**Atmospheric
Chemistry
and Physics**

Carbon source/sink information provided by column CO$_2$ measurements from the Orbiting Carbon Observatory

D. F. Baker[1,2], H. Bösch[3], S. C. Doney[1], D. O'Brien[2], and D. S. Schimel[4]

[1]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543-1543, USA
[2]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO 80523-1375, USA
[3]Department of Physics, University of Leicester, University Road, Leicester, LE1 7RH, UK
[4]National Ecological Observatory Network, Inc., 5340 Airport Boulevard, Boulder, CO 80301, USA

Abstract. We quantify how well column-integrated CO$_2$ measurements from the Orbiting Carbon Observatory (OCO) should be able to constrain surface CO$_2$ fluxes, given the presence of various error sources. We use variational data assimilation to optimize weekly fluxes at a $2° \times 5°$ resolution (lat/lon) using simulated data averaged across each model grid box overflight (typically every $\sim$33 s). Grid-scale simulations of this sort have been carried out before for OCO using simplified assumptions for the measurement error. Here, we more accurately describe the OCO measurements in two ways. First, we use new estimates of the single-sounding retrieval uncertainty and averaging kernel, both computed as a function of surface type, solar zenith angle, aerosol optical depth, and pointing mode (nadir vs. glint). Second, we collapse the information content of all valid retrievals from each grid box crossing into an equivalent multi-sounding measurement uncertainty, factoring in both time/space error correlations and data rejection due to clouds and thick aerosols. Finally, we examine the impact of three types of systematic errors: measurement biases due to aerosols, transport errors, and mistuning errors caused by assuming incorrect statistics.

When only random measurement errors are considered, both nadir- and glint-mode data give error reductions over the land of $\sim$45% for the weekly fluxes, and $\sim$65% for seasonal fluxes. Systematic errors reduce both the magnitude and spatial extent of these improvements by about a factor of two, however. Improvements nearly as large are achieved

over the ocean using glint-mode data, but are degraded even more by the systematic errors. Our ability to identify and remove systematic errors in both the column retrievals and atmospheric assimilations will thus be critical for maximizing the usefulness of the OCO data.

1 Introduction

The global carbon cycle plays a key role in the climatic response to anthropogenic forcing, yet our understanding of its dominant processes is still too weak to make accurate long-term predictions (IPCC, 2007). Atmospheric CO$_2$ measurements have revealed much of what we know about the functioning of the global carbon cycle. As our data coverage has increased, inverse methods have been used to optimize global sources and sinks of CO$_2$ and the process models that compute them (Enting et al., 1995; Bousquet et al., 2000; Rödenbeck et al., 2003; Baker et al., 2006a; Rayner et al., 2005).

So far, the "top-down" atmospheric inverse approach to validating carbon models has been only marginally successful: where the data are most dense, fluxes may be estimated at continental scales (Baker et al., 2006a), but not at the regional scales where they would be most useful for identifying flaws in the carbon models. Part of the problem is that the transport models have systematic mixing errors, notably in the vertical. The models also have great difficulty representing point measurements, particularly over the continents, using grid boxes 100s of km wide. The largest problem, however, is that the spatio-temporal density of the current in

situ measurement network is insufficient to correct the surface fluxes at regional scales. For the continental United States, for example, solving for fluxes at a 500 km resolution would require at least $7\,500\,000\,km^2/(500\,km)^2 \approx 30$ sites, each sampling air high enough in the column to have a footprint at least 500 km wide, with a frequency dictated by the cross-continental advection time scale.

Space-based measurements provide the most realistic opportunity to achieve global coverage at such regional scales. Recently, two satellites have been designed specifically to measure the column-averaged dry air mole fraction of CO_2 (X_{CO_2}): Japan's Greenhouse Gases Observing Satellite (GOSAT) and NASA's Orbiting Carbon Observatory (OCO). Their instruments measure CO_2 absorption in the near infra-red (IR) portion of the reflected solar beam and thus have sensitivity down to the surface, including the variable near-surface CO_2 concentrations most affected by the fluxes (Olsen & Randerson, 2004); previous instruments measuring in thermal IR bands sensed CO_2 concentrations mostly in the mid- to upper-troposphere, with little information about the surface fluxes (Chevallier et al., 2005a, b). Both missions also try to identify cloud-free scenes for their retrievals, since radiative transfer modeling problems associated with clouds can cause large errors in the retrieved CO_2 concentrations. Both use sun-synchronous orbits with early afternoon sun-lit equator crossing times and orbital inclinations near 98° (though, since their ascending nodes are 180° off, their paths cross only at the equator); subsequent orbits are separated by $\sim$25° in longitude, $\sim$99 min apart. In addition to nominal near-nadir pointing, both missions can also point at the sun glint spot, greatly increasing the signal over the oceans, which do not otherwise provide much reflection in the near IR (Miller et al., 2007). GOSAT was launched in January 2009, OCO in February 2009; GOSAT successfully reached its operational orbit, OCO did not. While GOSAT the measurements should greatly expand our knowledge of the global carbon cycle, the OCO design had certain strong points that have led to a push for a relaunch, possibly as early as 2012. OCO would measure more frequently than GOSAT (180 vs. 13.4 cross-scans per minute) with a smaller FOV ($\sim$2 km^2 vs. $\sim$100 km^2) and thus ought to find more cloud-free scenes (Crisp et al., 2004) with low X_{CO_2} retrieval errors.

In this study, we use an atmospheric inverse method to quantify how well X_{CO_2} measurements from OCO would help estimate sources and sinks of CO_2 at the surface. A tracer transport model relates simulated atmospheric CO_2 concentrations to the surface CO_2 fluxes at earlier times that determined them. Progressively higher layers in the atmospheric column reflect the influence of fluxes from increasing broad areas at the surface, due to atmospheric mixing. The transport model allows this X_{CO_2} measurement information, weighted properly in the vertical column, to be distributed appropriately to fill in the 25° gaps between subsequent OCO passes on any given day. Though OCO cannot

clarify the diurnal cycle of flux, it can shed light on flux variability due to synoptic-scale weather systems when they are modeled well by the transport model. Previous global CO_2 flux inversions using data from the global in situ measurement network have most often used the "Bayesian synthesis" inversion approach (Enting et al., 1995). This method has also been used to determine the information on surface CO_2 fluxes provided by satellite data (Rayner and O'Brien, 2001; Houweling et al., 2004; Miller et al., 2007), although only for monthly fluxes from fairly large emission regions ($\sim$2000 km on a side) since the number of fluxes solved for was limited by the inversion method. The density of OCO's data should permit fluxes to be estimated at a finer resolution than this, but a more computationally-efficient inversion method is required.

We use a state-of-the-art variational data assimilation scheme (Baker et al., 2006b) to solve for the CO_2 fluxes at the horizontal resolution of our transport model; optimized time-varying 3-D CO_2 concentration fields are also produced as a by-product. The fluxes are solved at a weekly resolution, though the measurements are modeled at the time step of the transport model (1 h). Our data assimilation approach is used to perform observing system simulation experiments (OSSEs) in which simulated data and measurement errors are input to produce statistics on the flux estimation errors and the improvement in the initial guess of the fluxes. Both Baker et al. (2006b) and Chevallier et al. (2007a) have done preliminary OSSEs for OCO using this approach before. For measurements, they assumed a single measurement per model grid box with a 1 or 2 ppm uncertainty value (1σ), respectively, and with a flat weighting versus pressure in the vertical. Here, we improve upon their assumptions in two ways. First, for each individual retrieval, we use new OCO X_{CO_2} retrieval uncertainties and averaging kernels (AKs) calculated as a function of surface type, solar zenith angle, aerosol optical depth (OD), and pointing mode (nadir vs. glint) using the OCO Level 2 X_{CO_2} retrieval scheme forced with radiances simulated by the OCO "full-physics" radiative transfer scheme, taken from Bösch, et al .(2010). Second, instead of assuming only a single valid retrieval per crossing of each model grid box (which takes $\sim$33 s for our 2° $\times$ 5° boxes), we collapse the information content of all valid retrievals across each grid box crossing into an equivalent multi-sounding measurement uncertainty, which is then used in the assimilation. Valid X_{CO_2} retrievals are only attempted for cloud-free conditions in which the aerosol OD is less than 0.30, in order to reduce associated radiative transfer modeling errors. We compute the number of valid retrievals for each grid box crossing based on the probability that such cloud-free and low-aerosol conditions exist for each retrieval; these probabilities are computed using climatological statistics from MODIS data. We attempt to account for along-track correlations in the X_{CO_2} measurements when specifying the equivalent measurement uncertainty for each model grid box crossing. Finally, we examine more types of systematic errors than
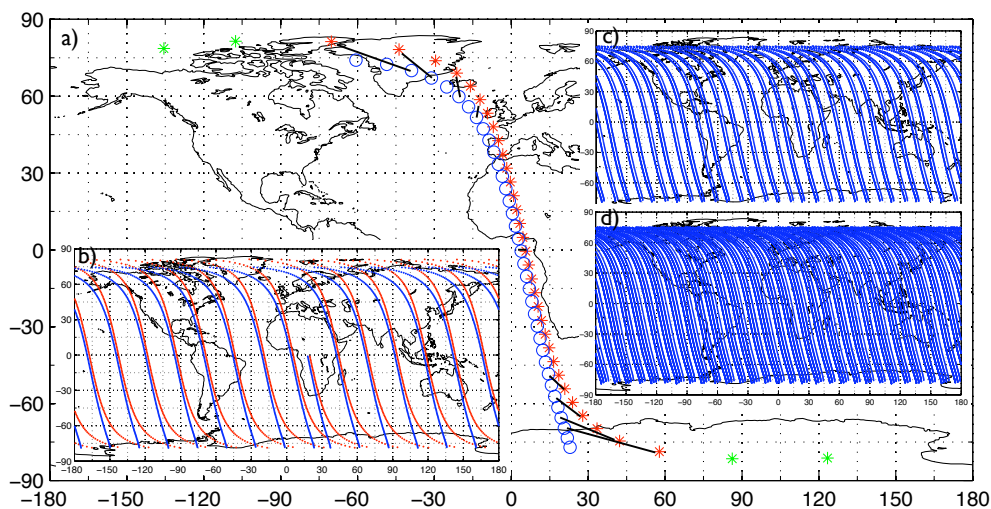
Fig. 1. (a) An example of the field-of-view (FOV) ground tracks for OCO for 21 March: 100 min of measurements for nadir pointing mode (asterisks) and glint (circles). Black lines connect nadir and glint FOVs at same time. The maximum SZA is taken as 85°/80° for nadir/glint. Green asterisks indicate positions where nadir SZA$\geq$85° and glint SZA$\leq$80°. **(b)** One-, **(c)** four-, and **(d)** seven-day coverage for nadir (red) and glint (blue) beginning 21 March.

these previous studies: measurement biases due to aerosols, transport errors, and errors caused by "mistuning" the inversion (i.e., assuming incorrect a priori flux and measurement error statistics). Feng et al. (2009) used the Bösch et al., OCO retrieval errors in an OCO OSSE study similar to this one, but with an ensemble Kalman filter approach. Chevallier et al. (2009) have recently performed a similar OSSE to evaluate the flux constraint provided by GOSAT, using variable measurement uncertainties appropriate for that satellite.

2 Method

2.1 OCO orbit and resolution choices

The OCO satellite measures X_{CO_2}, the column-averaged dry air fraction of CO_2, in the near-infrared (reflected solar) band with sensitivity down to the surface, but with a vertical weighting that varies with surface type, aerosol amount, and solar zenith angle (SZA) as described in Bösch et al. It samples eight fields of view (FOV), each with an area $\leq$2.8 km^2, every 333 milliseconds across an FOV ground track up to 10 km wide (Crisp et al., 2004), of which only four are downlinked. It is in a sun-synchronous orbit taking a single sun-lit pass of data per day every 24.7° in longitude; we asume a 13:18 local ascending node time here. Examples of the sunlit portion of the OCO FOV ground track are given in Fig. 1. The OCO ground track repeats precisely after 16 days, a fact that is useful for calibrating the measurements at fixed ground sites. However, as shown in Fig. 1, the ground tracks also achieve a somewhat uniform spatial coverage of ∼3.5° in longitude after only 7 days: we use this 7-day period as the discretization step for our solved-for fluxes, since it gives

good coverage over our transport model grid boxes, 5° wide in longitude. The latitudinal resolution of the model is chosen at 2° to match that of our meteorological products to give maximum resolution in the predominantly north/south (N/S) direction of the OCO ground tracks. Because the OCO data, sampled only once per day locally, provide little information on the diurnal cycle of X_{CO_2}, some assumption for the diurnal cycle of the surface CO_2 fluxes must also be made (see Sect. 2.4 below); this then allows multi-day flux blocks to be estimated in a reasonable way from the data.

2.2 Transport model

An off-line atmospheric transport model ("PCTM": see Kawa et al., 2004) is used to relate surface CO_2 fluxes to CO_2 concentrations. It is driven by pre-calculated meteorological fields (horizontal winds, surface pressure, vertical diffusion coefficient, and cloud-convective mass flux) from the GEOS4-DAS reanalysis (Bloom et al., 2005) for the year 1987, interpolated from the resolution normally input to PCTM (2.0° × 2.5° in lat/lon; 55 vertical layers) to the resolution of the model version used here (2° × 5° lat/lon; 25 vertical layers). The model uses a vertically-Lagrangian finite volume advection scheme (Lin, 2004) and has simple linear schemes for both dry and convective vertical mixing.

