface observations should in principle have a strong enough constraint to be able to steer the data assimilation in the right direction. But in order to get full advantage from this type of observations a denser network is required. Indeed, only a few locations will not have the desired impact taking into account the enormous amount of satellite observations. Although the impact of the large amount of satellite data is partially compensated by their larger errors (compared to the very accurate in situ data), the satellite data are also more likely to suffer from systematic errors. The in situ data could be an important "anchor" to the system in case they are able to provide sufficient weight against the satellite observations.

[32] Apart from providing input to flux inversions, the system can also be used as a test bed for new model developments. Both improvements in transport modeling

and in surface flux modeling can be tested directly against the various sources of satellite information. It is already envisaged to test prescribed natural biosphere fluxes from the ORCHIDEE model [*Krinner et al.*, 2005] as well as inline modeling of these fluxes using the C-TESSEL model. The improvement or degradation of the fit to the observations will provide crucial information for the improvement of these models.

[33] **Acknowledgments.** The research presented in this paper was funded by EU FP6 project GEMS. The authors wish to thank the many people at ECMWF who helped build the tracer data assimilation system as well as the partners within the GEMS project. Many thanks also to P. Tans, C. Sweeney, and T. Conway (NOAA/ESRL) for providing the aircraft and surface flask data and for carefully reading the manuscript. Nicolas Viovy is acknowledged for providing the ORCHIDEE fluxes.

References

Auligne, T., A. P. McNally, and D. P. Dee (2007), Adaptive bias correction for satellite data in a numerical weather prediction system, *Q. J. R. Meteorol. Soc.*, *133*, 631–642, doi:10.1002/qj.56.

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 Sens.*, *41*, 253–264.

Benedetti, A., and M. Fisher (2007), Background error statistics for aerosol, *Q. J. R. Meteorol. Soc.*, *133*, 391–405, doi:10.1002/qj.37.

Chahine, M., C. Barnet, E. T. Olsen, L. Chen, and E. Maddy (2005), On the determination of atmospheric minor gases by the method of vanishing partial derivatives with application to CO_2, *Geophys. Res. Lett.*, *32*, L22803, doi:10.1029/2005GL024165.

Chahine, M., L. Chen, P. Dimotakis, X. Jiang, Q. Li, E. T. Olsen, T. Pagano, J. Randerson, and Y. L. Yung (2008), Satellite remote sensing of mid-tropospheric CO_2, *Geophys. Res. Lett.*, *35*, L17807, doi:10.1029/2008GL035022.

Chédin, A., N. A. Scott, R. Armante, C. Pierangelo, C. Crevoisier, O. Fosse′, and P. Ciais (2008), A quantitative link between CO_2 emissions from tropical vegetation fires and the daily tropospheric excess (DTE) of CO_2 seen by NOAA-10 (1987–1991), *J. Geophys. Res.*, *113*, D05302, doi:10.1029/2007JD008576.

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, and P. Ciais (2005b), Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data, *J. Geophys. Res.*, *110*, D24309, doi:10.1029/2005JD006390.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth (1994), A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. R. Meteorol. Soc.*, *120*, 1367–1387.

Crevoisier, C., A. Chedin, and N. A. Scott (2003), AIRS channel selection for CO_2 and other trace-gas retrievals, *Q. J. R. Meteorol. Soc.*, *129*, 2719–274.

Denning, A. S., et al. (1999), Three-dimensional transport and concentration of SF6: A model intercomparison study (TransCom 2), *Tellus, Ser. B*, *51*, 266–297.

Dentener, F., et al. (2006), Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for aerocom, *Atmos. Chem. Phys.*, *6*, 4321–434.

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.

Fisher, M. (2003), Background error covariance modelling, paper presented at Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean, Eur. Cent. for Med. Range Weather Forecasts, Reading, U. K., 8–12 Sept.

Fisher, M. (2004), Generalised frames on the sphere, with application to background error covariance modelling, paper presented at Seminar on Recent Developments in Numerical Methods for Atmospheric and Ocean Modelling, Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K., 6–10 Sept.

Fisher, M. (2006), "Wavelet" J_b—A new way to model the statistics of background errors, *ECMWF Newsl.*, *106*, 23–28.

Fisher, M., M. Leutbecher, and G. A. Kelly (2005), On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation, *Q. J. R. Meteorol. Soc.*, *131*, 3235–3246, doi:10.1256/qj.04.142.

Freitas, S. R., et al. (2007), Including the sub-grid scale plume rise of vegetation fires in low resolution atmospheric transport models, *Atmos. Chem. Phys.*, *7*, 3385–3398.

Giglio, L., J. Descloitres, C. O. Justice, and Y. Kaufman (2003), An enhanced contextual fire detection algorithm for MODIS, *Remote Sens. Environ.*, *87*, 273–282.

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

Hollingsworth, A., et al. (2008), Toward a monitoring and forecasting system for atmospheric composition, *Bull. Am. Meteorol. Soc.*, *89*, 1147–1164, doi:10.1175/2008BAMS2355.1.

Hortal, M. (1994), Recent studies of semi-lagrangian advection at ECMWF, *Tech. Memo.*, 204, Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K.

