Atmos. Chem. Phys. Discuss., 8, 19917–19955, 2008
www.atmos-chem-phys-discuss.net/8/19917/2008/

Atmospheric Chemistry and Physics Discussions

This discussion paper is/has been under review for the journal *Atmospheric Chemistry and Physics (ACP)*. Please refer to the corresponding final paper in *ACP* if available.

Estimating surface CO_2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble Kalman Filter

L. Feng[1], P. I. Palmer[1], H. Bösch[2], and S. Dance[3]

[1]School of Geosciences, University of Edinburgh, King's Buildings, Edinburgh EH9 3JN, UK
[2]Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, UK
[3]Department of Mathematics and Department of Meteorology, University of Reading, Reading RG6 6BB, UK

Received: 8 September 2008 – Accepted: 30 September 2008 – Published: 28 November 2008

Correspondence to: L. Feng (lfeng@staffmail.ed.ac.uk)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

We have developed an ensemble Kalman Filter (EnKF) to estimate 8-day regional surface fluxes of CO_2 from space-borne CO_2 dry-air mole fraction observations (X_{CO_2}) and evaluate the approach using a series of synthetic experiments, in preparation for data from the NASA Orbiting Carbon Observatory (OCO). The 32-day duty cycle of OCO alternates every 16 days between nadir and glint measurements of backscattered solar radiation at short-wave infrared wavelengths. The EnKF uses an ensemble of states to represent the prior error covariance to estimate 8-day CO_2 surface fluxes over 144 geographical regions. We use a 12×8-day lag window, recognising that X_{CO_2} measurements include surface flux information from prior time windows. The observation operator that relates surface CO_2 fluxes to atmospheric distributions of X_{CO_2} includes: a) the GEOS-Chem transport model that relates surface fluxes to global 3-D distributions of CO_2 concentrations, which are sampled at the time and location of OCO measurements that are cloud-free and have aerosol optical depths <0.3; and b) scene-dependent averaging kernels that relate the CO_2 profiles to X_{CO_2}, accounting for differences between nadir and glint measurements, and the associated scene-dependent observation errors. We show that OCO X_{CO_2} measurements significantly reduce the uncertainties of surface CO_2 flux estimates. Glint measurements are generally better at constraining ocean CO_2 flux estimates. Nadir X_{CO_2} measurements over the terrestrial tropics are sparse throughout the year because of either clouds or smoke. Glint measurements provide the most effective constraint for estimating tropical terrestrial CO_2 fluxes by accurately sampling fresh continental outflow over neighbouring oceans. We also present results from sensitivity experiments that investigate how flux estimates change with 1) bias and unbiased errors, 2) alternative duty cycles, 3) measurement density and correlations, 4) the spatial resolution of estimated flux estimates, and 5) reducing the length of the lag window and the size of the ensemble.

1 Introduction

CO_2 surface fluxes inferred from atmospheric CO_2 concentrations by inverting models of atmospheric transport have led to substantial improvements in our quantitative understanding of the contemporary carbon cycle (e.g., Bousquet et al., 2000). Previous studies that employ these methods to estimate surface fluxes of CO_2 have tended to use accurate, but spatially sparse and heterogeneous, in situ measurements, which were not designed for the flux estimation problem, consequently limiting the extent of spatial disaggregation of fluxes that can be achieved (e.g., Houweling et al., 1999; Rödenbeck et al., 2003). Satellite measurements of CO_2 offer new constraints for estimating surface fluxes. The SCcanning Imaging Absorption spectroMeter for Atmospheric ChartographY (SCIAMACHY) satellite instrument (Bovensmann et al., 1999) has measured short-wave infra-red wavelengths (SWIR), with greatest sensitivity to CO_2 in the lower troposphere, since its launch in 2002. Current CO_2 column volume mixing ratio products from SCIAMACHY have an estimated measurement accuracy of between 1 and 5% (Schneising et al., 2008; Barkley et al., 2006, 2007). Uncharacterized systematic and random errors (e.g., Houweling et al., 2005), while the subject of ongoing research (Schneising et al., 2008), limit the application of these data for surface flux estimation. Top-down studies that use satellite measurements of CO_2 retrieved at thermal infra-red wavelengths, with greatest vertical sensitivity in the free troposphere, have concluded that uncharacterized observation and model biases compromise resulting surface flux estimates (Chevallier et al., 2005; Tiwari et al., 2006).

The NASA Orbiting Carbon Observatory (OCO) (Crisp et al., 2004) and the Japanese Greenhouse Observing SATellite (GOSAT) (Maksyutov et al., 2008), both due for launch in early 2009, will measure SWIR wavelengths, that are sensitive to CO_2 in the free and lower troposphere. OCO and GOSAT will operate two modes of observation: (1) nadir, and (2) glint, where the instrument boresight is directed off-nadir to the angle of specular reflection. The glint mode increases the signal to noise of measurements over the ocean. Dry-air CO_2 mole fractions (X_{CO_2}) will be retrieved from the

ACPD
8, 19917–19955, 2008

observed spectra to a precision of 1–2 ppm (Crisp et al., 2004), a level of precision necessary to improve upon constraints from existing in situ measurements (Rayner et al., 2002; Patra et al., 2003; Miller et al., 2007). We focus on OCO measurements of CO_2, but the general assimilation approach described here can easily be applied to GOSAT data.

Recent studies have used variational data assimilation methods with synthetic OCO observations to show that these data have the potential to estimate weekly and daily surface CO_2 fluxes at model grid scales of order $3.75^\circ \times 2.5^\circ$ (Baker et al., 2006; Chevallier et al., 2007a; Chevallier, 2007b). These studies (1) used a simple representation of measurements, assuming only glint-mode (Chevallier et al., 2007a) or making no distinction between nadir and glint observations (Baker et al., 2006); or (2) assumed a simple representation for measurement errors in which a uniform value is adopted (Chevallier et al., 2007a).

We have developed an Ensemble Kalman Filter (EnKF) (Evensen, 1994, 2003; Ehrendorfer, 2007) to estimate surface CO_2 fluxes from space-borne measurements of X_{CO_2} (Sect. 2). The EnKF, an independent and complementary approach to variational assimilation, has been developed in the physical oceanography and meteorology communities (e.g., Evensen, 1994; Houtekamer and Mitchell, 1998; Lorenc, 2003), and recently applied to carbon cycle research (Peters et al., 2005; Bruhwiler et al., 2005). The EnKF methodology we use is outlined in Sect. 3. We use the GEOS-Chem chemistry transport model to describe the relationship between surface CO_2 fluxes and 3-D atmospheric CO_2 concentrations, which are then sampled along the proposed OCO orbits and convolved with scene-dependent instrument averaging kernels as a function of observation modes, surface types, solar zenith angles, and optical depths (Sect. 2). This new, improved description of OCO measurements and their errors (Bösch et al., 2008, Sect. 2) is expected to provide more realistic descriptions of X_{CO_2} distributions, with which to infer more realistic flux estimates. We use the EnKF to explore the sensitivity of the surface flux inverse problem to changes in instrument configurations and the size of geographical regions over which fluxes are to be estimated (Sect. 4). We conclude the paper in Sect. 5.

2 Simulated OCO X_{CO_2} observations and uncertainties

The OCO instrument will be launched in early 2009 into the NASA EOS Afternoon Constellation (A-train), which is in a sun-synchronous polar orbit at an approximate altitude of 705 km. This orbit has 14.6 equator crossings per day, separated by 24.7° in longitude, resulting in a 16-day repeat cycle. OCO will have a local equatorial crossing time of 13:18. The OCO platform includes three, high-resolution grating spectrometers that measure absorptions of the reflected sunlight by using two CO_2 bands (1.61 and 2.06 μm) and the O_2 A-Band (0.765 μm) using nadir view geometry or glint view geometry in which the instrument will be pointed to the spot where solar radiation is specular reflected from the surface (Crisp et al., 2004). The 32-day duty cycle of OCO will alternate between 16-day cycles of nadir and glint modes.

We model OCO X_{CO_2} measurements in a four-step process, which constitutes the observation operator H that relates surface CO_2 fluxes to global distributions of X_{CO_2}. First, we use the GEOS-Chem chemistry transport model (v7-03-06) to relate surface fluxes to global 3-D CO_2 concentrations. For the purpose of these calculations we use the same flux inventory as described in Palmer et al. (2008) for 2003. For the sake of brevity, here we describe the model briefly and refer the reader to Appendix A and Palmer et al. (2008) for further details. We use a 2°×2.5° horizontal resolution for the experiments described here, with meteorological analyses from version 4 of the GEOS model from the NASA Goddard Global Modelling and Assimilation Office. We include CO_2 estimates for daily biospheric fluxes (Potter et al., 1993), monthly oceanic fluxes (Takahashi et al., 2002), monthly biomass burning fluxes from the second version of the Global Fire Emission Database (GFEDv2) for 2003, a climatological distribution of annual fossil fuel emissions that have been scaled to 2003 (Palmer et al., 2008), and climatological biofuel fluxes (Yevich and Logan, 2003).

Second, we sample the 3-D field of CO_2 concentrations at the time and the location of each nadir and glint measurement using the orbits of the Aqua satellite in 2006, which leads the A-train constellation with a local equatorial crossing time of 13:30. The retrievals are assumed to be made from binned scan pixels, so that two consecutive observations in one orbit are separated by 20 s, which corresponds to 1.2° in latitude.

Third, we use seasonal probability density functions (PDFs) of cloud and aerosol optical depths (AODs), derived from the MODIS and MISR instruments (Bösch et al., 2008), to remove cloudy scenes and scenes with AOD >0.3 that will not be retrieved, at least, initially from OCO. These restrictions remove about 50–60% of the daily available nadir measurements, and 60–70% of the daily glint measurements. Hereinafter, we refer to the resulting measurements as clear.