The modeled 3-D concentration fields are sampled in as similar a manner to the true OCO X_{CO_2} measurements as the transport model permits: vertically, using the averaging kernels computed by Bösch et al., as a function of surface type, SZA, aerosol OD, and nadir or glint viewing mode; horizontally, at the transport model's 2° × 5° resolution; and temporally, at the model's integration time step (1 h).

The adjoint of the transport model is needed in the assimilation scheme to move model-data misfit information backwards in time to compute the cost function gradient. The adjoint of the forward model has been computed in an efficient manner by running a linear version of the forward advection scheme backwards, and by computing the exact adjoint of the vertical mixing schemes' column mixing matrices. The adjoint is accurate to within ∼0.05% across a two-week run (as computed using the definition of the adjoint, i.e., comparing $(\mathbf{M}(x))^T \mathbf{M}(x)$ to $x^T \mathbf{M}^T (\mathbf{M}(x))$, for point perturbations in the initial concentration field x, where $\mathbf{M}$ represents the forward transport operator and $\mathbf{M}^T$ the adjoint). As shown in Baker et al. (2006b), this adjoint allows the true fluxes to be recovered to within 0.2% after 60 iterations in a perfect-model simulation with no measurement errors added.

2.3 Data assimilation scheme

We solve for weekly surface CO_2 fluxes at $2° \times 5°$ in lat/lon, using simulated X_{CO_2} measurements across a data span of 1 year. Both the number of fluxes to be solved for ($90 \times 72 \times 52 = \sim 35\,000$) and the number of data values used ($365 \times 1500 = \sim 50\,000$) are at least an order of magnitude larger than that used in typical past time-dependent CO_2 inversions of in situ data (e.g., Rödenbeck et al., 2003; Peylin et al., 2005b; Patra et al., 2005; Baker et al., 2006a; Rayner et al., 2008). Most of these previous inversions used the "Bayesian synthesis method", a batch least squares technique in which transport basis functions were constructed in separate model runs, either one for each solved-for flux or (backwards in time using the adjoint) one for each measurement, to fill a Jacobian matrix relating fluxes to concentrations. The resulting system of linear equations was solved directly to give both the optimal estimate and the accompanying covariance matrix describing the estimation errors. For problems of the size addressed here, this sort of direct (non-iterative) method is not computationally feasible and a more efficient approach is needed.

We have chosen to use a variational data assimilation approach to overcome these hurdles. It is similar to the "4-D Var" methods used in numerical weather prediction, except that instead of optimizing an initial condition (the atmospheric state) at the start of a relatively short assimilation window, we optimize time-varying boundary values (surface CO_2 fluxes) over a longer span. Baker et al. (2006b) outline the mathematical details and give some test results using simulated data. Rödenbeck (2005) has used a similar approach to estimate daily CO_2 fluxes from 20+ years of in situ CO_2 measurements, and Meirink et al. (2008) have recently used this method to estimate surface CH_4 fluxes on a fine grid from SCIAMACHY data. Rayner et al. (2005) have used a variational approach for solving directly for parameters in land biosphere carbon models, bypassing the surface fluxes. Over the past several years, a new class of ensemble filtering methods have also been applied to the tracer transport problem (Peters et al., 2005; Zupanski et al., 2007; Feng et al., 2009). Both the ensemble and variational methods achieve their computational savings in a similar fashion: by solving for only an approximate, low-rank version of the full a posteriori covariance matrix. The ensemble filters have the advantage of not requiring an adjoint and are easier to implement, but they also introduce approximations that may degrade the estimate. We have chosen to go with the proven computational savings of the variational methods for this study.

The variational method works in an iterative fashion, running an estimate of the surface fluxes forward in time through the transport model to derive modeled measurements, comparing these to the true measurements, and running these measurement residuals (weighted using assumed measurement error statistics) backwards in time through the adjoint of the transport model to obtain flux corrections, then repeating. The flux inversion is posed mathematically as a minimization problem, with the adjoint run providing the gradient to the measurement portion of the cost function. We use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method to solve it.

2.4 Simulation approach

The assimilation seeks to drive an initial (a priori) guess of the fluxes towards the real-world ("true") fluxes, using the measurements. In our simulations here, we generate measurements with different error sources added on that attempt to describe the real errors OCO will encounter when it actually flies, then process the measurements with the assimilation method in the same way that we would do with the real data. Since we know the fluxes used in generating the data, we can compare the estimated fluxes to these "true" values to get actual estimation errors. If only random estimation errors are added to the data (see Experiments 1 and 2, Sect. 2.6), the statistics of these estimation errors should be consistent with what would be given by the full-rank covariance matrix, if one were computed. To approximate the uncertainties that would be given by the covariance matrix, we accumulate our random estimation error statistics over seasons (13 weekly flux values) and over a full year (52 values).

Our simulation approach has the added benefit of allowing us to quantify the impact of systematic errors, such as measurement biases or errors in the transport model, with the same statistics as for the random error experiments. In the first case, the biases are added when simulating the true measurements; in the second case, different winds and vertical mixing parameters are used in the optimization than are used to generate the truth.

For our true fluxes, we use monthly land biospheric fluxes from the LPJ model (Sitch et al., 2003) and monthly ocean fluxes from a biospheric run of the NCAR ocean model (Doney et al., 2006; Najjar et al., 2007); both are interpolated to daily values. For our a priori fluxes, we use similar fluxes from the CASA land biosphere model (Randerson et
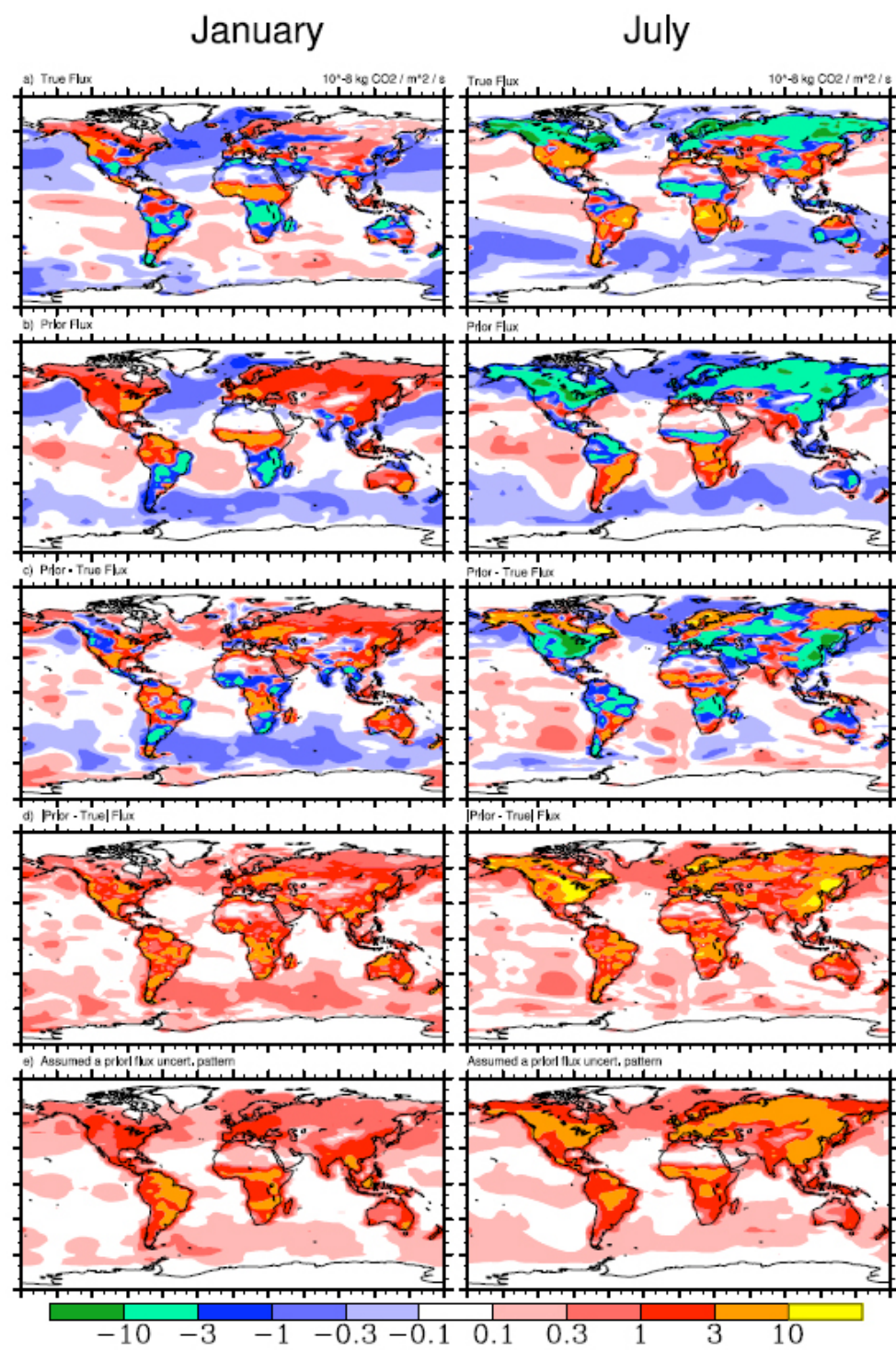
Fig. 2. January (left) and July (right) mean values for **(a)** the "true" surface CO$_2$ fluxes (LPJ land + NCAR ocean); **(b)** the a priori CO$_2$ fluxes (CASA land + Takahashi ocean); **(c)** the prior-truth flux difference; and **(d)** |prior-truth|. The values in **(d)** are used in the assumed a priori flux error covariance matrix for all experiments except Experiment 3, the mistuning experiment, which used the values in **(e)**. All in $[10^{-8}$ kg CO$_2$ m^{-2} s$^{-1}]$.

al., 1997) and the Takahashi et al. (1999) ocean CO_2 flux product. Figure 2a–c gives snapshots of both sets of fluxes for January and July, as well as their difference. While both sets of fluxes show similar features (e.g., the seasonal cycle of net photosynthesis minus respiration in both the northern and tropical land vegetation, uptake of CO_2 by the extratropical oceans versus outgassing by the tropical oceans), their timing and spatial details vary enough that the prior-truth difference (Fig. 2c) is often as large as the fluxes from either model: there is much room for improvement, even if the models appear to be doing a fair job, superficially.

The prior-truth flux differences (Fig. 2c) show systematic spatial and temporal correlations. The spatial correlations are often at fine scales, many times associated with deserts and mountain ranges: thin lines of $\pm$ values running parallel to the Canadian Rockies, for example. Because of the physical basis of these differences, we have some hope that the differences between our two sets of models will bear some resemblance to the difference between any one model and the real-world fluxes. The Bayesian prior in our cost function performs the useful function of damping out spurious noise in the estimate due to noise in the measurements (or, more accurately, in the model-measurement mismatches). However, inaccuracies in our knowledge of the a priori flux error covariance, $\mathbf{P}_o$, including both correlations and the overall magnitude of the variances, will degrade the final assimilated estimate. We use a diagonal $\mathbf{P}_o$ with variances set equal to the square of the actual weekly prior-truth flux difference (Fig. 2d) in most of our assimilation experiments (see Sect. 2.6), but also use an less precise estimate (Fig. 2e, based on the magnitude and variability of the prior fluxes) in a sensitivity experiment to examine the impact of realistic errors in the assumed $\mathbf{P}_o$. It is possible that we could have constructed a $\mathbf{P}_o$ with off-diagonal elements (correlations) that would better represent our prior-truth flux difference; since this would presumably lead to better-converged results, we should obtain conservative results using our diagonal $\mathbf{P}_o$.

We have not included fossil fuel fluxes in these simulations: errors in our best estimate of the fossil fuel source are thought to be small at our $2° \times 5°$ resolution. The net flux uncertainties we obtain over land should thus be thought of as applying to the sum of the fossil and land biospheric fluxes. Similarly, the diurnal cycle of flux is not modeled here, since the OCO data, taken at a single local time per day, cannot resolve it. Insofar as the OCO data are biased with respect to daily mean X_{CO_2}, the resulting CO_2 flux estimates will be biased as well; this error term is not quantified here.

2.5 X_{CO_2} measurement errors and averaging kernels

The assimilation requires a statistical description of the errors in individual X_{CO_2} measurement retrievals, as well as knowledge of the averaging kernel (AK— how strongly each vertical layer contributes to the column average). Bösch et al. have obtained new estimates of both quantities as a func-

tion of surface type, SZA, aerosol OD, and pointing mode (nadir vs. glint) (Fig. 3). They used a detailed radiative transfer scheme to simulate the radiances seen in the measured OCO spectral bands, then fed these through the OCO "full-physics" X_{CO_2} retrieval scheme, testing sensitivities to various error sources. We use these error and AK estimates, along with surface FOV locations and SZAs taken from an accurate OCO orbit generator for both nadir and glint pointing modes, to calculate realistic values single-sounding X_{CO_2} retrieval errors and AKs around the orbit.