Hurst, D., et al. (2006), Continuing emissions of restricted halocarbons in the USA and Canada: Are they still globally significant?, *J. Geophys. Res.*, *111*, D15302, doi:10.1029/2005JD006785.

Krinner, G., et al. (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cycles*, *19*, GB1015, doi:10.1029/2003GB002199.

Labonne, M., F.-M. Bréon, and F. Chevallier (2007), Injection height of biomass burning aerosols as seen from a spaceborne lidar, *Geophys. Res. Lett.*, *34*, L11806, doi:10.1029/2007GL029311.

Law, R. M., et al. (2008), TransCom model simulations of hourly atmospheric CO_2: Experimental overview and diurnal cycle results for 2002, *Global Biogeochem. Cycles*, *22*, GB3009, doi:10.1029/2007GB003050.

Lintner, B. R., A. B. Gilliland, and I. Y. Fung (2004), Mechanisms of convection-induced modulation of passive tracer interhemispheric transport interannual variability, *J. Geophys. Res.*, *109*, D13102, doi:10.1029/2003JD004306.

Lorenc, A. C. (1986), Analysis methods for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, *112*, 1177–1194.

Matricardi, M. (2003), RTIASI-4, a new version of the ECMWF fast radiative transfer model for the infrared atmospheric sounding interferometer, *Tech. Memo.*, 425, Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K.

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.

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

McNally, A. P., P. D. Watts, J. A. Smith, R. Engelen, G. A. Kelly, J. N. Thépaut, and M. Matricardi (2006), The assimilation of AIRS radiance data at ECMWF, *Q. J. R. Meteorol. Soc.*, *132*, 935–957, doi:10.1256/qj.04.171.

Meirink, J. F., et al. (2008), Four-dimensional variational data assimilation for inverse modeling of atmospheric methane emissions: Analysis of SCIAMACHY observations, *J. Geophys. Res.*, *113*, D17301, doi:10.1029/2007JD009740.

Olivier, J. G. J., and J. Berdowski (2001), Global emissions, sources and sinks, in *The Climate System*, edited by J. Berdowski, R. Guicherit, and B. Heij, pp. 33–78, A. A. Balkema, Lisse, Netherlands.

Olsen, S. C., and J. T. Randerson (2004), Differences between surface and column atmospheric CO_2 and implications for carbon cycle research, *J. Geophys. Res.*, *109*, D02301, doi:10.1029/2003JD003968.

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

Patra, P. K., M. Takigawa, G. S. Dutton, K. Uhse, K. Ishijima, B. R. Lintner, K. Miyazaki, and J. W. Elkins (2008a), Transport mechanisms for synoptic, seasonal and interannual SF6 variations in troposphere, *Atmos. Chem. Phys.*, *8*, 12,737–12,767.

Patra, P. K., et al. (2008b), TransCom model simulations of hourly atmospheric CO_2: Analysis of synoptic-scale variations for the period 2002–2003, *Global Biogeochem. Cycles*, *22*, GB4013, doi:10.1029/2007GB003081.

Peylin, P., et al. (2007), Evaluation of Television Infrared Observation Satellite (TIROS-N) Operational Vertical Sounder (TOVS) spaceborne CO_2 estimates using model simulations and aircraft data, *J. Geophys. Res.*, *112*, D09313, doi:10.1029/2005JD007018.

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.

Rind, D., J. Lerner, J. Jonas, and C. McLinden (2007), Effects of resolution and model physics on tracer transports in the NASA Goddard Institute for Space Studies general circulation models, *J. Geophys. Res.*, *112*, D09315, doi:10.1029/2006JD007476.

Strow, L. L., and S. E. Hannon (2008), A 4-year zonal climatology of lower tropospheric CO_2 derived from ocean-only Atmospheric Infrared Sounder observations, *J. Geophys. Res.*, *113*, D18302, doi:10.1029/2007JD009713.

Takahashi, T., et al. (2002), Global sea-air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, *Deep Sea Res., Part II*, *49*, 1601–1622.

Tans, P. P. (1996), Carbon cycle (group report), in *Summary Report 1994–1995*, vol. 23, edited by D. J. Hoffman, J. Peterson, and R. M. Rosson, pp. 29–49, U. S. Dep. of Commer, Boulder, Colorado.

van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano Jr. (2006), Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmos. Chem. Phys.*, *6*, 3423–3441.

Watts, P. D., and A. P. McNally (2004), Identification and correlation of radiative transfer modelling errors for atmospheric sounders: AIRS and AMSU-A, paper presented at Workshop on Assimilation of High Spectral Resolution Sounders in NWP, Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K., 28 June to 1 July.

F. Chevallier, Laboratoire des Sciences du Climat et de l'Environnement, L'Orme des Merisiers, Bat 701, Point courrier 129, F-91191 Gif-sur-Yvette CEDEX, France. (frederic.chevallier@lsce.ipsl.fr)

R. J. Engelen and S. Serrar, European Centre for Medium-Range Weather Forecasts, Reading RG2 9AX, UK. (richard.engelen@ecmwf.int; soumia.serrar@ecmwf.int)