Finally, we apply scene-dependent averaging kernels, which account for the vertical sensitivity of OCO, to map from the 1-D CO_2 concentrations profiles to X_{CO_2} (Connor et al., 2008):

$$X_{CO_2} = X_{CO_2,a} + \mathbf{A}\left(\frac{1}{1-\mathbf{w}}\right)\left(M(\mathbf{x}^t) - \mathbf{f}_a\right). \quad (1)$$

Lower case variables in bold denote vectors and upper case variables in bold denote matrices. The subscript a denotes a priori; $M(\mathbf{x}^t)$ is the GEOS-Chem chemistry transport model driven by "true" surface fluxes of CO_2 ($\mathbf{x}^t$); $\mathbf{w}$ denotes GEOS-4 water mole fractions that are used to map from CO_2 concentrations to dry mole fraction; and $\mathbf{f}_a$ is climatological dry CO_2 mole fractions that will be used to retrieve CO_2 profile information from OCO, and $X_{CO_2,a}$ is the associated column amount. We use an annual zonal mean for $\mathbf{f}_a$. The column averaging kernel $\mathbf{a}$ is given by $\mathbf{t}^T\mathbf{A}$, where $\mathbf{A}$ is the averaging kernel, $\mathbf{t}$ is the column integration operator that integrates a vertical profile to a column and superscript T denotes the matrix transpose operation.

We use averaging kernels as a function of two view modes (nadir and glint), five surface types (snow, ocean, soil, conifer, and desert), ten solar zenith angles (SZA) (from 10° to 85° for nadir measurements, and from 10° to 72° for glint measurements), and seven AODs from 0 to 0.3 (Bösch et al., 2008).

Figure 1a and b shows averaging kernels for five different surface types at a solar zenith angle of 10° under a clear-sky with an AOD of 0.1. In general, OCO averaging kernels peak in the mid and lower troposphere. The instrument sensitivity to changes in CO_2 near the surface is particularly important for flux estimation. Using nadir view geometry, the oceans are relatively dark at the SWIR wavelengths measured by OCO, and the resulting averaging kernels below 400 hPa are lower than over other surface types. In contrast, glint view measurements over the oceans take advantage of specular reflection, resulting in a large signal to noise and an averaging kernel close to unity below 400 hPa.

The uncertainty associated with the simulated X_{CO_2} also depends on the scene characterization. Figure 1c and d shows observation errors over 5 different surface types as a function of SZA. The error over land is usually <0.5 ppm for a single nadir measurement at SZAs <40°, but increases with SZA, eventually reaching 1.2 ppm at a SZA of 85°. Observation errors for nadir measurements over ocean are typically >3.0 ppm for all SZAs. In contrast, the error for a single glint measurement over ocean is typically <0.4 ppm, smaller than nadir errors over land, which as we show later has important implications for flux estimation using these data.

Figure 2 summarises the resulting 2°×2.5° distribution and uncertainties of nadir and glint X_{CO_2} measurements between 17th January and 1st February 2003. Continental regions at mid and low latitudes typically have uncertainties of <0.2 ppm (Fig. 2b), well within the target precision of OCO (Crisp et al., 2004), while oceans have an observation error >3 ppm. Over tropical regions, scenes are frequently obscured by clouds during the wet seasons, and frequently obscured by smoke aerosol from biomass burning during the dry seasons. Glint measurements are generally restricted to a smaller latitude range (85° S to 55° N for the months shown) than nadir measurements because they are used over a small range of SZAs (<72° for glint versus <85° for nadir). There are typically less glint observations than nadir over the tropics due to their larger view spot (≃25 km^2 at the extreme view angles versus ≃3 km^2 for the nadir view). The larger view spot increases the probability of cloud obscuration. A similar method is used to define "model" X_{CO_2} distributions for the observation system simulation experiment (OSSE) in Sect. 4.

3 The Ensemble Kalman Filter

3.1 Basic formulation

We have developed an ensemble data assimilation system based on the Ensemble Transform Kalman Filter (ETKF) technique (Bishop et al., 2001) to simultaneously assimilate consecutive X_{CO_2} observations. At each assimilation cycle, we assimilate 8-day OCO observations $\mathbf{y}_{obs}$ to improve the prior estimation of regional surface CO_2 fluxes via:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}[\mathbf{y}_{obs} - H(\mathbf{x}^f)] \tag{2}$$

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T [\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R}]^{-1}, \tag{3}$$

where $\mathbf{x}^f$ is the a priori state vector; $\mathbf{x}^a$ is the a posteriori; H is the observation operator that describes the relationship between the state vector and the observations (Sect. 2); and $\mathbf{K}$ is the Kalman gain matrix that determines the adjustment to the a priori based on the difference between model and observations and their uncertainties. $\mathbf{R}$ is the observation error covariance matrix, and $\mathbf{P}^f$ is the a priori error covariance matrix. $\mathbf{H}$, the Jacobian of the observation operator H, maps $\mathbf{P}^f$ into observation space.

The observation error covariance $\mathbf{R}$ includes measurement (instrument + retrieval) error, model (transport) error and representation error (Peylin et al., 2002). Quantifying model error is non-trivial, and for simplicity we have assumed uniform model and representation errors: 2.5 ppm over land regions and 1.5 ppm over oceans (Rödenbeck et al., 2003), both of which are uncorrelated with the measurement errors. Also, we assume that $\mathbf{R}$ is either diagonal (i.e., no observation correlation, Sect. 4.1), or has a simple block structure for correlations between successive observations (Sect. 4.4).

Here, the state vector includes 8-day mean CO_2 surface fluxes from 144 regions that span the globe (Fig. 3), based on the 22 regions used by TransCom-3 (T3) (Law et al., 2003; Gurney et al., 2002). We divide each T3 land region into 9 near-equal areas (resulting in 9×11 land regions), and divide each T3 ocean region into 4 near-equal areas (resulting in 4×11 ocean regions). We have included one region (region 1) to represent all other low-emission regions, typically covered by snow. Recognising that observed CO_2 concentration variations contain a history of source/sink signatures, the state vector $\mathbf{x}$ consists of these regional fluxes at the current assimilation timestep and those from the previous 11 timesteps. This corresponds to solving the regional fluxes over a 96-day period (3×32-day OCO duty cycles), resulting in 12×144 control variables.

We estimate the associated forecast error ε_i for an individual geographical region i at the current assimilation step by rescaling the annual-mean T3 regional a priori error ε_{T3} (e.g., Patra et al., 2003) to an 8-day mean flux error over an area of 1/9 (for lands) or 1/4 (for oceans) size of the parent T3 region:

$$\varepsilon_i = \varepsilon_{T3} \left(\sqrt{\frac{365}{8}} \right) \left(\sqrt{\frac{A_l}{A_{T3}}} \right), \tag{4}$$

where the first bracketed term represents the scaling from the TransCom annual mean error to the 8-day period, and the second bracketed term represents the scaling from T3 area A_{T3} to the regional area A_l. Table 1 summarises the forecast errors for the regions 2 to 144. We assume that the snow region (region 1) has zero emissions with an uncertainty of 0.1 GtC yr^{-1}. We also assume no spatial correlation between these regional flux forecasts, so that their (sub) error covariance matrix is diagonal. The other 11×144 variables (for previous 8-day periods), having passed through a number of assimilation cycles, include spatial and temporal correlations.

To avoid computing the Jacobian of the observation operator H, the EnKF approximates the a priori error covariance by introducing an ensemble of perturbation states

$\Delta\mathbf{X}^f = [\Delta\mathbf{x}_1, \Delta\mathbf{x}_2, ..., \Delta\mathbf{x}_{N_e}]$ (Evensen, 1994), so that

$$\mathbf{P}^f = \Delta\mathbf{X}^f(\Delta\mathbf{X}^f)^T, \tag{5}$$

where we have absorbed the normalization factor $1/(N_e-1)$ into $\Delta\mathbf{X}^f$ (Zupansi, 2005). As a result, $\mathbf{K}$ can now be approximated by the ensemble gain matrix $\mathbf{K}_e$:

$$\mathbf{K}_e = \Delta\mathbf{X}^f(\Delta\mathbf{Y})^T[\Delta\mathbf{Y}(\Delta\mathbf{Y})^T + \mathbf{R}]^{-1}, \tag{6}$$

$$\Delta\mathbf{Y} = H(\Delta\mathbf{X}^f). \tag{7}$$

In the above equations, we have made use of $H(\Delta\mathbf{X}^f)=H(x^f+\Delta\mathbf{X}^f)-H(x^f)$ for the linear observation operator H.

Another advantage of the EnKF is its ability to directly calculate the analysis error covariance. We use the revised, unbiased Ensemble Transform Kalman Filter (ETKF) algorithm (Wang et al., 2004; Livings et al., 2008) to determine the analysis ensemble $\Delta\mathbf{X}^a$ and the a posteriori error covariance, $\mathbf{P}^a$:

$$\Delta\mathbf{X}^a = \Delta\mathbf{X}^f\mathbf{T}, \tag{8}$$

and

$$\mathbf{P}^a = \Delta\mathbf{X}^f\mathbf{T}(\Delta\mathbf{X}^f\mathbf{T})^T. \tag{9}$$

The transform matrix $\mathbf{T}$ given by

$$\mathbf{T}(\mathbf{T})^T = \mathbf{I} - (\Delta\mathbf{Y})^T[\Delta\mathbf{Y}(\Delta\mathbf{Y})^T + \mathbf{R}]^{-1}\Delta\mathbf{Y}. \tag{10}$$

We simplify the calculation of $\mathbf{T}$, $\mathbf{K}_e$, and $[\Delta\mathbf{Y}(\Delta\mathbf{Y})^T+\mathbf{R}]$, which is quite large due to the dense OCO observations, by using singular value decomposition (SVD) of the scaled model observation ensemble $\Delta\mathbf{Y}^T\mathbf{R}^{-1/2}$ (Livings, 2005).

3.2 A priori error and its representation

We construct an ensemble of perturbation states to reflect the a priori error covariance matrix, using eigenvalue decomposition:

$$\mathbf{P}^f = \mathbf{V}_x \mathbf{p}^{1/2} \left(\mathbf{V}_x \mathbf{p}^{1/2} \right)^T, \tag{11}$$

where $\mathbf{V}_x$ and $\mathbf{p}$ are the eigenvector matrix and the eigenvalue diagonal matrix of the error covariance, respectively.