There are potentially hundreds of separate measurements (with FOV areas $\leq 2.8\,km^2$) along the FOV ground track swath for any single crossing of our $2° \times 5°$ atmospheric model grid boxes. Since these measurements are taken over an often heterogeneous surface with different reflective properties and CO_2 emissions, with varying cloud and aerosol amounts interfering with the retrieval, the measurement errors along the swath could be quite variable. When averaged across the grid box, the uncorrelated portion of these errors could be expected to cancel out significantly. We make an attempt here to estimate what portion of this error cancels out and what does not, to quantify the effective measurement error of all the valid retrievals inside each model grid box. In computing this effective error, we consider the probability of obtaining cloud-free retrievals with aerosol ODs lower than a 0.30 cutoff, and we model correlations along the orbit as a function of SZA. The along-orbit computation of the AKs and single- and multi-sounding retrieval uncertainties are done first at a $1° \times 1°$ resolution, then translated to the $2° \times 5°$ model grid box resolution used in the assimilation based on the time spent in each $1° \times 1°$ area inside the $2° \times 5°$ box. We show annual mean plots here for the uncertainties and quantities used to compute them, but they vary monthly in the simulations (see the Supplementary Material for seasonal plots; http://www.atmos-chem-phys.net/10/4145/2010/acp-10-4145-2010-supplement.pdf).

2.5.1 Single-sounding X_{CO_2} errors and supporting fields

The calculation of the SZA and the FOV location on the surface, required for the X_{CO_2} error and AK calculations, both depend on an accurate orbit propagation. For nadir mode, the FOV is located at the sub-satellite point. For glint mode, the surface normal at the glint spot is computed by iteration until the surface normal is the same angle from the sun and the satellite position vectors, in the plane they define. In both pointing modes, the surface normal is computed assuming the Earth is an oblate spheroid. The orbit is taken as sun synchronous, with a 13:18 local time of ascending node, $a=7083.45\,km$, $e=0.0012$, $i=98.2°$. The anomaly is chosen arbitrarily to have the spacecraft crossing north across the equator at 00:00:00 on 1 January.

Figure 4a gives the distribution of the five surface types used to calculate the X_{CO_2} errors and AKs: ocean/water,
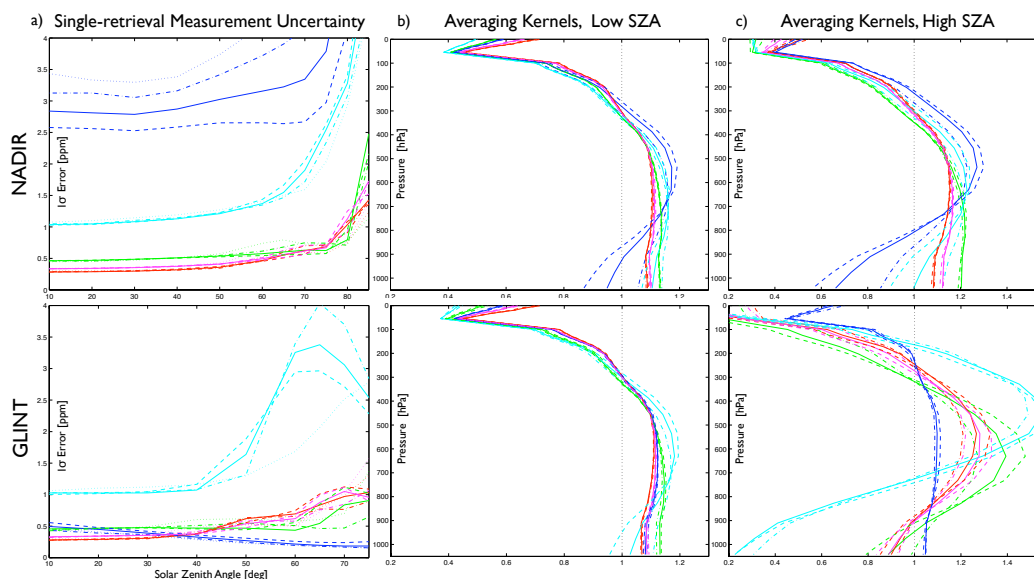
Fig. 3. A summary of **(a)** the single-sounding OCO X_{CO_2} uncertainties [ppm] and **(b)** and **(c)** normalized averaging kernels (AKs) for nadir (top) and glint (bottom), from Bösch et al., for five surface types (conifer, green; desert, red; sparse vegetation, magenta; snow, cyan; and ocean, blue) and for four 760 nm aerosol ODs (0.00, dotted; 0.10, dot-dash; 0.20, solid; and 0.30, dashed). The AKs in (b) have been averaged over SZAs of $20° − 50°$; those in (c) over $60° − 75°$ for glint and $60° − 80°$ for nadir.

snow/ice, desert, conifer (representing all types of dense vegetation), and sparse vegetation/exposed soil. Figure 4b gives median total aerosol ODs derived from Aqua/MODIS data. The aerosol OD histograms used to compute these medians are described in more detail in Bösch et al. Computed solar zenith angles as a function of latitude for four seasons are given in Fig. 5a. Finally, the OCO single-sounding X_{CO_2} retrieval uncertainties calculated from these fields are given in Fig. 6a for both nadir and glint pointing modes. The most noticeable feature of Fig. 6a is how much lower the uncertainties are over the oceans in glint mode as compared to nadir mode. Note also, however, that they are somewhat lower over the land in nadir mode compared to glint.

2.5.2 Computing effective multi-sounding X_{CO_2} errors

Our ability to represent the OCO X_{CO2} retrievals is limited by the fairly coarse spatial resolution of our transport model: our ∼220 km wide grid boxes cannot represent the X_{CO_2} variability occurring in the real world at shorter spatial scales. However, for the purposes of estimating CO2 concentrations and fluxes at scales of 100s to 1000s of km, there is no need to model every ∼2 km^2 X_{CO_2} retrieval correctly. The real question is: how close does the average of all the X_{CO_2} measurements taken inside a model-scale grid box come to the average of all true X_{CO_2} values across the full area of that grid box (not just inside the ∼10 km-wide OCO FOV track)? We model the latter quantity.

The first point to note is that even if the X_{CO_2} measurements are perfect and complete (no data gaps due to clouds or aerosols) across the full length of the 10 km-wide FOV ground track, there will still be a difference between this perfect ground track average and the average X_{CO_2} across the full grid box. Second, the perfect X_{CO_2} measurements may not even get the ground track average correct, because of non-uniform coverage (data gaps) due to clouds and aerosols. And, third, the X_{CO_2} measurements are obviously not perfect, but are subject to the measurement errors discussed above. When all the X_{CO_2} measurements inside a grid box are averaged together, their errors may cancel out to some extent in the average, but there will still be a remaining error between the average measurement and the true X_{CO_2} value for the measured portion of the ground track. All three of these errors – track-to-box representation error, along-track representation error, and average effective measurement error – must be combined to get the model-measurement mismatch error that should be fed into the flux error simulations.

The first two of these error sources have been examined by Corbin et al. (2008). They did detailed simulations of X_{CO_2} variability inside domains of $1° × 1°$ and $4° × 4°$ using a mesoscale atmospheric transport model, comparing the X_{CO_2} averages along an OCO-like FOV ground track to the average values across the full domain to obtain estimates of the track-to-box representation errors. They also simulated the effect of clouds on the availability of OCO retrievals, coming up with realistic estimates of the along-track representation errors. For the two sites they examined, they concluded that the along-track representation error was small compared to the track-to-box representation error. They also concluded
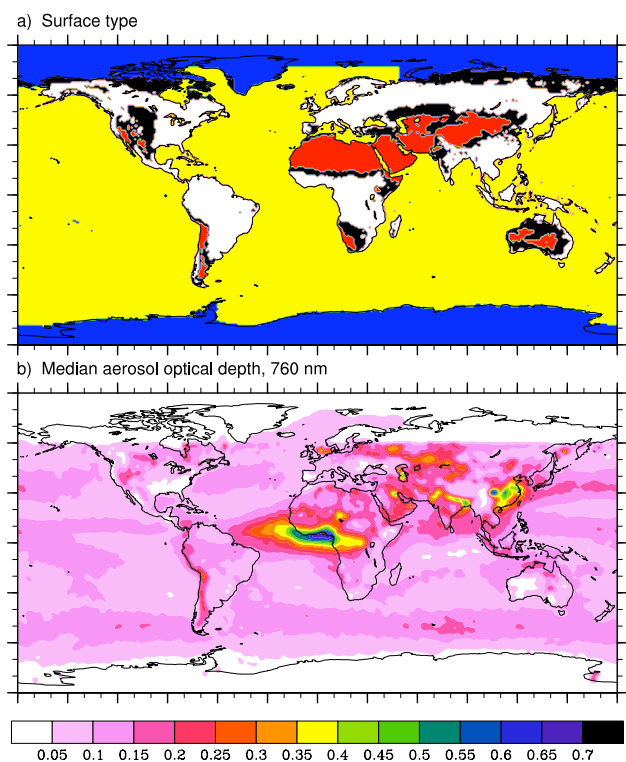
a) Surface type

b) Median aerosol optical depth, 760 nm

0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5 0.55 0.6 0.65 0.7

Fig. 4. (a) The five surface cover types assumed: desert (red), conifer (white), ocean/water (yellow), snow (blue), and soil/sparse vegetation (black). **(b)** The median aerosol OD at 760 nm computed from Aqua/MODIS data according to the procedure outlined in Bösch et al. (annual mean of four seasonal medians).

that the track-to-box error was, in turn, largely random and relatively small compared to the measurement errors. In our study here, we neglect the along-track errors, and extrapolate the Corbin et al., track-to-box representation errors from their two sites to the full globe using a fit proportional to the absolute value of the net ocean or land biosphere flux from our monthly-varying a priori flux model inside each $1° \times 1°$ grid box (Fig. 6c, with a proportionality factor of $2.5 \cdot 10^6$ ppm/(kg CO$_2$ m^{-2}s^{-1})). These track-to-box representation errors are taken to be unbiased and gaussian, and are added in all the simulation cases presented here.

The third error source, the effective joint error of all the individual X_{CO_2} measurements inside a grid box, is the largest over almost all of the globe at all times of the year. To compute it, one must factor in data gaps due to cloud coverage or aerosol ODs greater than 0.30 (the level beyond which the OCO retrievals will not be routinely performed). Furthermore, one must estimate the error correlation along the ground track of near-by measurements. Here we assume that errors from aerosols and clouds will dominate the correlated errors (both directly by causing single-sounding retrieval biases that are correlated along-track, and indirectly by introducing data gaps of finite extent that cause representation er-
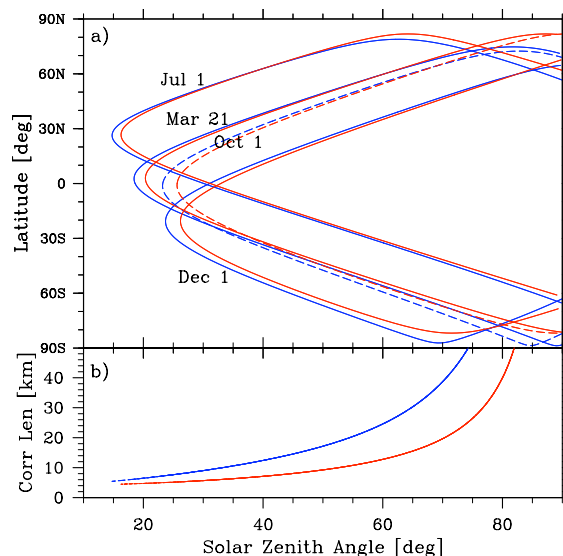

Fig. 5. (a) The solar zenith angles (SZA) encountered in nadir (red) and glint (blue) pointing modes for four times of the year, plotted against FOV latitude. (The 1 October–12 March difference reflects the east/west shift in the Sun's position in the analemma). **(b)** The correlation length L beyond which measurement errors are assumed to be independent, for nadir (red) and glint (blue), as given by Eq. (1).

rors) and that their correlation lengths increase with SZA and path in atmosphere. We represent this with a simple ad hoc correlation length L (Fig. 5b):

$$L^2 = (c_w^2 + (c_h P \tan(SZA))^2) \tag{1}$$

where c_w is a fine-scale cloud width (taken here as 4 km), c_h is a typical average cloud height (taken here as 7 km), and P is a path-length factor (taken as 1 for nadir pointing mode and 2 for glint). The maximum number of possible independent measurements inside a $1° \times 1°$ grid box is then taken to be $N_{max} = l_{1 \times 1}/L$, where $l_{1 \times 1}$ is the OCO FOV ground track path length inside the box. This maximum value is reduced by the availability of data due to clouds and aerosols, giving N_{eff}, the effective number of independent X_{CO_2} measurements inside the $1° \times 1°$ grid box, as

$$N_{eff} = P_{cloud-free}(1 - P_{HiAeroOD})l_{1 \times 1}/L \tag{2}$$

where $P_{HiAeroOD}$ is the probability of aerosol ODs exceeding the 0.30 value beyond which OCO X_{CO_2} retrievals are not attempted, and $P_{cloud-free}$ is the probability of finding at least one cloud-free scene in a swath of OCO FOV ground track of length L. $P_{HiAeroOD}$ is computed from the same aerosol OD histograms as the median aerosol ODs, from Bösch et al. $P_{cloud-free}$ is computed from climatologies of Aqua/MODIS and Terra/MODIS data, sampled in 10 km-wide swaths, as detailed in the Appendix.

Both aerosol and cloud coverage are calculated using data from the MODIS instrument aboard NASA's Aqua satellite,
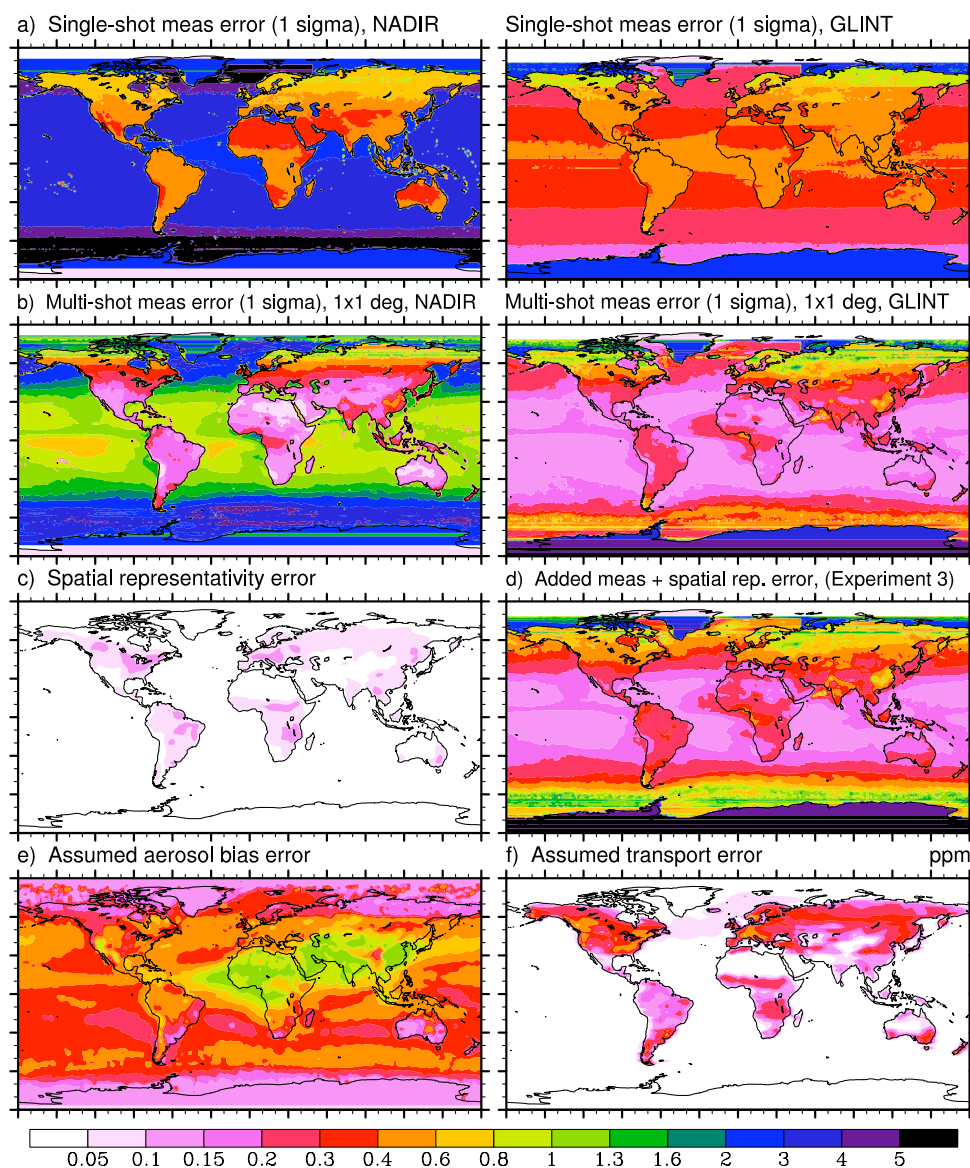
Fig. 6. **(a)** The single-sounding OCO X_{CO_2} retrieval uncertainties σ_{1shot} computed in Bösch et al., for both nadir (left) and glint (right) viewing modes. **(b)** The effective multi-sounding OCO X_{CO_2} measurement uncertainty σ_{eff}, computed as $\sigma_{eff} = \sigma_{1shot}/\sqrt{N_{eff}}$. using N_{eff} from Fig. 7a, **(c)** The assumed spatial representation error, extrapolated from Corbin et al. (2008). **(d)** The random measurement error added to the data (in place of σ_{eff} in Fig. 6b) in Experiment 3, the mistuning experiment. The extra measurement uncertainty assumed to account for the impact of **(e)** aerosol biases and **(f)** transport errors.

which flies in the same "A-train" orbit as OCO will. MODIS has a 1×1 km FOV that, being close to the ~2 km^2 OCO FOV area, should give realistic idea of cloud free areas and aerosol amounts over most areas. Since the MODIS instrument scans up to 45° off-nadir, the sensed radiation actually passes though a slightly longer path than that for OCO in nadir mode, encountering if anything more clouds and aerosols. For OCO in glint mode, however, the path length of the radiation in the atmosphere can be quite a bit longer than that sensed by MODIS. To account for the increased prob-

ability of encountering clouds and aerosols at SZAs greater than 20° in glint mode, we use:

$$P_{cloud-free} = P_{cloud-free\_MODIS}^{(2/(1+\cos(SZA)/\cos(20°)))} \qquad (3)$$

while $P_{HiAeroOD}$ is recomputed by shifting the 0.30 OD cut-off to a lower value of $0.30 \cdot (2/(1+\cos(SZA)/\cos(20°)))$ and summing aerosol OD histogram to the right of this new value. Once N_{eff} is calculated, the effective measurement error accounting for all X_{CO_2} measurements inside each $1° \times 1°$ grid box, is given as: $\sigma_{eff,1\times1} = \sigma_{1shot}/\sqrt{N_{eff}}$. The effective

a) N_Effective per 1x1 deg box, NADIR   N_Effective per 1x1 deg box, GLINT

b) Prob of a cloud-free shot across L, NADIR   Prob of a cloud-free shot across L, GLINT

c) Probability aerosol optical depth > 0.30 (NADIR)   Probability aerosol optical depth > 0.30 (GLINT)

0.05 0.1 0.2 0.3 0.4 0.6 0.8 1 1.4 2 4 6 8 10 12 15 20
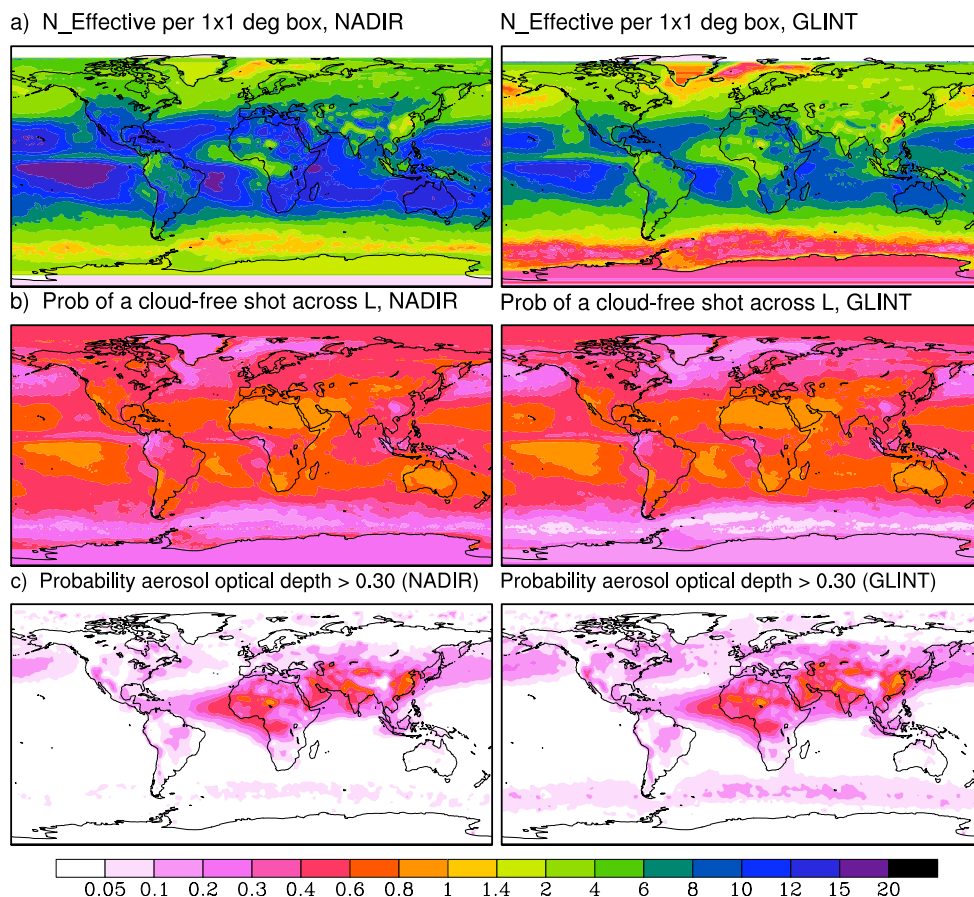
Fig. 7. **(a)** The effective number of independent X_{CO_2} measurements N_{eff} in each $1°$ latitude band for a single sun-lit pass of the OCO orbit for both nadir (left) and glint (right), computed with Eq. (2). **(b)** The probability $P_{cloud-free}$ of finding at least one cloud-free X_{CO_2} measurement across an OCO FOV ground track of length L (Fig. 5b), calculated from MODIS data according to the procedure outlined in the Appendix. **(c)** The probability $P_{HiAeroOD}$ of encountering 760 nm aerosol ODs greater than 0.30, from Aqua/MODIS data.

measurement uncertainties at $2° \times 5°$ resolution used in the assimilation are then computed from these $1° \times 1°$ values, based on the distance $l_{1\times1}$ and $l_{2\times5}$ inside each $1° \times 1°$ and $2° \times 5°$ box, as:

$$l_{2\times5}/\sigma^2_{\text{eff},2\times5} = \sum_i l_{1\times1,i}/\sigma^2_{\text{eff},1\times1,i}.$$

Figure 6b gives the distribution of $\sigma_{\text{eff},1\times1}$ and Fig. 7, N_{eff}, along with the $P_{\text{cloud-free}}$ and P_{HiAeroOD} values used to compute them. Figure 7b, c shows that both persistent cloudiness and areas of high aerosol contamination significantly reduce the availability of OCO measurements in this approach. The $\sigma_{\text{eff},1\times1}$ values in Fig. 6b are substantially higher than the track-to-box representation errors given in Fig. 6c, by generally more than a factor of 5. The areas of low error in Fig. 6b, c show where the measurements with the greatest information content will occur; the assimilation convolutes these with transport to determine where the flux constraints will be the strongest.

2.6 Flux estimation simulations

The main objective of our study is to perform a series of OSSEs meant to represent how well our data assimilation system will estimate surface CO_2 fluxes, given the presence of various error sources. We somewhat arbitrarily divide these errors into purely random ones (modeled as unbiased, gaussian noise) and biases constant in space and time. In reality, of course, there is a spectrum of errors that are correlated in both space and time that fall between these extremes, due to correlations in such error-causing factors as scattering due to aerosols and undetected clouds, spectral effects, and surface reflectance properties. We have attempted to account for some of these terms above by transforming the correlated errors into the corresponding purely random problem using the idea of "effective independent measurements". Since the finest-resolution unit that the atmospheric transport model, and thus the atmospheric flux assimilation, can deal with is the transport model grid box at the model time step, both random and systematic errors are quantified at that scale: what

is the net bias or random error between the weighted average of all measurements in a grid box (from a single crossing) and the true concentration in that box?

Table 1 outlines a series of assimilation experiments we perform, with the error sources that have been added in each case. Two of the sources of error described above – the "track-to-box" representation errors and the random measurement errors – have been added in all the experiments as gaussian noise. Biases due the representation errors were found to be small in Corbin et al. (2008) and are not added here at all. Systematic errors in the measurements have been added onto true measurements in Experiments 4–6 (Table 1) as described below. Whenever these systematic errors are added, we increase the uncertainties assumed in the measurement error covariance matrix, **R**, in an attempt to account for them. Although it is not formally valid, statistically, to represent systematic errors with random ones, it is often done and is certainly better than not attempting to account for the biases, since in that case the measurements would be given too much weight vis-a-vis the prior and the impact of the biases would be greater than if the measurements had been de-weighted (Chevallier, 2007c). In all experiments, both the measurement error and a priori flux error covariance matrices, **R** and **P**$_o$, are diagonal: we account for measurement correlations inside a grid box by computing the effective number of independent measurements and adjusting the multi-sounding measurement uncertainties accordingly; measurement correlations between grid boxes are neglected; both time and space correlations between the estimated weekly fluxes are neglected, since using a 2° × 5° grid box already effectively imposes a fairly coarse correlation length.