At the limit of using the full rank matrix, as we do here, the ensemble of perturbation states is defined as:

$$\Delta \mathbf{X}^f = \mathbf{V}_x \mathbf{p}^{1/2}, \tag{12}$$

where the matrix $\Delta \mathbf{X}^f$ has a size of $N_x \times N_e$, with N_x=12×144, and the ensemble size N_e being equal to N_x.

The most time-consuming part in our flux inversions is the projection of these perturbation states to the observation space by using the observation operator that includes running a global transport model (Sect. 2). To avoid running the transport model for every different observation configuration used in our experiments, we define one diagonal matrix $\Delta \mathbf{X}_0^f$ of the same size as $\Delta \mathbf{X}^f$, with each column only specifying an emission occurring in one of the twelve 8-day periods over one of the 144 regions. We then calculate the variations in the observed X_{CO_2} caused by these emissions through the observation operator H

$$\Delta \mathbf{Y}_0 = H(\Delta \mathbf{X}_0^f). \tag{13}$$

Because H is linear, we can calculate $\Delta \mathbf{Y}$ for any given a priori ensemble $\Delta \mathbf{X}^f$ by:

$$\Delta \mathbf{Y} = H(\Delta \mathbf{X}^f) = \Delta \mathbf{Y}_0 \left([\Delta \mathbf{X}_0^f]^{-1} \Delta \mathbf{X}^f \right). \tag{14}$$

In principle, we only need to retain part of column vectors given by Eq. (12) and ignore those associated with small amplitudes to provide a good approximation of the error covariance. However, such a reduced representation usually requires re-running GEOS-Chem model for every different observation configuration (see Appendix B), while tending to underestimate the a posteriori uncertainties (Sect. 4.6), on which this study is focused.

4 Results

We evaluate our EnKF approach using an observing system simulation experiment (OSSE) framework, which is illustrated in Fig. 4. Observed X_{CO_2} distributions are described in Sect. 2, which we regard as the "truth". Model X_{CO_2} distributions are defined similarly but we assume the prior flux estimates to be 80% higher than the "true" values.

First, we present results from a control experiment for a 7-month period from 1st January to 31st July 2003, during which the OCO instrument is assumed to operate at the nominal 32-day duty cycle with alternating 16-day nadir and glint measurements. We then assess the sensitivity of the a posteriori flux estimates to 1) systematic (bias) and random (unbiased) errors; 2) observation error, density and correlations; 3) alternative duty cycles; 4) the spatial resolution of the state vectors; and 5) the length of the lag window and the size of the ensemble.

We evaluate the performance of the EnKF by using an error reduction γ

$$\gamma = 1 - \sigma^a/\sigma^f, \tag{15}$$

where σ^f and σ^a denote the a priori and a posteriori variance uncertainties, respectively. For each 8-day mean regional flux, we calculate its σ^f from the a priori error covariance at the time when it first enters the lag window, and calculate σ^a from the a posteriori error covariance at the time when it leaves the lag window.

In principle, the error reduction γ reflects only the quality of the observations, and not any assumption we have made about the "true" or prior fluxes. However, approximations in EnKF approaches may lead to underestimation of the a posteriori uncertainties (Livings et al., 2008), in particular, when a reduced representation of the error covariance is used (see Sect. 4.6).

4.1 Control experiment

Figure 5 presents the error reduction in the estimates for 8-day mean fluxes over 144 regions. The results have been averaged over a 32-day period from 17th January to 17th February 2003.

During the northern winter, nadir measurements cover the latitudes between 90° S and 60° N, while glint measurements only reach 55° N. As a result, over most land regions between 30° S and 50° N, OCO measurements reduce uncertainties in the flux estimates by more than 70% (Fig. 5a), while errors over the boreal latitudes decrease by 20–65%. The widespread error reduction reflects the coverage and the precision of nadir and glint measurements.

We find that most of the error reduction occurs when the continental signal is younger than 3 weeks and still distinct from the slowly varying background. Over regions with a dense distribution of observations, the error reduction can reach saturation well within three (8-day) assimilation cycles. Conversely, at the northern high latitudes during winter, when there is a low observation density, saturation of error reduction requires more time but is still within 5–6 assimilation cycles (<2 months). These results suggest that our 3-month lag window is more than sufficient (Sect. 4.6).

We find that OCO measurements reduce the uncertainties in oceanic CO_2 flux estimates by 10–60%, despite small a priori errors. Most of these reductions are attributed to the accurate glint measurements over the oceans (Fig. 1). To highlight this point, Fig. 5b shows the error reduction when the glint measurements over the oceans are excluded from the assimilation. Without these measurements, the error reduction over ocean reaches 10–30%. We also find that omitting glint measurements over ocean also leads to lower error reduction over tropical continents. For example, the error reduction for region 27 over tropical South America is 50%, compared to 70% with all

ACPD
8, 19917–19955, 2008

clear glint measurements (Fig. 5a). The results reflect the additional constraints on continental surface CO_2 flux from accurate glint measurements of continental outflow over the surrounding oceans.

The results for the time period from 1st May to 1st June 2003 (Fig. 5c) show similar significant error reduction over lands. In particular, the northern high latitudes are now fully covered by OCO measurements, and the corresponding error reduction reaches 40–90%.

4.2 Sensitivity to bias and unbiased error

We generate a random observation error for each OCO measurement through randomly sampling a Gaussian probability distribution function with the variance equal to the measurement uncertainty. Figure 6 shows that including random errors to the "true" OCO observations (see Fig. 4) leads to departures from the control flux estimates. These departures are well within the a posteriori errors, and usually become even smaller when averaged over a longer period (not shown).

The impacts from small-scale or scene-dependent measurement biases on the source/sink estimation are of greater interest. Previous work has implemented coherent bias by relating it with to the sub-micron aerosols (Chevallier et al., 2007a). We use a similar approach to include an observation bias (in ppm) twice that of the obscured AODs. This results in a maximum bias for clear observations of about 0.6 ppm, a magnitude similar to the typical observation error. Using this approach, we find significant positive biases over Eurasia, and over Europe, accompanied by negative and positive biases over Pacific, similar to Chevallier et al. (2007a). However, these systematic differences in the estimated fluxes are still smaller than the a posteriori errors.

4.3 Sensitivity to measurement duty cycle

The OCO satellite repeats its sun-synchronous orbit every 16 days and the current instrument configuration is to switch between nadir and glint mode at the same frequency (Crisp et al., 2004), which can be reprogrammed within orbit, if necessary. Consequently, there are a number of nadir-glint measurement combinations that could form a 32-day duty cycle over the nominal two-year OCO mission. Here we assess the impact of two alternative duty cycles on estimating surface CO_2 fluxes: the nadir-only cycle and the glint-only cycle.

Figure 7a compares the results for the nadir-only and glint-only duty cycles with the control experiment. We have averaged the results over a 32-day cycle from 17th January to 17th February 2003. For all the three duty cycles, the geographical pattern of the resulting error reduction is similar, showing significant reductions (40–85%) over lands, and moderate reductions (10–65%) over oceans. Because of the wider observation coverage, the nadir-only cycle has better performance over northern high latitudes than the other duty cycles. However, glint-only measurements lead to slightly larger error reductions over the terrestrial tropics, although nadir measurements theoretically represent better constraints for terrestrial sources and sinks by sampling overhead. As mentioned previously, we generally find that tropical land masses are typically characterized by extensive and persistent cloud cover during the wet season and by smoke aerosol during the dry season so the observation density of nadir measurements is low. High-precision glint measurements, sampling continental outflow over the oceans, provide important constraints for estimating land flux estimates.

Nadir measurements provide little constraint on ocean CO_2 flux estimates, as expected. Glint measurements lead to significant reductions of flux errors over the oceans, reaching 40–60% over the tropics. The 16-day nadir/glint switch leads to a moderate performance between the glint-only and nadir-only duty cycles.

4.4 Sensitivity to observation density and correlation

Figure 7b shows that because of the high observation density, reducing the clear observation number by 20% only slightly degrades (i.e., increases) the uncertainties of the estimated fluxes.

To investigate the impact of measurement correlations, we assume a distancedependent spatial correlation between observations from the same satellite orbits so that the off-diagonal term $R(m_1, m_2)$ for two measurements m_1 and m_2 in one orbit is given as

$$R(m_1, m_2) = \sqrt{R(m_1, m_1)\, R(m_2, m_2)}\, \exp\left(-l(m_1, m_2)/l_{cor}\right), \tag{16}$$

where l_{cor}=300 km is the characteristic spatial correlation length scale, and $l(m_1, m_2)$ is the distance between the two measurements m_1 and m_2. Here we assume that these correlations between successive X_{CO_2} arise from both the model and observation errors. Figure 7b shows that imposing a spatial correlation weakens the measurement constraint on flux estimations, as expected. We find the largest impacts from including observation correlations are over the oceans where there is a greater density of cloud and aerosol-free measurements, in agreement with Chevallier (2007b). Successive clear measurements over most land regions are sparse and consequently strong correlations are rare. The shown degradation reflects a weaker but possibly more realistic measurement constraint, but does not suggest that it is a beneficial practice to ignore the existing observation correlations in data assimilation (Stewart et al., 2008).

4.5 Sensitivity of state vector resolution

To investigate the sensitivity of our results to the spatial resolution of the state vector, we estimate 8-day surface fluxes over South American tropical region during January 17th to 17th February 2003 at 4 different spatial resolutions: 1) the standard T3 region, 2) the 1/4 T3 region (about 2 300 000 km^2), and 3) the 1/9 T3 region (about 1 100 000 km^2), and 4) at 4°×5° resolution grid boxes (about 220 000 km^2). To reduce the computational costs, we represent the rest of world using the other standard T3 regions (plus one low emission region, see Sect. 3).

Figure 8 shows that the mean error reduction (i.e., the averaged error reduction over tropical South America) decreases rapidly as the the spatial resolution increases. All these experiments lead to almost the same spatial pattern in the aggregated errors over T3 regions as our control run with 144 regions (not shown here).