Our control experiments (1 and 2) examine the impact of only random measurement errors in the nadir and glint mode data. There is no transport error: the same model that was used to generate the true data is used in the assimilation. There are no measurement biases added, only random measurement errors. And the assimilation is well "tuned": both the assumed measurement error covariance matrix and the assumed a priori flux estimation error covariance matrix are chosen to be consistent with the statistics of the added measurement errors and of the prior-truth flux errors, respectively. With these assumptions, the flux errors that result from the assimilation should agree with the error statistics that would be given by the a posteriori flux covariance matrix of inverse methods that produce one (our assimilation here does not produce a full rank covariance matrix, only a low-rank approximation not useful for quantitative error analyses at the fine scales examined here). Such a posteriori covariance matrices are often the end product of error analyses and are useful for quantifying the precision of the assimilation (the standard deviation of errors about the mean estimate), though not the accuracy (the standard deviation of errors about the truth) since they do not quantify the impact of systematic errors. The variances in the a priori flux error

covariance matrix were taken to be the square of the actual prior-truth flux difference given in Fig. 2c.

The remainder of the tests were done only for glint viewing mode; Experiments 3–5 differ from Experiment 2 in that a different source of systematic error was added in each case. In Experiment 3, we add more realism by "mistuning" the assimilation, adding realistic errors to both the assumed a priori flux error and measurement error covariance matrices. Instead of making the a priori flux uncertainties proportional to the actual prior-truth flux difference (Fig. 2d), we use uncertainties based only on our a priori flux patterns (Fig. 2e) since, in real world simulations, we have no knowledge of the true fluxes. To mistune the assumed measurement error covariance matrix, **R**, we actually change the added measurement uncertainties from the glint mode values in Fig. 6b to those shown in Fig. 6d; we keep the assumed values the same as in the other experiments to allow the cost function values to be compared with the other experiments more readily. To obtain the values in Fig. 6d, we simplified the SZA-dependent glint mode X_{CO_2} retrieval errors (Fig. 3a) as follows: for the conifer and sparse vegetation surface types, the measurement errors were taken to be 0.60 and 0.50 ppm, respectively, for SZAs under 55°, and 0.70 and 0.90 ppm over 55°; over deserts and snow, 0.40 and 1.10 ppm under 45°, and 0.75 and 3.00 ppm over 45°; and over water, 0.40 ppm for all SZAs.

Biases due to aerosols are expected to cause the main systematic errors in the OCO X_{CO_2} retrievals (Connor et al., 2008). In Experiment 4, we add a bias of $+\alpha \cdot$aero_OD to all measurements over land and ice-covered areas, and a bias of $-\alpha \cdot$aero_OD over the ocean, where aero_OD is the seasonally-varying median aerosol OD (Fig. 4b) and $\alpha = 2$ ppm/OD; the maximum bias is ±0.6 ppm, since no X_{CO_2} retrievals are attempted for aerosol ODs greater than 0.3. The magnitude of these assumed aerosol biases is generally larger than the (1σ) multi-sounding random measurement uncertainties over land, especially over Africa and central/southern Asia. To account for this extra error in the assimilation, we add the aerosol bias uncertainties given in Fig. 6e to the assumed multi-sounding random measurement uncertainties (Fig. 6b and c) in quadrature. (The values added to the assumed errors (Fig. 6e) are actually twice as high as the added biases to account for two effects: a) the assumed errors at 1° × 1° in Fig. 6e will drop by a factor of $\sqrt{2}$ when averaged across the 2°-wide grid boxes on which scale the biases are added, and (b) 50% of the area under a gaussian curve falls withing ±0.676σ, requiring a larger 1σ value when attempting to represent a bias; $\sqrt{2}/0.676 = 2.09 \approx 2$.)

Atmospheric transport models have a variety of inaccuracies, not only in their representation of the broad-scale general circulation, but also in their smaller-scale mixing processes (especially between the planetary boundary layer and the free atmosphere) and in their ability to represent fine scale in situ or satellite data, that impact the inverted flux estimates. In Experiment 5 we add a simple approximation of

Table 1. The errors added to the true measurements and the random error sources assumed in the assimilation for the various OSSEs. Figure numbers are given for annual summary plots of the various added or assumed errors (e.g., "6b"). N=nadir, G=glint.

Experiment	#	Nadir/Glint	Measurement Noise Added and Assumed	Prior Flux Uncert. Assumed	Aerosol/Transport Biases Added?	Extra Aerosol/ Transport Noise Assumed
Control, Nadir	1	N	6b, 6b	2d		
Control, Glint	2	G	6b, 6b	2d		
Mistuned	3	G	6d, 6b	2e		
Aerosol bias	4	G	6b, 6b	2d	✓/✗	6e/–
Transport error	5	G	6b, 6b	2d	✗/✓	–/6f
Mistuned+aerosols+transport	6	G	6d, 6b	2e	✓/✓	6e/6f
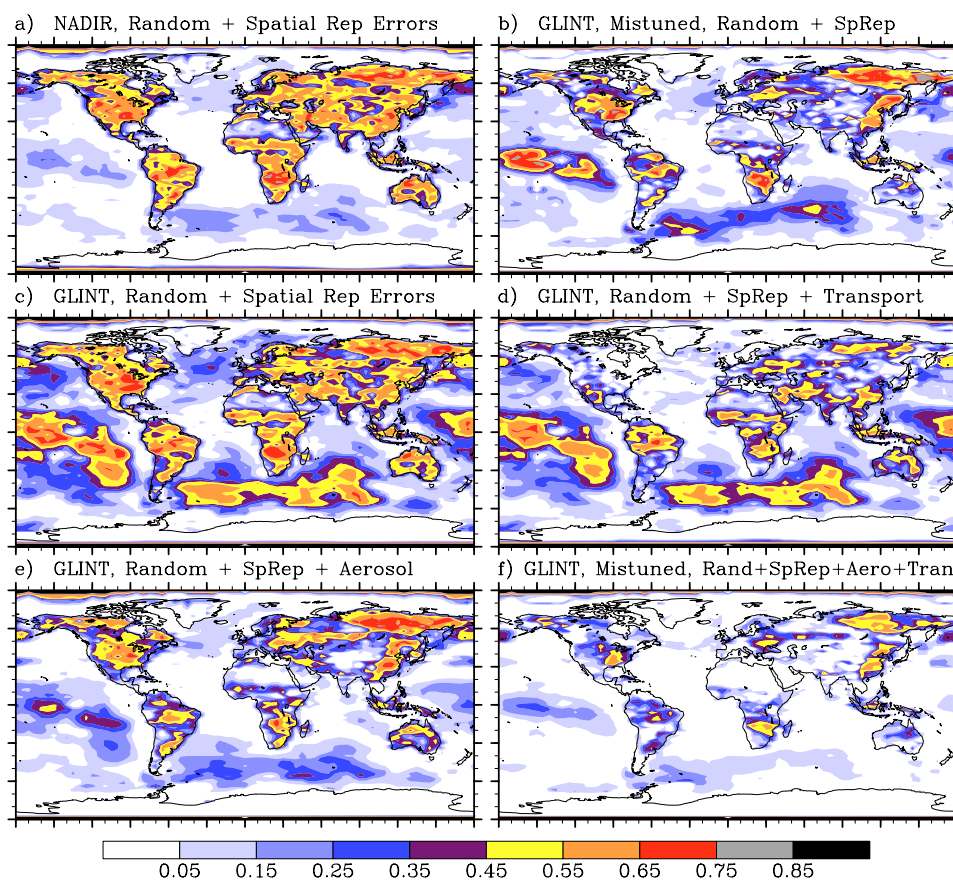
Fig. 8. Fractional error reductions in 7-day CO_2 fluxes, using full-year RMS errors, after 50 iterations of the optimization method, for experiments: **(a)** #1, control case (random measurement errors and spatial representation errors only) using nadir data; **(c)** #2, control case using glint data, **(b)** #3, mistuned case, glint, **(d)** #4, random errors + aerosol biases, glint, **(e)** #5, random + transport errors, glint, and **(f)** #6, random errors + aerosol biases + transport errors + mistuning effects, glint.

these errors by shifting the meteorology products driving the transport model forward by 18 h when generating the truth as compared to those used in the assimilation. This captures errors in both the synoptic meteorology as well as in the timing of the diurnal cycle of mixing. At the same time,

we add the transport uncertainties in Fig. 6f to the assumed measurement uncertainties to account for the transport errors; these are taken as the mean of the absolute values of the true and prior fluxes (Fig. 2a, b), divided by a factor of 10^{-7} kg CO_2 m^{-2} s^{-1} ppm^{-1}. This ad hoc estimate is based

on the idea that the largest transport errors occur where the surface flux variability is the greatest, and that this occurs where the fluxes themselves are the greatest.

Finally, Experiment 6 examines the combined effect of all three systematic error sources: the mistuning effects, aerosol biases, and transport errors of Experiments 3, 4, and 5, respectively.

3 Results

We use the root mean square (RMS) difference, RMS_{post}, between the estimated and true fluxes to assess the assimilation results. This is presented here in terms of the fractional error reduction statistic, given by $(RMS_{prior}-RMS_{post})/RMS_{prior}$, which puts areas of large and small flux variability on more equal footing (the RMS values themselves are given in the Supplemental Material http://www.atmos-chem-phys. net/10/4145/2010/acp-10-4145-2010-supplement.pdf). RMS_{prior} quantifies the initial difference between the prior and true flux models; no attempt to incorporate the information provided by the current in situ measurement network into RMS_{prior} has been made, since Baker et al. (2006b) suggest that its constraint is weak at the $2° \times 5°$ resolution examined here. The RMS values for the estimated 7-day fluxes given here are computed across the full year (see Supplemental Material for a seasonal breakdown http://www.atmos-chem-phys.net/10/4145/2010/ acp-10-4145-2010-supplement.pdf); RMS statistics for seasonal means computed from the 7-day fluxes are also given.

3.1 Control experiments

A posteriori RMS 7-day flux error reductions obtained using data from nadir- and glint-mode OCO observations (Experiments 1 and 2) after 50 descent iterations of the assimilation algorithm are presented in Fig. 8a, c. The nadir observations provide little improvement over the oceans (in agreement with the very high measurement errors there) but impressive improvements over the land – 45% or more in most areas, especially where the initial flux errors (Fig. 2d) are largest. The glint mode improvement over land is nearly as good as that of nadir mode – surprisingly, given that the effective glint mode measurement uncertainties are larger over land than the nadir ones (Fig. 6b). Apparently, enough land flux information blows out over the ocean for the more precise glint mode measurements there to compensate for the less precise and/or less available glint mode measurements over the adjacent land regions. As might be expected, the ocean flux improvement in glint mode is much better than in nadir; in fractional terms, it is as large as the improvement over the land, over 45%, in the areas where the initial errors are the largest. Since glint mode measurements give lower flux errors over a broader area than nadir mode (i.e,

over both land and ocean), we focus on glint mode in the remaining experiments.

Improvements are less impressive in the areas with low initial flux errors – the background flux estimation error due to the measurement noise masks improvements there. The assimilation corrects the largest flux errors during initial descent steps of the optimization, moving to progressively finer-scale corrections later. While lack of improvement in the low-flux areas could thus also be due to not running out the optimization method for enough iterations, we have been careful to converge adequately and feel that this is not the case here.

In Fig. 9, we plot the seasonal flux error reductions (computed from the RMS of four 13-week values) corresponding to the 7-day flux error reductions given in Fig. 8. For the control experiments (Fig. 9a and c), the initial errors are reduced by over 65% almost everywhere over land, as compared to only over 45% for the 7-day fluxes for similar areas. In glint mode, the ocean improvements are also greater.

The a posteriori error statistics given by these control experiments correspond to those from a single draw from the a posteriori estimation error covariance matrix, if our method were to compute one. While they do not include systematic errors, they provide a useful "best case" error estimate – if the measurements are not precise enough to provide useful information in this view, they will never be when all the other systematic error sources are added in. We address these other errors next.

3.2 Estimation errors with a "mistuned" assimilation

When the measurement noise and a priori flux error covariance matrices assumed in the assimilation ($\mathbf{R}_a$ and $\mathbf{P}_{o,a}$) are not equal to those corresponding to the true measurement noise added ($\mathbf{R}_t$) and the true statistics of the prior-truth flux fields ($\mathbf{P}_{o,t}$), then we call the assimilation "mistuned". For a basic Bayesian cost function $J = (\mathbf{H}x - z)^T \mathbf{R}_a^{-1}(\mathbf{H}x - z) + (x - x_o)^T \mathbf{P}_{o,a}^{-1}(x - x_o)$, where x and x_o represent the estimated and a priori state vector, z the measurements, and $\mathbf{H}$ the linearized measurement matrix, the true a posteriori covariance matrix in that case is given by

$$\mathbf{P}_x = [\mathbf{H}^T \mathbf{R}_a^{-1} \mathbf{H} + \mathbf{P}_{o,a}^{-1}]^{-1} \left[\mathbf{H}^T \mathbf{R}_a^{-1} \mathbf{R}_t \mathbf{R}_a^{-1} \mathbf{H} + \mathbf{P}_{o,a}^{-1} \mathbf{P}_{o,t} \mathbf{P}_{o,a}^{-1} \right]$$
$$[\mathbf{H}^T \mathbf{R}_a^{-1} \mathbf{H} + \mathbf{P}_{o,a}^{-1}]^{-1} \tag{4}$$

and no longer reduces to the simplified form $\mathbf{P}_x = [\mathbf{H}^T \mathbf{R}_a^{-1} \mathbf{H} + \mathbf{P}_{o,a}^{-1}]^{-1} = \mathbf{P}_{x,a}$. To produce a posteriori error statistics corresponding to what would be given by a full-rank covariance matrix with our simulation setup (in the control experiments, 1 and 2), we had to set $\mathbf{R}_t = \mathbf{R}_a$ by adding measurement noise to the data using the statistics from $\mathbf{R}_a$, and we chose $\mathbf{P}_{o,a}$ to agree with the actual (known) prior-truth flux difference. However, in a real-world assimilation, one has only an imprecise idea of what $\mathbf{R}_t$ and $\mathbf{P}_{o,t}$ should be, so
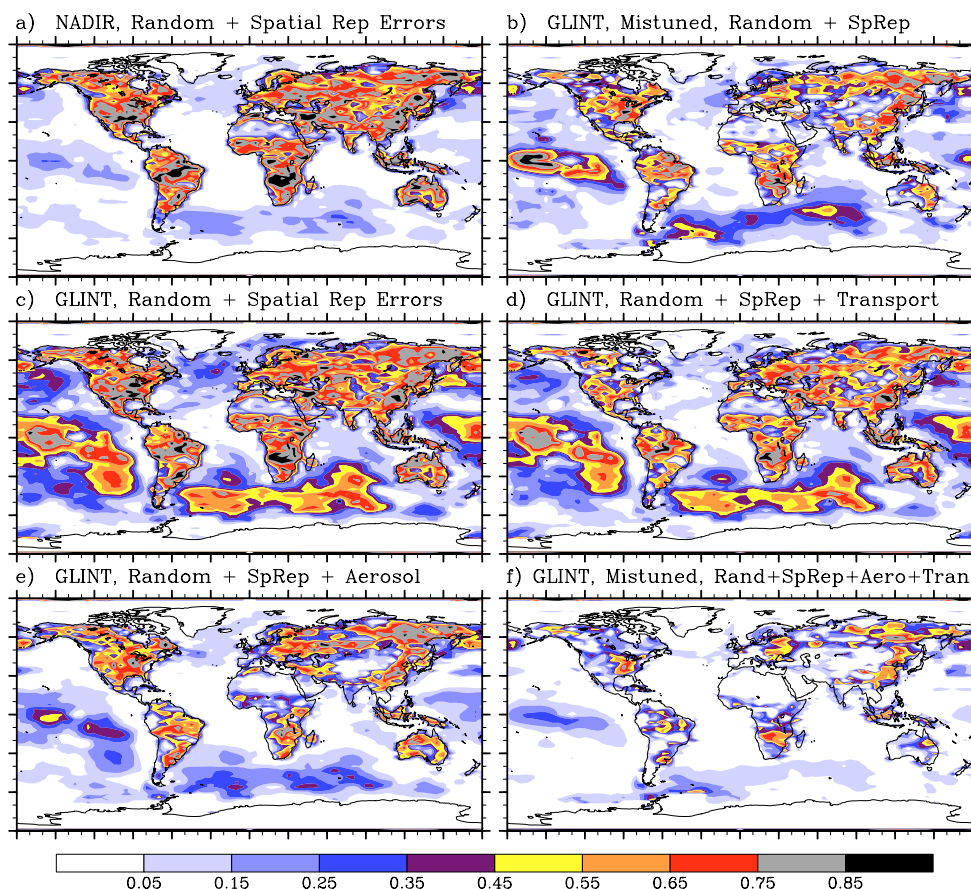
Fig. 9. Fractional error reductions in seasonal CO$_2$ fluxes, computed from the RMS of the four seasonal values (JFM, AMJ, JAS, OND) for experiments: **(a)** #1, **(b)** #3, **(c)** #2, **(d)** #4, **(e)** #5, and **(f)** #6, as in Fig. 8.

$\mathbf{R}_a \neq \mathbf{R}_t$ and $\mathbf{P}_{o.a} \neq \mathbf{P}_{o,t}$ and the covariance from Eq. (4) applies; this is captured in our error statistics when we mistune $\mathbf{P}_{o,a}$ and $\mathbf{R}_a$.

Mistuning both $\mathbf{P}_{o,a}$ and $\mathbf{R}_a$ (Experiment 3) degrades the flux estimate over most of the globe (compare Fig. 8b to c), especially in areas with lower initial flux differences. Areas in the center of broad regions of initially-large flux errors are affected the least by the mistuning. We have done a separate assimilation, not shown here, that verifies that most of this degradation is due to the mistuning of $\mathbf{P}_{o,a}$, rather than $\mathbf{R}_a$.

3.3 The impact of aerosol-related measurement biases

Adding a bias proportional to aerosol depth (Experiment 4) causes a significant degradation in the assimilated 7-day fluxes over land (compare Fig. 8e to c), most noticeably around the edges of the continents and around the high aerosol regions of Africa, western Asia, and India. Over the oceans, the impact is even larger, degrading the improvement by almost a factor of two in many places. The impact of the biases is at least this important for the seasonal fluxes (Fig. 9), but even so, there are still large areas over

land where improvements over 65% remain, particularly in the interior of the continents.

3.4 Impact of transport errors

The 18 hour shift in winds added in Experiment 5 greatly degrades the estimated fluxes over the extra-tropics (compare Fig. 8d to c), especially over North America and east Asia where the jets are the strongest, and has a somewhat lesser impact in the tropics. The near-surface winds in the extra-tropics are predominantly horizontal, so transport errors there lead to horizontal errors in where the flux corrections are placed. Over the tropics, however, wind motions are more vertical, due to the weak Coriolis force and the dominance of convection; transport errors affect more where concentrations are distributed in the column (having little impact on the column-integrated measurement) and less the horizontal assignment of the fluxes. Interestingly, the degradation in the estimates is weakest over the extratropical southern oceans, where horizontal winds are strong: transport here may be more predictable, or else the lower flux variability here may account for the difference.

The degradation of the 7-day flux estimates due to transport error is much less than that due to the aerosol biases over the oceans, and greater over the northern land (compare Fig. 8d and e). The impact on the seasonal flux error reductions (Fig. 9), however, is different: the transport errors generally have a smaller impact than the aerosol bias errors everywhere, except over North America, where they are similar. Unlike the aerosol biases applied here, which vary slowly across the year, the transport errors are more variable and their effect on the inverted fluxes cancels out more when averaged over longer spans.

3.5 Impact of all three systematic error sources

When the effects of all three systematic error sources (mistuning, transport error, and measurement biases) are considered together (Experiment 6), most of the flux improvements are lost. In terms of the weekly flux error reductions (Fig. 8f), there are still areas over land with improvements of 45% or higher, though these are restricted geographically to some of the areas with the largest initial errors, or to broad regions of homogeneous flux (eastern Siberia). Error reductions over the oceans are less encouraging, under 15% for most areas. Improvements in the seasonal fluxes (Fig. 9f) are 10–20% higher over the land than for the weekly fluxes but just as restricted geographically, and are similarly low over the ocean.

3.6 Impact of systematic errors at coarser scales

For climate research, flux averages over annual scales (or longer) are of more interest than the weekly and seasonal fluxes discussed above. The annual mean fractional error reductions we obtain are noisy – we simulated only a single year of data here, so random errors do not cancel out – but they tend to be at least as large the seasonal error reductions in Fig. 9. This suggests that the more-statistically-significant fractional reductions we obtain for the seasonal flux errors (Fig. 9) may be a good proxy for the annual mean error reductions across the full globe. It was not clear that this would be the case before doing these tests: the magnitude of the a priori errors in the seasonal fluxes is generally higher than in the annual means, especially over land, and since these magnitudes are in the denominator of the error reductions, one might think that the seasonal error reductions would be higher.

The seasonal errors from the control experiments (see Supplemental Material; http://www.atmos-chem-phys.net/10/4145/2010/acp-10-4145-2010-supplement.pdf) are characterized by alternating regions of counterbalancing errors over the global land areas, on scales of ~1000–2000 km. The ocean errors vary across longer scales but are weaker. For the experiments with systematic errors added, the errors grow and take on coarser scale patterns over the land regions. Much of the alternating ± errors over land cancel out when integrated over larger regions. In Fig. 10, we integrate the

seasonal and annual mean flux errors across the 22 globe-spanning regions from the Transcom3 (T3) flux inversion intercomparison project (see Fig. 1 from Baker et al., 2006a for a map). The RMS seasonal errors (plotted below the axis as negative values) for the 11 land regions drop from a priori values of ~0.5–2.0 PgC/year to ~0.1–0.2 PgC/year for the control experiments. When the systematic errors in the problem are added on, however, these land errors increase to ~0.3–0.6 PgC/year, still low enough to give a significant improvement over the a priori estimates, but much worse than the control experiment statistics would indicate. For the annual mean errors (absolute values plotted above the axis) over land, a priori errors in the range of ~0.1–0.5 PgC/year are reduced to generally below 0.1 PgC/year in the control experiments, but rise back up to ~0.1–0.3 PgC/yr when the systematic errors are considered. For those T3 regions with the largest initial errors, the errors are halved at least, while those with the smallest initial errors see little to no improvement. Over the oceans, where the seasonal cycles are less pronounced, error reductions of up to 50% are obtained for both seasonal and annual mean errors in the control experiment with glint mode data, but little improvement is obtained when the systematic errors are also considered.

4 Summary and discussion

We have simulated how well X_{CO_2} measurements from the OCO satellite could constrain the surface sources and sinks of CO_2, using a variational data assimilation technique that treats the measurements at the time and place they occur, averaged only over the time step and grid resolution of the transport model. The fluxes are solved at a coarser time resolution – weekly – to get adequate measurement density at our $2° \times 5°$ spatial resolution. We have used improved measurement information: new estimates of single-retrieval error uncertainties and averaging kernels calculated as a function of surface type, aerosol OD, and viewing geometry. And we combine the information from all valid retrievals for each ~33 second grid box crossing to get the measurement uncertainty used in the assimilation, accounting for measurement correlations as well as data dropout from both clouds and aerosol.

We first computed best case flux error estimates in our control experiments using X_{CO_2} measurements affected only by random errors. These error statistics correspond to those that would be given by a full-rank a posteriori covariance matrix, were one to be calculated. Nadir- and glint-mode measurements give similar flux improvements over the land: generally over 45/65% for weekly/seasonal fluxes. The weekly flux error reductions are larger than those obtained by Chevallier et al. (2007a) by almost a factor of two, despite the fluxes being solved for at a similar resolution: this is to be expected, since our measurement uncertainties (Fig. 6b) are several times lower than the 2 ppm values they assumed.
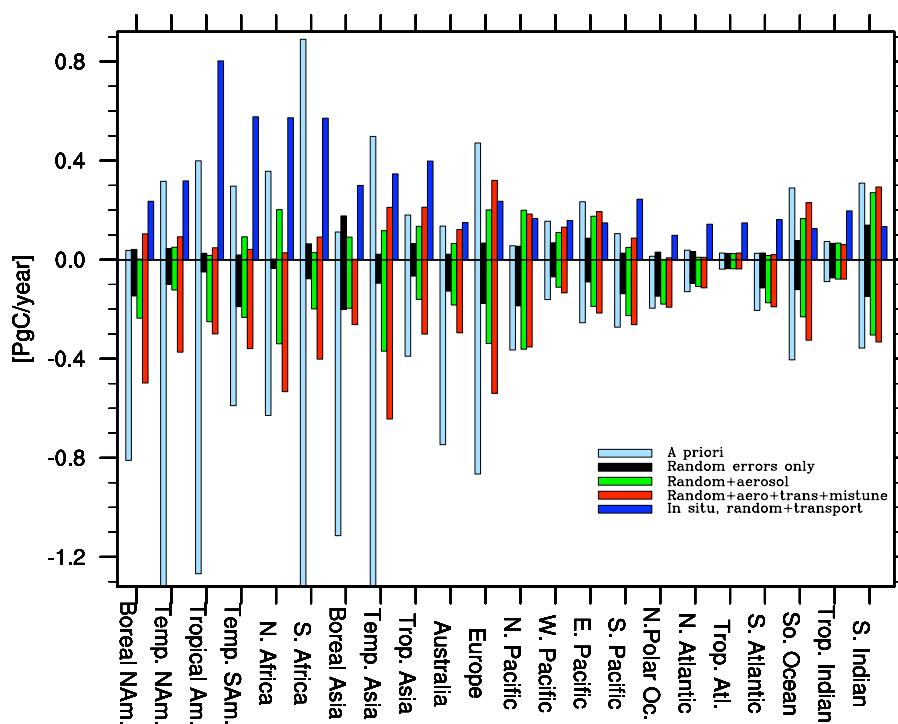
Fig. 10. Annual mean flux errors and RMS seasonal flux errors [PgC/year] integrated over the areas of the 22 Transcom3 emission regions. The absolute values of the annual mean errors are plotted above the axis as positive values, while the RMS of four 13-week seasonal values are plotted below it as negative values. A posteriori errors from three glint mode experiments are given: #2 (black bars), in which only random measurement errors are added, #4 (green) in which aerosol biases are also added, and #6 (red) in which random errors, aerosol bias, and transport errors are all added, as well as mistuning effects. Also given: the a priori flux errors (light blue) and the a posteriori errors given by assimilating only data from the in situ CO$_2$ montoring network of the 1990s (dark blue), computed as the root sum square of the "Post. Error" and "Model Error" columns from Table 4 of the Transcom3 CO$_2$ flux interannual variability study (Baker et al., 2006a).