ACPD

8, 19917–19955, 2008

4.6 Sensitivity to lag window and ensemble size

Figure 9a compares the estimated fluxes in the control experiment with the results for the inversions with a shorter lag window (8×8 days vs. 12×8), or with a smaller ensemble size (6×144 vs. 12×144, see Appendix B). The differences in the resulting fluxes are typically smaller than 0.05 GtC yr^{-1}.

Figure 9b shows that when the lag window shrinks to 8×8 days, the error reductions reach almost the same values as the control run, indicating our results are not sensitive to the length of the lag time window. When the size of the a priori ensemble is halved from 12×144 to 6×144, the error reduction is generally overestimated, due to the loss of the a posteriori variations by gradually removing perturbation states associated with small flux uncertainties.

5 Conclusions

We developed an ensemble Kalman Filter (EnKF) to estimate 8-day regional surface fluxes of CO_2 from space-borne CO_2 dry-air mole fraction observations (X_{CO_2}) and evaluated the approach using a series of synthetic experiments, in preparation for data from the NASA Orbiting Carbon Observatory (OCO). The 32-day duty cycle of OCO alternates between nadir and glint (specular reflection) measurements of backscattered solar radiation at short-wave infrared wavelengths. Our EnKF represents a complementary approach to the variational techniques that have already been developed for interpreting the space-borne X_{CO_2} data (e.g., Chevallier et al., 2007a). The main advantages of the EnKF is that it does not require the linearized versions of the forecast and observation operators and provides a framework to estimate the uncertainty of a posteriori fluxes. We use the ensemble transform Kalman Filter algorithm to determine the ensemble analysis and its error covariance.

For this work, we estimate 8-day CO_2 surface fluxes over 144 geographical regions (corresponding to 1 100 000 km^2 over land), based on the TransCom-3 experiments

ACPD
8, 19917–19955, 2008

(Gurney et al., 2002). We use a 12×8-day lag window, taking into account that X_{CO_2} measurements include surface flux information from prior time windows. The observation operator relates surface CO_2 fluxes to the global distributions of the "observed" X_{CO_2}. First, we use the GEOS-Chem transport model to relate surface fluxes to global 3-D distributions of CO_2 concentrations. Second, these distributions are sampled at the time and location of OCO measurements, using geolocation data from the Aqua satellite that precedes OCO by 15 min. Third, we use seasonal probability density functions, determined from MODIS and MISR, to remove cloudy scenes and scenes with aerosol optical depth (AOD)>0.3. Finally, we use scene-dependent averaging kernels, determined using detailed radiative transfer modelling specific to OCO (Bösch et al., 2008), to relate the CO_2 profiles to X_{CO_2} as a function of land-type, AOD, solar zenith angle, and nadir and glint view modes. We use the scene-dependent measurement errors that correspond to the averaging kernels. These scene-dependent calculations provide us with the most realistic simulation of X_{CO_2} distributions to date, with which to understand potential of OCO to estimate surface CO_2 fluxes. We use the same observation operator to model atmospheric distributions of X_{CO_2}, but with an 80% bias in the prior surface emissions.

We show that OCO X_{CO_2} measurements significantly reduce the uncertainty of surface CO_2 flux estimates. We find that nadir measurements are better at estimating land-based fluxes and glint measurements are generally better at constraining ocean fluxes. Nadir X_{CO_2} measurements over the terrestrial tropics are typically sparse throughout the year because of either widespread and persistent cloud cover during the wet season or smoke aerosol associated with extensive biomass burning during the dry season. We find that glint measurements over the oceans provide the most effective constraint for estimating terrestrial CO_2 fluxes by accurately sampling fresh continental outflow over neighbouring oceans.

We also presented the results from sensitivity experiments to investigate how flux estimates change with 1) bias and unbiased errors, 2) alternative duty cycles, 3) measurement density and correlations, 4) the spatial resolution of estimated flux estimates,

ACPD
8, 19917–19955, 2008

and 5) reducing the length of the lag window and the size of the ensemble. We find that biases in the observations, which we introduce by scaling the error from AOD (by a factor of two), cause large perturbations to some of the a posteriori fluxes but they are still within the a posteriori uncertainty of the control experiment. We find that either the current 32-day duty cycle (alternating 16-day cycle between glint and nadir measurements) or one that uses only glint view measurements will address the primary science objectives of the OCO mission, a reflection of the importance of glint measurements in constraining tropical terrestrial fluxes. A modest 20% reduction in the number of available clear observations does not affect a posteriori flux estimates, reflecting the high measurement density. Introducing a spatial correlation between successive measurements effectively reduces the number of independent observations. We find that spatial correlations mainly affect glint measurements over the oceans where there is a greater number of neighbouring scenes that are cloud-free and have AODs <0.3. We find that reducing the size of the geographical regions over which to estimate surface fluxes much below 1 million km^2 introduces large correlations between neighbouring regional estimates. In the control experiment, we simultaneously estimate surface fluxes at the time of the assimilation and at times up to 3 months prior. We find that surface flux estimates for a particular 8-day period typically converge after ingesting 4–6 weeks of data. To improve the speed of the EnKF we halved the number of ensemble states used to determine the a priori error covariance and showed that the flux estimates were close to the control experiment but using a reduced number of ensemble states, we generally underestimated the associated error.

The analysis we presented here can easily be applied to CO_2 and CH_4 data from the Japanese Greenhouse gas Observing SATellite (GOSAT). Implementing the EnKF for GOSAT will require detailed information about the orbit, averaging kernels, and associated error characteristics. Using OCO and GOSAT data together would improve the spatial and temporal coverage of the atmosphere and would likely 1) reduce the assimilation lag window and 2) increase the spatial resolution of resulting independent flux distributions.

Appendix A

Description of the GEOS-Chem Model of Atmospheric CO_2

We use the GEOS-Chem global 3-D chemistry transport model (v7-03-06) to calculate column concentrations of CO_2 from prescribed surface CO_2 fluxes described below. We used the model with a horizontal resolution of $2^\circ \times 2.5^\circ$, and 30 vertical levels (derived from the native 48 levels) ranging from the surface to the mesosphere, 20 of which are below 12 km. The model is driven by GEOS-4 assimilated meteorology data from the Global Modeling and Assimilation Office Global Circulation Model based at NASA Goddard. The 3-D meteorological data is updated every six hours, and the mixing depths and surface fields are updated every three hours. The CO_2 simulation is based on Suntharalingam et al. (2005) and Palmer et al. (2006, 2008).

We use gridded fossil fuel emission distributions, representative of 1995 (Suntharalingam et al., 2005), which we have scaled to 2003 values using regional budget estimates for the top 20 emitting countries in 2003 from the Carbon Dioxide Information Analysis Center (Marland et al., 2007). Biofuel emission estimates are taken from Yevich and Logan (2003) and represent climatological values. Monthly mean biomass burning emission estimates are taken from the second version of the Global Fire Emission Database (GFEDv2) for 2003 (van der Werf et al., 2006), which are derived from ground-based and satellite observations. Daily mean land biosphere fluxes are taken from the CASA model for 2001 (Randerson et al., 1997), in the absence of corresponding fluxes for 2003. We do not explicitly account for the contribution of fuel combustion CO_2 from the oxidation of reduced carbon species (Suntharalingam et al., 2005) as they make only a small contribution to the CO_2 column. Monthly mean air-sea fluxes of CO_2 are taken from Takahashi et al. (1999).

CO_2 concentrations for January 2002 were initialized from a previously evaluated model run (Palmer et al., 2006), which we integrate forward to January 2003. We include an additional initialization to correction for the model bias introduced by not accounting for the net uptake of CO_2 from the terrestrial biosphere. We make this downward correct by comparing the difference between GLOBALVIEW CO_2 data (GLOBALVIEW-CO2) and model concentrations over the Pacific during January 2003. Differences range from 1 to 4 ppmv with a median of 3.5 ppmv, and we subtract this value globally, following Suntharalingam et al. (2005). From January 2003 the total CO_2 tracer becomes the "background" CO_2 concentration and is only subject to atmospheric transport. At that time, we also introduce additional model tracers, initialized with a uniform value (for numerical reasons and which is subtracted in subsequent analyses), that account for the monthly production and loss of CO_2 originating from specific geographical regions and surface processes ("tagged" tracers). The linear sum of these monthly tagged tracers (and the "background") is equivalent to the total CO_2.

Appendix B

Description of the reduced representation of a priori error covariance

The most computationally expensive part in our approach is the projection of the a priori ensemble to the observation space. Such cost can be reduced by using a smaller ensemble to approximate the a priori error covariance. Because in this study, we have adopted a lag window of 12×8 days, the algorithm to construct the perturbation states (of a reduced number) at the beginning of each assimilation cycle from previous analysis and the current forecasts is eventually required to be equally applicable in constructing the corresponding variations in the 3-D CO_2 distributions at this time being to avoid rerunning the GEOS-Chem model from 12×8 days ago for these newly constructed perturbation states.

To construct the prior ensemble of size N_e for one assimilation cycle j to assimilate observations during day d to $d+8$, we start with the analysis ensemble at the end of the previous assimilation cycle $j-1$, which is given by Eq. (8):

$$\Delta \mathbf{X}^a_{j-1} = \Delta \mathbf{X}^f_{j-1} \mathbf{T}_{j-1}. \tag{B1}$$

For clarification, we have explicitly given the subscript $j-1$ to denote the assimilation cycle. The resulting matrix $\Delta \mathbf{X}^a_{j-1}$ has the same size of the matrix $\Delta \mathbf{X}^f_{j-1}$, and consists of N_e perturbation states, each of which has $N_x = 12 \times 144$ elements for perturbations in fluxes during the past 12×8 days.