Also, we do not solve for both day and night fluxes for each span as they do, resulting in fewer degrees of freedom and a somewhat tighter flux constraint. It is more difficult to compare our results with those of Miller et al. (2007) because they both used higher measurement uncertainties (1 ppm) and solved for larger flux regions (effectively adding strong spatial correlations): our flux uncertainties are larger over the land (except over Australia where they use smaller regions) and smaller over the oceans (in both nadir and glint modes). Our results, like those of Baker et al., 2006b and Miller et al., 2007, indicate that the OCO data should provide a much better constraint on the CO$_2$ fluxes than the current in situ network, in this random-errors-only view. On the scale of the 22 global Transcom3 regions, our seasonal error reductions are generally similar to the 32-day values of Feng et al. (2009); like them, we see a tendency towards lower improvements at high latitudes in the winter hemisphere, when few glint-mode measurements are available.

In our simulations, glint mode data give land flux error reductions that are nearly as great as with nadir data, despite the larger glint measurement uncertainties over land, apparently because the more precise glint measurements over the

ocean contain much information on the land fluxes, enough to make up the difference. Feng et al. (2009) found a similar compensation, using an entirely different approach for assessing data availability and aggregated measurement error. The difference between glint and nadir results over land is more noticeable here than in Feng et al., however, perhaps because we decrease the probability of finding clear and low-aerosol scenes at high SZAs (using the factor in Eq. 3) more than they do. Over the oceans, the more precise glint measurements lead to much larger flux error reductions than the nadir data: over 45% across broad swaths of the tropical and southern oceans, versus under 15% in nadir. Because the glint data provide more of an overall constraint on the surface fluxes (both land and ocean), in this random-errors-only view OCO would collect more information on the global carbon cycle overall by remaining in glint mode at all times rather than by switching between glint and nadir modes (but see discussion below).

While the control experiment error analyses provide a useful metric for comparing different sets of observations, they provide an overly-optimistic view of how well the OCO data actually will improve our flux estimates. On one hand, the

actual random retrieval errors are likely to be higher than those assumed here, since the analysis of Bösch et al., does not capture all possible radiative transfer errors (e.g. those due to the vertical distribution, size, and shape of scatterers, the absorption line shape, line mixing, etc.). Probably more importantly, though, a variety of systematic errors will prevent the improvement from being this large. It is difficult to know beforehand which systematic errors will be most important for a mission; the crude representations added here give only a rough idea of what may actually occur.

First of all, we found that mistuning the assimilation (assuming incorrect patterns for the a priori flux error covariance and measurement error covariance matrices) by a realistic amount degrades the error reductions significantly, especially in areas where the initial flux differences are lower. This error source is unavoidable: the assimilation must be constrained by a realistic prior to damp out the worst effects of the random measurement errors (Baker et al., 2006b), and yet there is little chance of modeling the details of the a priori uncertainties correctly to avoid the mistuning (Chevallier et al., 2006); the same modeling challenges apply to the assumed measurement error covariance, as well.

Second, we added measurement biases proportional to aerosol OD, since aerosol-related radiative transfer modeling errors are expected to be an important source of model-measurement mismatches. With these biases added, the flux error reductions over the oceans are degraded by about a factor of two compared with the unbiased values; over land, flux improvements as high as in the unbiased case are still often achieved, but the spatial extent of such improvements are degraded by about a factor of two. Weekly flux error reductions as high as 65% are still achieved in a few areas, especially eastern Siberia. We obtain aerosol-related annual mean flux biases on the scale of the 22 Transcom3 regions that are generally smaller than Chevallier et al. (2007a) obtain: they are never greater than 0.2 PgC/year (look at the difference between the green and black bars on the top of Fig. 10). The two largest biases from Chevallier et al. (0.73 and 0.57 PgC/year for Temp. Eurasia and Europe, respectively; see their Fig. 4) seem to be due to the use of aerosol biases as high as 1.0 ppm or higher over those regions; the largest biases we applied were only 0.6 ppm (this, too, is likely to be over-optimistic).

Finally, we examined the impact of transport model errors in the assimilation with the ad hoc approach of shifting the winds used to generate the truth by 18 h. These degraded the 7-day flux improvements more strongly over land than the aerosol bias experiment, especially in the extra-tropical north, but had a much smaller impact over the oceans. The impact on the seasonal flux error reductions was much less: apparently, the transport errors that we added largely average out in time, something that may not occur with more realistic transport errors.

When all three systematic error sources (mistuning, transport, and aerosol biases) are added at the same time, most of the improvement seen in the control experiments is lost: the OCO data improve the weekly flux estimates by more than 45% in only a few restricted areas over the land (roughly corresponding to those areas where our a priori uncertainty is the largest) and generally under 15% over the oceans.

Our simulations suggest that the precision of OCO's X_{CO_2} measurements is more than adequate for estimating weekly grid-scale CO_2 fluxes at scientifically-useful levels. Knowing annual mean CO_2 fluxes to within 0.1 PgC/yr for most of the 22 Transcom3 regions (Fig. 10) would constrain the key sources and sinks of CO_2 well on a global scale. The real challenge, however, appears to be in identifying and removing systematic errors, both in deriving the X_{CO_2} values and in processing these values with an atmospheric assimilation method. For the level of systematic errors considered here, annual mean flux errors rise as high as 0.2–0.3 PgC/year for many of the Transcom3 regions, a level which, while better than that given by the current in situ network, still would leave much uncertainty in the global carbon budget. Since the value of the X_{CO_2} data fall off rapidly if systematic errors are much higher than this, more effort must devoted to quantifying them. We have addressed the systematic errors only in a very rough fashion here. The OCO X_{CO_2} retrievals will likely be corrupted by a variety of measurement error sources, spectrographic and radiative transfer modeling errors, and other errors besides the aerosol scattering effects considered approximately here. Simulation studies might be able to help characterize the impact of these error sources, once they are identified. These are not simply of academic interest, to be forgotten once the spacecraft begins returning real data; rather, they will be critical for interpreting the data once it arrives. A more detailed assessment of transport errors must also be performed. The transport errors could be quantified by running the identical fluxes (including fossil fuel input at fine spatial scales and diurnally-varying land biospheric fluxes) through multiple transport models, sampling the resulting concentration fields with realistic averaging kernels along realistic OCO orbits, and then comparing the resulting X_{CO_2} values in an approach similar to what the Transcom group has done for continuous in situ and aircraft profile data (Law et al., 2008; Patra et al., 2008; Pickett-Heaps et al., 2010) and is currently doing for satellite measurements (S. Maksyutov, lead). Finally, our mistuning experiment illustrates the importance of having a good a priori flux model to help partition the flux corrections properly: we must continue to improve our flux process models, just as we must improve our transport models.

If the systematic errors in the problem can be beaten down to below the levels used here, then the OCO measurements should provide much useful new carbon cycle science. Improvements in seasonal fluxes of ~50% or more over the tropical and northern forests, when viewed over the course of multiple years, will begin to resolve the processes driving the global interannual variability of CO_2. Similar improvements in weekly fluxes will help clarify the response of ecosystems
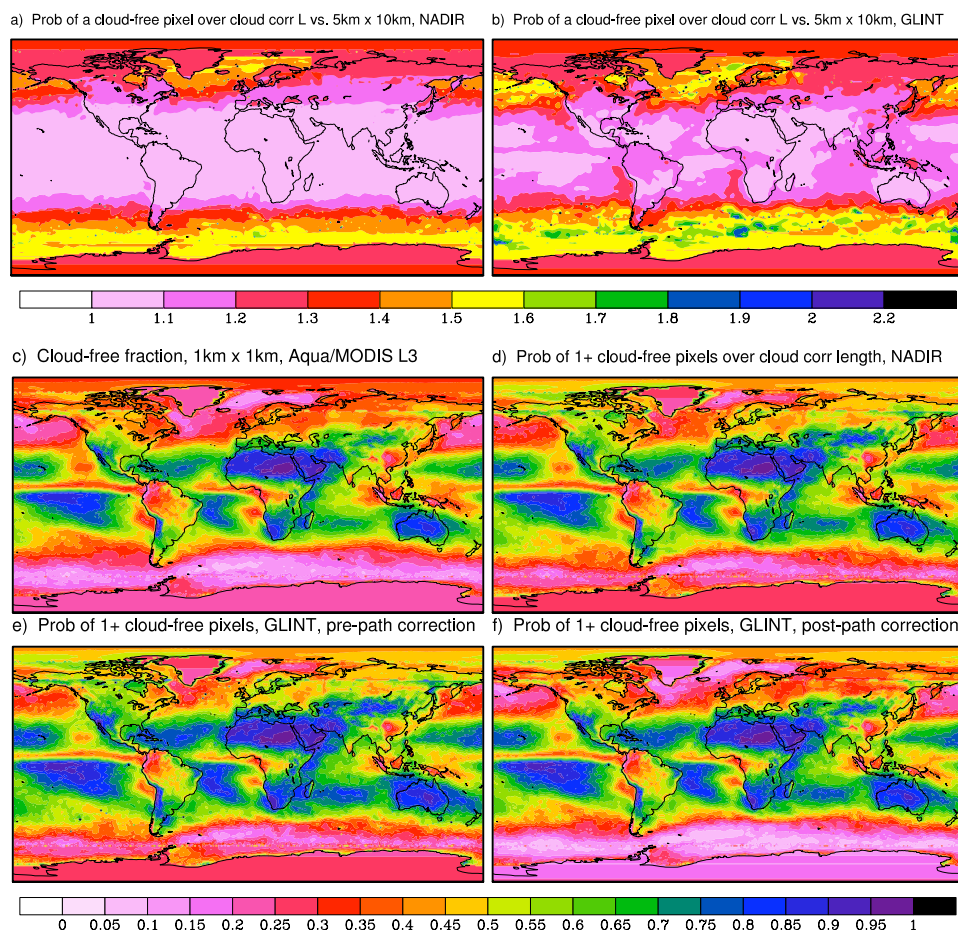
Fig. A1. Computation of climatological cloud-free pixel availability from Terra/MODIS and Aqua/MODIS data. The ratio of the probability of finding at least one cloud-free sounding across a ground track swath of length L (Fig. 5b) over the same probability for a swath only 5 km long, calculated by sampling 10 km-wide Terra/MODIS Level 2 data swaths in the along-track direction, using L for a) nadir- and **(b)** glint. **(c)** The cloud-free probability at $1 km \times 1 km$ resolution, taken from the Aqua/MODIS Level 3 cloud-mask product. **(d)** and **(e)**: the probability of finding at least one cloud free sounding in an OCO ground track swath of length L (nadir and glint) found by multiplying (c) by (a) and (b). **(f)** The glint-mode cloud-free probability from (e) corrected for the greater atmospheric path length at high SZAs according to Eq. (3). Note that the probabilities in (e) are higher than in (d) because L is about two times longer in glint than nadir (Fig. 5b); because they are divided by L in Eq. (2), however, the resulting N_{eff} values are lower in glint than nadir, even without the glint path correction.

to fast disturbances (like fire) and variability in the weather-related drivers. Improvements over the ocean may be as great as over land, depending on the nature of the aerosol biases, especially. Perhaps the greatest impact will come where our current observations are the worst, such as over the tropical forests, which are thought to play in driving global CO_2 variability (Baker et al., 2006a). Further, the global distribution of the improvements should help clarify the partitioning of the global sink between the tropics and extra-tropics, and help pin down the longitudinal distribution of the northern CO_2 sink.

Appendix A

Cloud coverage calculations

Figure 7b shows the probability of finding at least one cloud-free scene inside an OCO ground track swath of length L, the cloud influence length defined by Eq. (1) and plotted in Fig. 5b. This appendix describes how climatological values for this probability are derived from MODIS observations.

The cloud fraction parameter from the Aqua/MODIS Level 3 MYD08_M3 product gives the monthly average probability that any single MODIS scene will be clear of detectable clouds, at $1° \times 1°$ resolution. This is computed from data at $1 km \times 1 km$ resolution, close to the OCO FOV size, and is available for the same orbit as OCO. This should

give a very good idea of the probability that any single OCO sounding will see cloud-free conditions. Because of along-track spatio-temporal correlations, however, it is not clear how to compute the probability of finding at least one cloud-free scene in an OCO ground track swath of length L from these single-sounding probabilities. Obtaining that information requires examining the Level 2 MODIS data from which the Level 3 monthly averages were computed.