To maintain the ensemble size N_e at the following assimilation cycle j, we rewrite $\Delta \mathbf{X}^a_{j-1}$ by using the singular value decomposition

$$\Delta \mathbf{X}^a_{j-1} = \mathbf{U}^a_{j-1} \Sigma^a_{j-1} (\mathbf{V}^a_{j-1})^T, \tag{B2}$$

where $\mathbf{U}^a_{j-1}$, and $\mathbf{V}^a_{j-1}$ are two orthogonal matrices of size $N_x \times N_x$ and $N_e \times N_e$, respectively, and Σ^a_{j-1} is a diagonal matrix of size $N_x \times N_e$, with non-zero diagonal elements representing singular values of matrix $\Delta \mathbf{X}^a_{j-1}$ in the descending order. We can now obtain a base matrix $\Delta \mathbf{X}_b$ by applying $\mathbf{V}^a_{j-1}$ to Eq. (B2),

$$\Delta \mathbf{X}_b = \Delta \mathbf{X}^a_{j-1} \mathbf{V}^a_{j-1} = \mathbf{U}^a_{j-1} \Sigma^a_{j-1}. \tag{B3}$$

Note that the matrix $\Delta \mathbf{X}_b$ has the same size as $\Delta \mathbf{X}^a_{j-1}$, and satisfies

$$\mathbf{P}^a_{j-1} = \Delta \mathbf{X}_{j-1} (\Delta \mathbf{X}_{j-1})^T = \Delta \mathbf{X}_b (\Delta \mathbf{X}_b)^T \tag{B4}$$

The first $N_e - 144$ columns of the matrix $\Delta \mathbf{X}_b$ have the largest amplitudes, and form the basis to represent the error covariances for flux estimates before day d. By concatenating them with other 144 new perturbation vectors for the uncertainties associated with flux forecasts for day d to $d+8$, we can now generate the required the a priori ensemble of size N_e for assimilation cycle j.

The GEOS-Chem model simulations for the source/sink perturbations represented by $\Delta \mathbf{X}^f_{j-1}$ also provide an ensemble of variations in the 3-D CO_2 distributions at day

d, i.e., the beginning of the new assimilation cycle j. Due to the linearity of the atmospheric transport processes, the transformation by matrix $\mathbf{T}_{j-1}\mathbf{V}^a_{j-1}$ can be applied to these atmospheric CO_2 variations, and the resulting first N_e-144 fields are attributed to the source/sink perturbations represented by the first N_e-144 columns in $\Delta\mathbf{X}_b$.

During the new assimilation cycle j, these N_e-144 perturbed atmospheric CO_2 distributions are first added to the atmospheric CO_2 "analysis" $f^{CO_2}_{d-1}$, which corresponds to the sources/sinks represented by the a posteriori fluxes x^a_{j-1}, and then are propagated from from day d till day $d+8$ by using the GEOS-Chem model forced with the new (day d to day $d+8$) surface flux forecasts. We also make another set of model simulations from a common initial atmospheric CO_2 distribution $f^{CO_2}_{d-1}$ from day d till day $d+8$. These simulations are forced by the new surface flux forecast or by one of the associated 144 perturbation states. As a result, at each assimilation cycle, we run the GEOS-Chem model simulations for only N_e+1 tagged "tracers", instead of the N_x+1 ones, to obtain required perturbation model observations $\Delta\mathbf{Y}$, as well as the initial conditions for the following assimilation cycles.

Acknowledgements. This study is funded by the UK Natural Environment Research Council under NE/F000014/1. Work at the Jet Propulsion Laboratory (JPL), California Institute of Technology was carried out under a contract with the National Aeronautics and Space Administration.

References

Baker, D. F., Doney, S. C., and Schimel, D. S.: Variational data assimilation for atmospheric CO_2, Tellus Ser. B, 58, 359–365, 2006. 19920

Barkley, M. P., Monks, P. S., Frieß, U., Mittermeier, R. L., Fast, H., Körner, S., and Heimann, M.: Comparisons between SCIAMACHY atmospheric CO_2 retrieved using (FSI) WFM-DOAS to ground based FTIR data and the TM3 chemistry transport model, Atmos. Chem. Phys., 6, 4483–4498, 2006, http://www.atmos-chem-phys.net/6/4483/2006/. 19919

Barkley, M. P., Monks, P. S., Hewitt, A. J., Machida, T., Desai, A., Vinnichenko, N., Nakazawa, T., Yu Arshinov, M., Fedoseev, N., and Watai, T.: Assessing the near surface sensitivity of SCIAMACHY atmospheric CO_2 retrieved using (FSI) WFM-DOAS, Atmos. Chem. Phys., 7, 3597–3619, 2007, http://www.atmos-chem-phys.net/7/3597/2007/. 19919

Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects, Mon. Weather Rev., 129, 420–436, 2001. 19924

Bösch, H., Baker, D., Connor, B., O'Brien, D., Crisp, D., and Miller, C.: Global Characterization of X_{CO_2} Retrievals from OCO Observations, in preparation, 2008. 19920, 19922, 19934

Bousquet, P., Peylin, P., Ciais, P., Quere, C. L., Friedlingstein, P., and Tans, P. P.: Regional changes in carbon dioxide fluxes of land and oceans since 1980, Science, 290, 1342–1346, 2000. 19919

Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V., Chance, K. V., and Goede, A.: SCIAMACHY: Mission objectives and measurement modes, J. Atmos. Sci., 56, 127–150, 1999. 19919

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P.: An improved Kalman Smoother for atmospheric inversions, Atmos. Chem. Phys., 5, 2691–2702, 2005, http://www.atmos-chem-phys.net/5/2691/2005/. 19920

Chevallier, F.: Impact of correlated observation errors on inverted CO_2 surface fluxes from OCO measurements, Geophys. Res. Lett., 34, L24804, doi:10.1029/2007GL030463, 2007b. 19920, 19932

Chevallier, F., Bréon, F.-M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of CO_2 sources and sinks: Theoretical study in a variational data assimilation framework, J. Geophys. Res., 112, D09307, doi:10.1029/2006JD007375, 2007a. 19920, 19930, 19933

Chevallier, F. M. F., Peylin, P., Bousquet, S. S. P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res., 110, D24309, doi:10.1029/2005JD006390, 2005. 19919

Connor, B. J., Boesch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: Inverse method and prospective error analysis, J. Geophys. Res., 113, D05305, doi:10.1029/2006JD008336, 2008. 19922

Crisp, D., Atlas, R. M., Breon, F.-M., et al.: The Orbiting Carbon Observatory (OCO) Mission, Adv. Space. Res., 34(4), 700–709, 2004. 19919, 19920, 19921, 19923, 19931

Ehrendorfer, M.: A review of issues in ensemble-based Kalman filtering, Meteorol. Z., 16, 795–

818, 2007. 19920
Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99(C5), 10 143–10 162, 1994. 19920, 19926
Evensen, G.: The Ensemble Kalman Filter: Theoretical formulation and practical implementation, Ocean Dynam., 53, 343–367, 2003. 19920
GLOBALVIEW-CO_2: Cooperative Atmospheric Data Project Carbon Dioxide, CD-ROM, NOAA GMD, Boulder, Colorado, USA (also available via anonymous FTP to ftp.cmdl.noaa.gov, path:/ccg/co2/GLOBALVIEW, 2006. 19937
Gurney, K. R., Law, R. L., Denning, A. S., et al.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, 2002. 19925, 19934
Houtekamer, P. L. and Mitchell, H. L.: Data assimilation using an ensemble Kalman filter technique, Mon. Weather Rev., 126, 796–811, 1998. 19920
Houweling, S., Kaminski, T., Dentener, F., Lelieveld, J., and Heimann, M.: Inverse modeling of methane sources and sinks using the adjoint of a global transport model, J. Geophys. Res., 104, 26 137–26 160, 1999. 19919
Houweling, S., Hartmann, W., Aben, I., Schrijver, H., Skidmore, J., Roelofs, G.-J., and Breon, F.-M.: Evidence of systematic errors in SCIAMACHY-observed CO_2 due to aerosols, Atmos. Chem. Phys., 5, 3003–3013, 2005, http://www.atmos-chem-phys.net/5/3003/2005/. 19919
Law, R. M., Chen, Y. H., and Gurney, K. R.: TransCom-3 CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices, Tellus, Ser. B, 55, 580–595, 2003. 19925
Livings, D. M.: Aspects of the Ensemble Kalman Filter, mSc thesis, Department of Mathematics, University of Reading, UK, 34–37, 2005. 19926
Livings, D. M., Dance, S. L., and Nichols, N. K.: Unbiased ensemble square root filters, Physica D., 237/8, 1021–1028, 2008. 19926, 19929
Lorenc, A. C.: The potential of the ensemble Kalman filter for NWP–A comparison with 4D-Var, Q. J. Roy. Meteorol. Soc., 129, 3183–3203, 2003. 19920
Maksyutov, S., Kadygrov, N., Nakatsuka, Y., Patra, P. K., Nakazawa, T., Yokota, T., and Inoue, G.: Projected impact of the GOSAT observations on regional CO_2 flux estimations as a function of total retrieval error, Journal of Remote Sensing Society of Japan, in press, 2008. 19919
Marland, G., Boden, T. A., and Andres, R. J.: Global, Regional, And National CO_2 Emissions, in

Trends: A Compendium of Data on Global Change, Tech. Rep. 2007, 7346, Carbon Dioxide Information Analysis Center Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tenn., USA, 2007. 19936

Miller, C. E., Crisp, D., DeCola, P. L., et al.: Precision requirements for space-based X_{CO_2} data, J. Geophys. Res, D10314, doi:10.1029/2006JD007659, 2007. 19920

Palmer, P. I., Suntharalingam, P., Jones, D. B. A., Jacob, D. J., Streets, D. G., Fu, Q., Vay, S. A., and Sachse, G. W.: Using CO_2:CO correlations to improve inverse analyses of carbon fluxes, J. Geophys. Res., 111, D12318, doi:10.1029/2005JD006697, 2006. 19936

Palmer, P. I., Barkley, M. P., and Monks, P. S.: Interpreting the variability of space-borne CO_2 column-averaged volume mixing ratios over North America using a chemistry transport model, Atmos. Chem. Phys., 8, 5855–5868, 2008, http://www.atmos-chem-phys.net/8/5855/2008/. 19921, 19936

Patra, P. K., Maksyutov, S., Sasano, Y., Nakajima, H., Inoue, G., and Nakazawa, T.: An evaluation of CO_2 observations with Solar Occultation FTS for Inclined-Orbit Satellite sensor for surface source inversion, J. Geophys. Res., 108(D24), 4759, doi:10.1029/2003JD003661, 2003. 19920, 19925

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, J. Geophys. Res., 110, D24304, doi:10.1029/2005JD006157, 2005. 19920

Peylin, P., Baker, D., Sarmiento, J., Ciais, P., and Bousquet, P.: Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO_2 data, J. Geophys. Res., 107(D19), 4385, doi:10.1029/2001JD000857, 2002. 19924

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A.: Terrestrial ecosystem production: A process model based on global satellite and surface data, Global Biogeochem. Cy., 7(4), 811–842, 1993. 19921

Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y., and Field, C. B.: The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, Global Biogeochem. Cy., 11(4), 535–560, 1997. 19936

Rayner, P. J., Law, R. M., O'Brien, D. M., Butler, T. M., and Dilley, A. C.: Global observations of the carbon budget: 3. Initial assessment of the impact of satellite orbit, scan geometry, and cloud on measuring CO_2 from space, J. Geophys. Res., 107(D21), 4557, doi:10.1029/2001JD000618, 2002. 19920

ACPD

8, 19917–19955, 2008

Estimating surface flux

L. Feng et al.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 1919–1964, 2003, http://www.atmos-chem-phys.net/3/1919/2003/. 19919, 19924

Schneising, O., Buchwitz, M., Burrows, J. P., Bovensmann, H., Reuter, M., Notholt, J., Macatangay, R., and Warneke, T.: Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite – Part 1: Carbon dioxide, Atmos. Chem. Phys., 8, 3827–3853, 2008, http://www.atmos-chem-phys.net/8/3827/2008/. 19919

Stewart, L. M., Dance, S., and Nichols, N.: Information content and correlated observation errors, Int. J. Numer. Meth. Fl., 56, 1521–1527, 2008. 19932

Suntharalingam, P., Randerson, J. T., Krakauer, N., Logan, J. A., and Jacob, D. J.: The influence of reduced carbon emissions and oxidation on the distribution of atmospheric CO_2: implications for inversion analysis, Global Biogeochem. Cy., 19, GB4003, doi:10.1029/2005GB002466, 2005. 19936, 19937

Takahashi, T., Wanninkhof, R. T., Feely, R. A., Weiss, R. F., Chapman, D. W., Bates, N. R., Olafsson, J., Sabine, C. L., and Sutherland, C. S.: Net sea-air CO_2 flux over the global oceans, proceedings of the 2nd international symposium CO_2 in the oceans: CGER 1037, National Institute for Environmental Studies, Tsukuba, Japan, 915 pp., 1999. 19936

Takahashi, T., Sutherland, S. C., Sweeney, C., et al.: 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, 2002. 19921

Tiwari, Y. K., Gloor, M., Engelen, R. J., Chevallier, F., Rödenbeck, C., Körner, S., Peylin, P., Braswell, B. H., and Heimann, M.: Comparing CO_2 retrieved from Atmospheric Infrared Sounder with model predictions: Implications for constraining surface fluxes and lower-to-upper troposphere transport, J. Geophys. Res., 111, D17106, doi:10.1029/2005JD006681, 2006. 19919

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr., A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, Atmos. Chem. Phys., 6, 3423–3441, 2006, http://www.atmos-chem-phys.net/6/3423/2006/. 19936

Wang, X., Bishop, C. H., and Julier, S. J.: Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble, Mon. Weather Rev., 132, 1590–1605, 2004. 19926

Yevich, R. and Logan, J. A.: An assessment of biofuel use and burning of agricultural waste in the developing world, Global Biogeochem. Cy., 17, 1095, doi:10.1029/2002GB001952, 2003. 19921, 19936

Zupansi, M.: Maximum likelihood Ensemble Filter: Theoretical Aspects, Mon. Weather Rev., 133, 1710–1726, 2005. 19926

Table 1. Uncertainty (in GtC yr^{-1}) associated with TransCom-3 (T3) continental and ocean regions that have been subdivided for our EnKF inversion

T3 Region	Err	EnKF Region	Err
North American Boreal	0.73	Reg (002–010)	1.64
North American Temperate	1.50	Reg (011–019)	3.38
South American Tropical	1.41	Reg (020–028)	3.18
South American Temperate	1.23	Reg (029–037)	2.76
North Africa	1.33	Reg (038–046)	3.00
South Africa	1.41	Reg (047–055)	3.18
Eurasia Boreal	1.51	Reg (056–064)	3.41
Eurasia Temperate	1.73	Reg (065–073)	3.89
Tropical Asia	0.87	Reg(074–082)	1.95
Australia	0.59	Reg (083–091)	1.34
Europe	1.42	Reg (092–100)	3.20
North Pacific Temperate	0.27	Reg (101–104)	0.61
West Pacific Tropics	0.39	Reg (105–108)	0.88
East Pacific Tropics	0.37	Reg (109–112)	0.83
South Pacific Temperate	0.63	Reg (113–116)	1.42
Northern Ocean	0.35	Reg (117–120)	0.79
Northen Atlantic Temperate	0.27	Reg (121–124)	0.61
Atlantic Tropics	0.41	Reg (125–128)	0.92
South Atlantic Temperate	0.55	Reg (129–132)	1.24
South Ocean	0.72	Reg (133–136)	1.62
Indian Tropical	0.48	Reg (137–140)	1.08
South Indian Temperate	0.41	Reg (141–144)	0.92

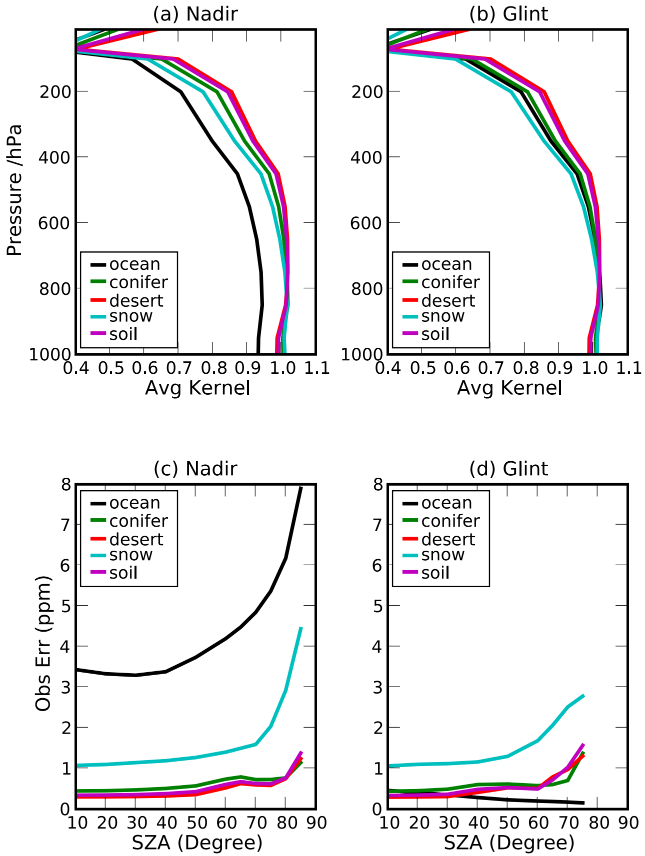


Fig. 1. Orbiting Carbon Observatory (OCO) instrument averaging kernels (dimensionless) associated with **(a)** nadir and **(b)** glint SWIR X_{CO_2} measurements as a function of pressure (hPa) for different land types, at a solar zenith angle (SZA) of 10° and an aerosol optical depth (AOD) of 0.1. Observation errors (ppm) associated with **(c)** nadir and **(d)** glint X_{CO_2} measurements as function of SZA for different land types and an AOD of 0.1.

ACPD
8, 19917–19955, 2008

Estimating surface flux

L. Feng et al.

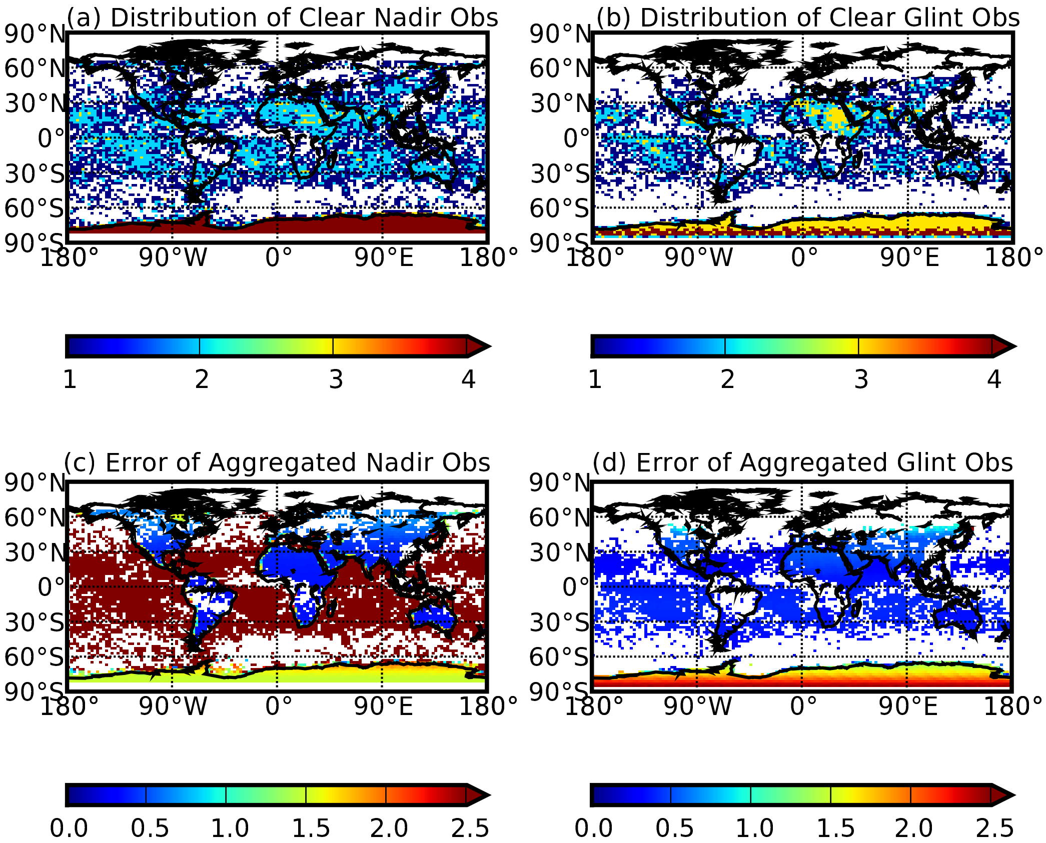


Fig. 2. Number of clear observations (aerosol optical depth <0.3 and cloud-free) and for **(a)** nadir and **(b)** glint X_{CO_2} measurements averaged over 16 days from 17th January to 1st February 2003, on a horizontal grid of $2^{\circ}\times2.5^{\circ}$. Associated aggregated errors (ppm) for the **(c)** nadir and **(d)** glint X_{CO_2} measurements