The Level 2 MODIS data come packaged in the form of "granules", approximately 5 min of measurements spanning roughly 2000 km in the along-track direction and 2330 km across-track (as swept out by a ±55° scan on either side of nadir). Rather than process this massive archive of data ourselves, we used a "climatology" of Level 2 MODIS cloud and cloud mask products (MOD06_L2 and MOD35_L2) that was compiled by Chang and Li (2005), albeit from the Terra satellite which has a somewhat different orbit than Aqua and OCO. To reduce the volume of data to process, Chang and Li processed 8 full days of data in each of the months of January, April, July, and October, spaced 4 days apart from each other. Among other cloud-related quantities, they saved a cloud mask value at 5 km×5 km resolution indicating whether the scene was "cloudy", "possibly cloudy", "probably clear", or "confident clear". For the "cloudy" boxes, an additional value was saved indicating the number of 1 km×1 km pixels inside the 5 km×5 km box (0–25) with measurable cloud optical depths (MODIS can generally detect clouds with ODs as thin as 0.10). This second quantity is valuable because it provides the frequency of occasional cloud gaps in areas with the cloudiest conditions, where OCO will have the most difficulty obtaining data, at a 1 km×1 km resolution that is close to that seen by OCO (nominally 2.8 km^2 when the sensor slit is oriented perpendicular to the direction of motion, less than that when the satellite "pirouettes" towards the sun to maintain its pointing in the sun/ground/satellite plane; overall, the average FOV size is ∼2 km^2).

We have sampled the Chang and Li data in 10 km-wide swaths of differing lengths (5, 10, 20, 40, 100, and 200 km) in the along-track direction, accumulating statistics on the probability of finding at least one cloud-free scene at 1 km×1 km resolution inside the swaths of differing lengths for each month. The probabilities increase with increasing swath length. We normalize the probabilities at each swath length by those at the 5 km length. This normalized multiple represents how many more time likely it is to find at least one 1 km×1 km cloud-free scene in a swath of length L than it is inside a box of 5 km×10 km, accounting for realistic correlations in cloud amount along the track, or cloud "clumpiness". Figure A1a, b gives maps of this multiple interpolated to the actual swath lengths for nadir and glint modes (Fig. 5b) corresponding to the true solar zenith angles around the orbit. At high solar zenith angles, including the near-polar areas where it will be the most difficult to penetrate through the clouds, this multiple is generally over 2.

In the final step of this process, we interpolate these multiples across the full year from the four months examined by Chang and Li, and multiply them by the single-sounding cloud-free probabilities of the Level 3 Aqua/MODIS product (Fig. A1c) to obtain the probability of a cloud-free sounding per cloud correlation length L shown in Fig. 5b for the nadir case. For glint mode, these cloud-free probabilities are further reduced to account for the greater path-length in the atmosphere according to Eq. (3).

Our approach here is actually somewhat conservative, since the probability of finding a cloud-free sounding inside a box of 5 km×10 km (the value we normalize our Level 2 multiple by) should be higher than the single-sounding cloud-free probability. Another factor to consider is that our Level 2 MODIS multiples are computed using data from the Terra satellite, which has a 10:30 a.m. local ascending node time and thus may not exactly capture the cloud properties that OCO will see in the early afternoon.

References

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, Global Biogeochem. Cy., 20, GB1002, doi:10.1029/2004GB002439, 2006a.

Baker, D. F., Doney. S. C., and Schimel, D. S.: Variational data assimilation for atmospheric CO_2, Tellus B, 58(5), 359–365, doi:10.1111/j.1600-0889.2006.00218.x, 2006b.

Bloom, S., da Silva, A., Dee, D., Bosilovich, M., Chern, J.-D., Pawson, S., Schubert, S., Sienkiewicz, M., Stajner, I., Tan, W.-W., and Wu, M.-L.: Documentation and Validation of the Goddard Earth Observing System (GEOS) Data Assimilation System – Version 4. NASA Technical Report Series on Global Modeling and Data Assimilation 104606, v26, 2005.

Bösch, H., Baker, D., Connor, B., Crisp, D. and Miller, C.: Global Characterization of X_{CO_2} Retrievals from Shortwave-infrared Satellite Observations, in preparation, 2010.

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

Chang, F.-L. and Li, Z. Q.: A near-global climatology of single-layer and overlapped clouds and their optical properties retrieved from Terra/MODIS data using a new algorithm, J. Climate, 18(22), 4752–4771, 2005.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, 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.-Atmos., 110(D24), D24309, doi:10.1029/2005JD006390, 2005a.

Chevallier, F., Engelen, R. J., and Peylin, P.: The contribution of AIRS data to the estimation of CO_2 sources and sinks, Geophys. Res. Lett., 32(23), L23801, doi:10.1029/2005GL024229, 2005b.

Chevallier, F., Viovy, N., Reichstein, M., and Ciais, P.: On the assignment of prior errors in Bayesian inversions of CO_2 surface fluxes, Geophys. Res. Lett., 33(13), L13802, doi:10.1029/2006GL026496, 2006.

Chevallier, F., Breon, 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.-Atmos., 112(D9), D09307, doi:10.1029/2006JD007375, 2007a.

Chevallier, F.: Impact of correlated observation errors on inverted CO_2 surface fluxes from OCO measurements, Geophys. Res. Lett., 34(24), L24804, doi:10.1029/2007GL030463, 2007b.

Chevallier, F., Maksyutov, S., Bousquet, P., Breon, F.-M., Saito, R., Yoshida, Y., and Yokota, T.: On the accuracy of the CO_2 surface fluxes to be estimated from the GOSAT observations, Geophys. Res. Lett., 36(19), L19807, doi:10.1029/2009GL040108, 2009.

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.-Atmos., 113, D05305, doi:10.1029/2006JD008336, 2008.

Corbin, K. D, Denning, A. S., Lu, L., Wang, J. W., and Baker, I.: Possible representation errors in inversions of satellite CO_2 retrievals, J. Geophys. Res.-Atmos., 113(D2), D02301, doi:10.1029/2007JD008716, 2008.

Crisp, D., Atlas, R. M., Breon, F.-M., Brown, L. R., Burrows, J. P., Ciais, P., Connor, B. J., Doney, S. C., Fung, I. Y., Jacob, D. J., Miller, C. E., O'Brien, D., Pawson, S., Randerson, J. T., Rayner, P., Salawitch, R. J., Sander, S. P., Sen, B., Stephens, G. L., Tans, P. P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) Mission, Adv. Space Res., 34, 700–709, 2004.

Doney, S. C., Lindsay, K., Fung, I., and John, J.: Natural variability in a stable 1000 year coupled climate-carbon cycle simulation, J. Climate, 19(13), 3033–3054, 2006.

Enting, I. G., Trudinger, C. M., Francey, R. J., and Granek, H.: A synthesis inversion of the concentration and $\delta^{13}C$ of atmospheric CO_2, Tellus B, 47, 35–52, 1995.

Feng, L., Palmer, P. I., Bösch, H., and Dance, S.: Estimating surface CO_2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble Kalman Filter, Atmos. Chem. Phys., 9, 2619–2633, 2009,
http://www.atmos-chem-phys.net/9/2619/2009/.

Houweling, S., Breon, F.-M., Aben, I., Rödenbeck, C., Gloor, M., Heimann, M., and Ciais, P.: Inverse modeling of CO_2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time, Atmos. Chem. Phys., 4, 523–538, 2004,
http://www.atmos-chem-phys.net/4/523/2004/.

IPCC: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp., 2007.

Kawa, S. R., Erickson, D. J., Pawson, S., and Zhu, Z.: Global CO_2 transport simulations using meteorological data from the NASA data assimilation system, J. Geophys. Res.-Atmos., 109(D18), D18312, doi:10.1029/2004JD004554, 2004.

Law, R. M., Peters, W., Rodenbeck, C., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C., Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Patra, P. K., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Takigawa, M., Vautard, R., Vermeulen, A. T., and Zhu, Z: TransCom model simulations of hourly atmospheric CO_2: experimental overview and diurnal cycle results for 2002, Global Biogeochem. Cy., 22(3), GB3009, doi:10.1029/2007GB003050, 2008.

Lin, S.-J.: A vertically Lagrangian finite-volume dynamical core for global models, Mon. Weather Rev., 132, 2293–2307, 2004

Meirink, J. F., Bergamaschi, P., Frankenberg, C., d'Amelio, M. T. S., Dlugokencky, E. J., Gatti, L. V., Houweling, S., Miller, J. B., Röckmann, T., Villani, M. G., and Krol, M. C.: Four-dimensional Variational Data Assimilation for Inverse Modelling of Atmospheric Methane Emissions: 2. Analysis of SCIA-MACHY Observations, J. Geophys. Res.-Atmos., 113(D17), doi:10.1029/2007JD009740, 2008.

Miller, C. E., Crisp, D., DeCola, P. L., Olsen, S. C., Randerson, J. T., Michalak, A. M., Alkhaled, A., Rayner, P., Jacob, D. J., Suntharalingam, P., Jones, D. B. A., Denning, A. S., Nicholls, M. E., Doney, S. C., Pawson, S., Boesch, H., Connor, B. J., Fung, I. Y., O'Brien, D., Salawitch, R. J., Sander, S. P., Sen, B., Tans, P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., and Law, R. M.: Precision requirements for space-based X-CO_2 data, J. Geophys. Res.-Atmos., 112(D10), D10314, doi:10.1029/2006JD007659, 2007.

Najjar, R. G., Jin, X., Louanchi, F., Aumont. O., Caldeira, K., Doney, S. C., Dutay, J.-C., Follows, M., Gruber, N., Joos, F., Lindsay, K., Maier-Reimer, E., Matear, R. J., Matsumoto, K., Monfray, P., Mouchet, A., Orr, J. C., Plattner, G.-K., Sarmiento, J. L., Schlitzer, R., Slater, R. D., Weirig, M.-F., Yamanaka, Y., and Yool, Y.: Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from Phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2), Global Biogeochem. Cy., 21, GB3007, doi:10.1029/2006GB002857, 2007.

Olsen, S. C. and Randerson, J. T.: Differences between surface and column atmospheric CO_2 and implications for carbon cycle research, J. Geophys. Res.-Atmos., 109(D2), D02301, doi:10.1029/2003JD003968, 2004.

Patra, P. K., Maksyutov, S., Ishizawa, M., Nakazawa, T., Takahashi, T., and Ukita, J.: Interannual and decadal changes in the sea-air CO_2 flux from atmospheric CO_2 inverse modeling, Global Biogeochem. Cy., 19, GB4013, doi:10.1029/2004GB002257, 2005.

Patra, P. K., Law, R. M., Peters, W., Rodenbeck, C., Takigawa, M., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P., Brandt,

J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C., Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Vautard, R., Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO_2: analysis of synoptic scale variations for the period 2002–2003, Global Biogeochem. Cy., 22, 4, doi:10.1029/2007GB003081, 2008.

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.-Atmos., 110(D24), D24304, doi:10.1029/2005JD006157, 2005.

Peylin, P., Rayner, P. J., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P., and AEROCARB contributors: Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, Atmos. Chem. Phys., 5, 3173–3186, 2005a,
http://www.atmos-chem-phys.net/5/3173/2005/.

Peylin, P., Bousquet, P., Le Quéré, C., Sitch, S., Friedlingstein, P., McKinley, G., Gruber, N., Rayner, P., and Ciais, P.: Multiple constraints on regional CO_2 flux variations over land and oceans, Global Biogeochem. Cy., 19(1), GB1011, doi:10.1029/2003GB002214, 2005b.

Pickett-Heaps, C. A., Rayner, P. J., Law, R. M., Bousquet, P., Peylin, P., Patra, P., Maksyutov, S., Marshall, J., Roedenbeck, C., Ciais, P., Langenfelds, R., Tans, P., Steele, P., and Francey, R.: Atmospheric CO_2 inversion cross-validation using vertical profile measurements, J. Geophys. Res., submitted, 2010.

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, 535–560, 1997.

Rayner, P. J. and O'Brien, D.: The utility of remotely sensed CO_2 concentration data in surface source inversions, Geophys. Res. Lett., 28, 175–178, 2001.

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), Global Biogeochem. Cy., 19, GB2026, doi:10.1029/2004GB002254, 2005.

Rayner P. J., Law, R. M., Allison, C. E., Francey, R. J., Trudinger, C. M., and Pickett-Heaps, C.: Interannual variability of the global carbon cycle (1992–2005) inferred by inversion of atmospheric CO_2 and $\delta^{13}CO_2$ measurements, Global Biogeochem. Cy., 22, GB3008, doi:10.1029/2007GB003068, 2008.

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/.

Rödenbeck, C.: Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Max-Planck-Institut fr Biogeochemie: Technical Paper 6, 2005.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Glob. Change Biol., 9(2), 161–185, 2003.

Takahashi, T., Wanninkhof, R. H., Feely, R. A., Weiss, R. F., Chipman, D. W., Bates, N., Olafsson, J., Sabine, C., and Sutherland, S. C.: Net sea-air CO_2 flux over the global oceans: an improved estimate based on the sea-air pCO_2 difference, in: Proceedings of the 2nd International Symposium: CO_2 in the Oceans, the 12th Global Environmental Tsukuba, 18–22 January 1999, Tsukuba Center of Institutes, Y. Nojiri, ed., Natl. Inst. for Envir. Studies, Envir. Agency of Japan, 1999.

Zupanski, D., Denning, A. S., Uliasz, M., Zupanski, M., Schuh, A. E., Rayner, P. J., Peters, W., and Corbin, K. D.: Carbon flux bias estimation employing maximum likelihood ensemble filter (MLEF), J. Geophys. Res.-Atmos., 112(D17), D17017, doi:10.1029/2006JD008371, 2007.