ACPD
8, 19917–19955, 2008

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
◀
▶
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

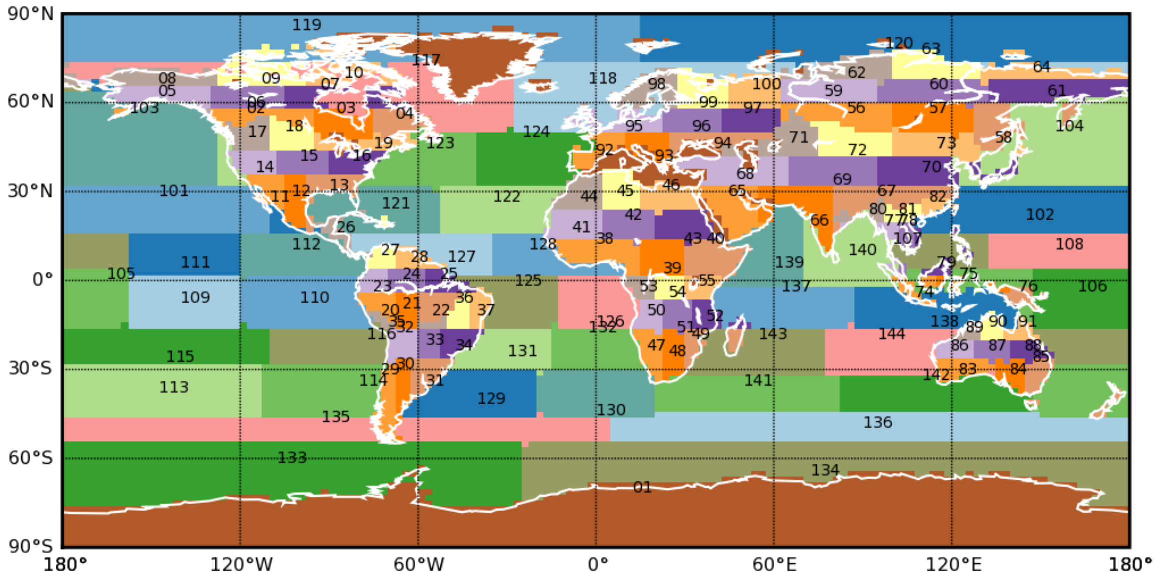
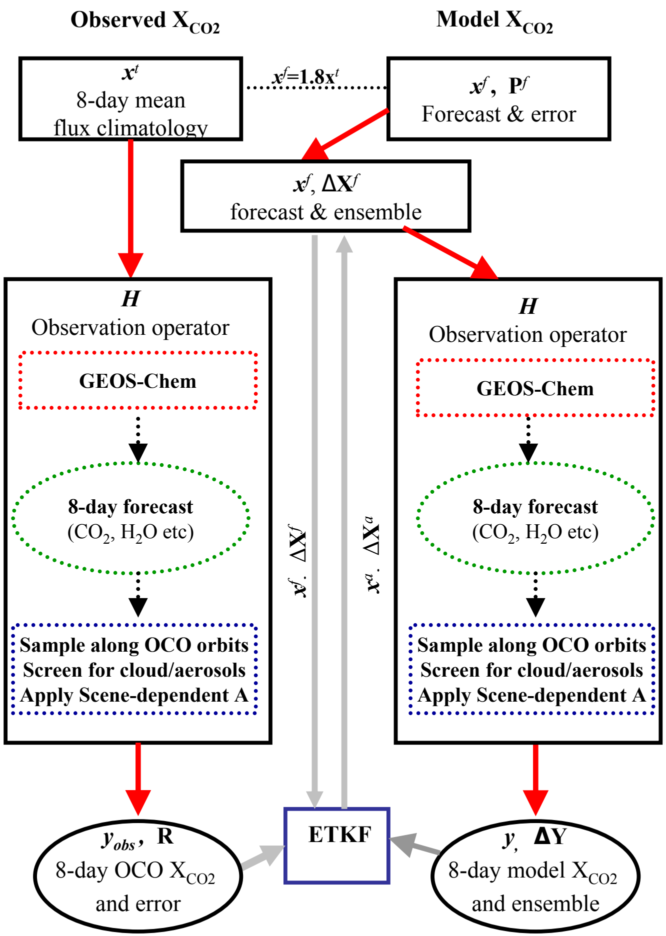


Fig. 3. The continental and ocean regions used to estimate CO_2 source and sinks, based on a coarser distribution from the TransCom-3 experiment (see Table 1 in the text and Gurney et al., 2002).



Observed X_{CO2}
Model X_{CO2}
x^t
8-day mean
flux climatology
x^f=1.8x^t
x^f, P^f
Forecast & error
x^f, ΔX^f
forecast & ensemble
H
Observation operator
GEOS-Chem
8-day forecast
(CO_2, H_2O etc)
Sample along OCO orbits
Screen for cloud/aerosols
Apply Scene-dependent A
x^f, ΔX^f
x^a, ΔX^a
H
Observation operator
GEOS-Chem
8-day forecast
(CO_2, H_2O etc)
Sample along OCO orbits
Screen for cloud/aerosols
Apply Scene-dependent A
y_{obs}, R
8-day OCO X_{CO2}
and error
ETKF
y, ΔY
8-day model X_{CO2}
and ensemble

Fig. 4.

Fig. 4. Schematic diagram of the OCO X_{CO_2} Observing System Simulation Experiment (OSSE). The left column describes the simulation of OCO X_{CO_2} measurements: $\mathbf{x}^t$ denotes the "true" fluxes; and H is the observation operator for mapping surface fluxes to X_{CO_2} observations $\mathbf{y}_{obs}$. H includes the GEOS-Chem transport model that relates surface fluxes to global 3-D CO_2 distributions, which are then sampled along OCO orbits. Scenes with cloud or aerosol optical depths >0.3 are removed. The resulting profiles are mapped to X_{CO_2} using scene-specific averaging kernels, with associated scene-specific error $\mathbf{R}$. The right column describe the simulation of model X_{CO_2} measurements using prior fluxes $\mathbf{x}^f$ (80% larger than $\mathbf{x}^t$) and the associated error covariance $\mathbf{P}^f$, which is approximated by the perturbation state vector ensemble $\Delta\mathbf{X}^f$. Mapping $\mathbf{x}^f$ and $\Delta\mathbf{X}^f$ to the observation space by observation operator H results in the model observation $\mathbf{y}$, and the associated variations $\Delta\mathbf{Y}$. The middle column shows that the Ensemble Transform Kalman Filter (ETKF) algorithm generates the optimal estimate $\mathbf{x}^a$, and the a posteriori error covariance $\mathbf{P}^a$ by comparing the model forecasts with observations.

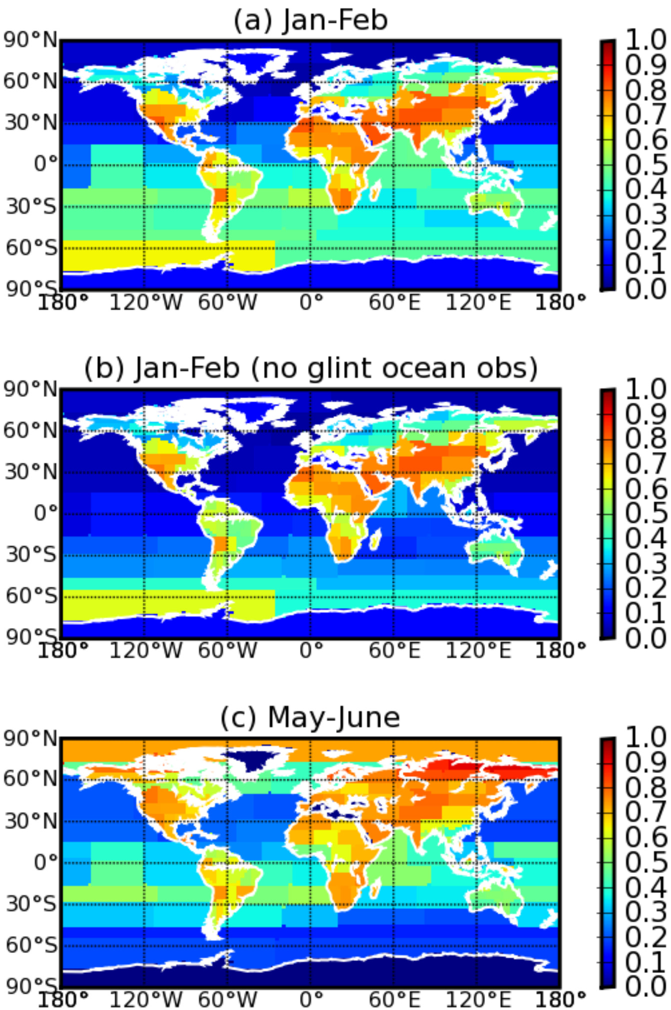


Fig. 5. (a) aposteriori error reduction, $\gamma = 1 - \sigma^a / \sigma^f$, associated with CO_2 flux estimation using OCO X_{CO_2} observations over one duty cycle, including alternate 16-day period of nadir and glint measurements. **(a)** Over 17th January – 17th February 2003; **(b)** over 17th January – 17th February 2003 without glint measurements over the ocean; and **(c)** over 1st May – 1st June 2003

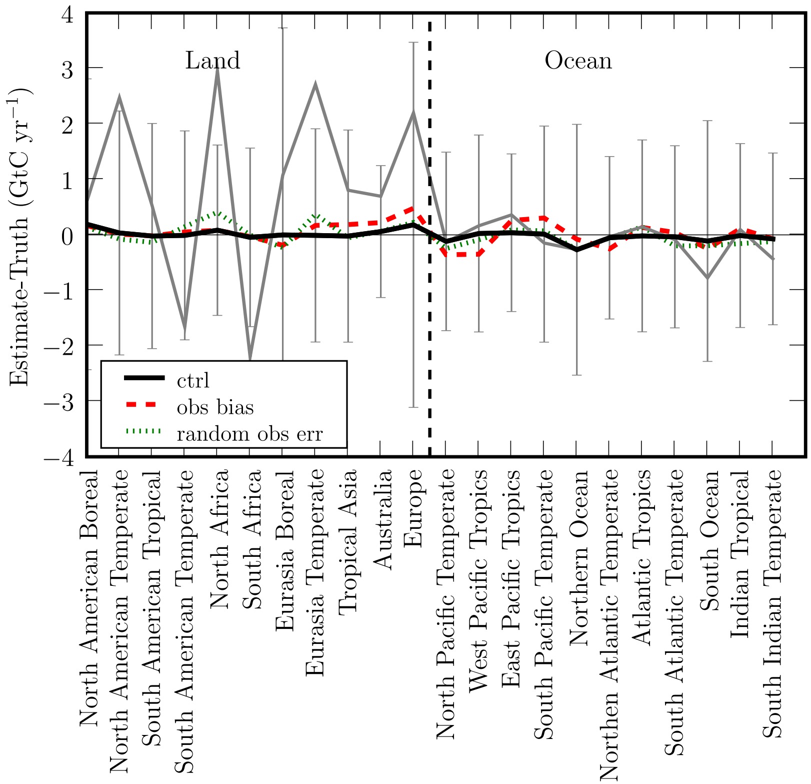


Fig. 6. CO_2 flux errors (GtC yr^{-1}) over one duty cycle from 17th January to 17th February 2003. The results have been aggregated from 144 regions (Fig. 3) to the 22 TransCom-3 regions (see Table 1 in the text and Gurney et al., 2002). Grey and black lines denote a priori and a posteriori errors. Vertical lines denote the standard deviation of the error. Dashed and dotted lines are the a posteriori flux errors associated estimates that include systematic and (larger) random errors. The vertical dashed line separates land and ocean fluxes.

ACPD
8, 19917–19955, 2008

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

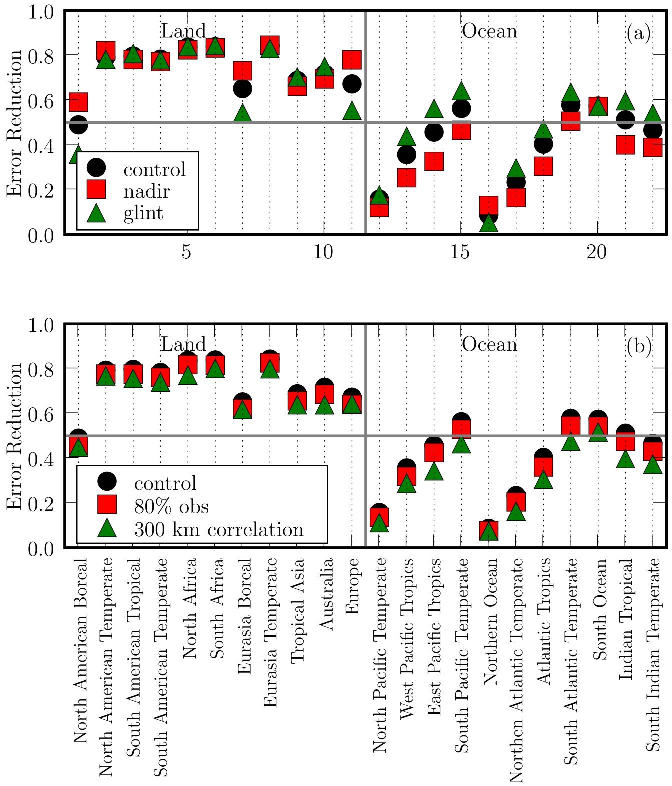


Fig. 7. Regional a posteriori error reduction, $\gamma=1-\sigma^a/\sigma^f$, associated with CO_2 flux estimation using OCO X_{CO_2} observations over one 32-day duty cycle (17th January – 17th February 2003). Results have been aggregated from 144 regions (Fig. 3) to the 22 TransCom-3 regions (see Table 1 in the text and Gurney et al., 2002). The vertical line separates land and ocean flux estimates. Significant error reduction is above the horizontal line at $\gamma=0.5$. **(a)** Circles denote results from the control run, squares denote results from using only nadir measurements, and triangles denote results from using only glint measurements; **(b)** squares denote results from using 80% of available measurements, and triangles denote results from including spatial correlations in the measurement error covariance **R** with an e-folding length scale of 300 km.

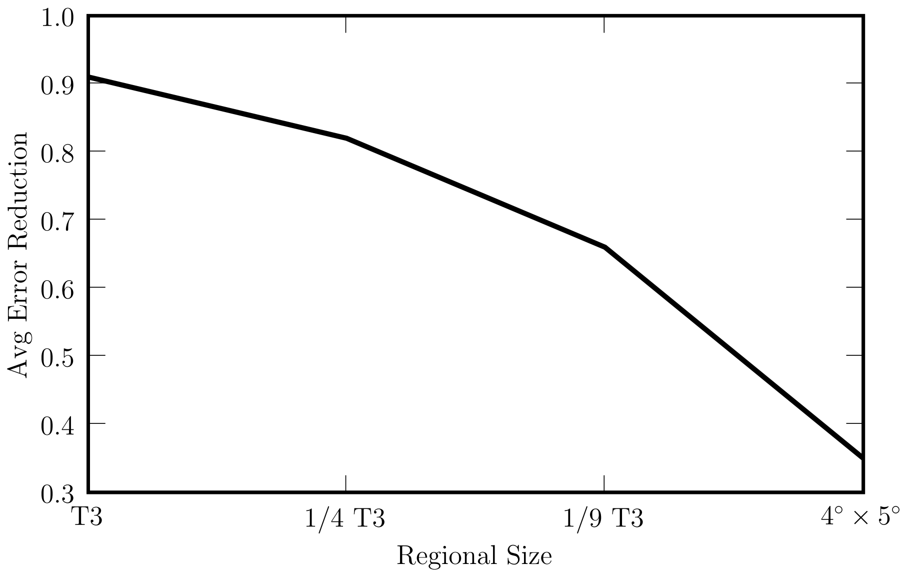


Fig. 8. The sensitivity of error reduction, $\gamma=1-\sigma^a/\sigma^f$, associated with CO_2 flux estimation using OCO X_{CO_2} observations over one 32-day duty cycle (17th January – 17th February 2003) to changes in the spatial resolution of the state vector. T3 denotes TransCom-3 regions that are approximately 9 500 000 km^2

ACPD
8, 19917–19955, 2008

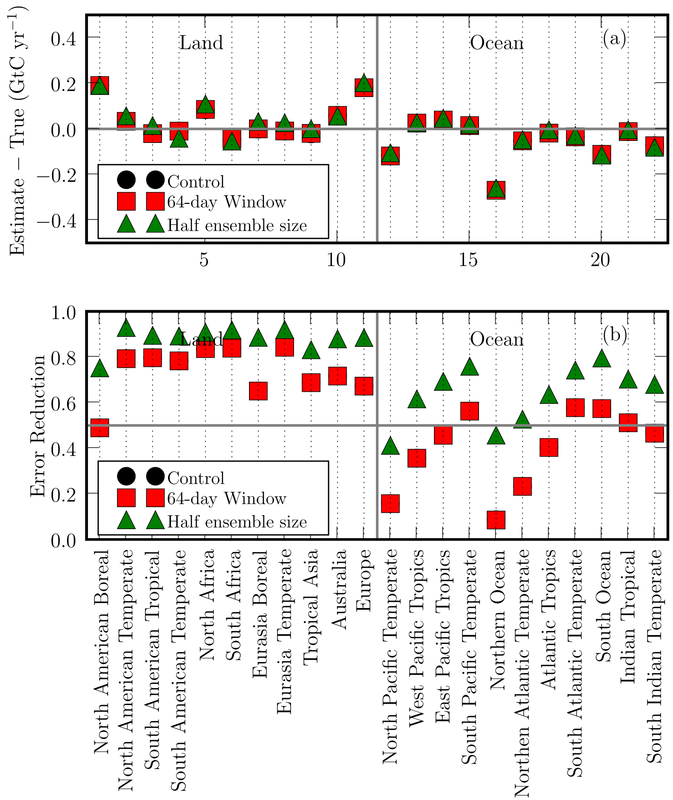


Fig. 9. **(a)** CO_2 flux errors (GtC yr^{-1}) over one duty cycle from 17th January to 17th February 2003. The results have been aggregated from 144 regions (Fig. 3) to the 22 TransCom-3 regions (see Table 1 in the text and Gurney et al., 2002). The circles denote the results for the control experiment, the squares the experiment with a shorter (8×8 days vs. 12×8 days) lag window, and the triangles the experiment using half the ensemble size (6×144 vs. 12×144); **(b)** the a posteriori error reduction, $\gamma = 1 - \sigma^a/\sigma^f$, associated with regional fluxes shown in (a).