Biogeosciences Discuss., 6, 10195–10241, 2009
www.biogeosciences-discuss.net/6/10195/2009/
© Author(s) 2009. This work is distributed under
the Creative Commons Attribution 3.0 License.

Biogeosciences Discussions

This discussion paper is/has been under review for the journal Biogeosciences (BG). Please refer to the corresponding final paper in BG if available.

A regional high-resolution carbon flux inversion of North America for 2004

A. E. Schuh[1], A. S. Denning[1], K. D. Corbin[1,*], I. T. Baker[1], M. Uliasz[1], N. Parazoo[1], A. E. Andrews[2], and D. E. J. Worthy[3]

[1]Colorado State University, Fort Collins, Colorado, USA
[2]National Oceanic and Atmospheric Administration Earth System Research Laboratory, 325 Broadway R/GMD1, Boulder, CO 80305, USA
[3]Environment Canada, 4905 Dufferin Street, Toronto, Ontario, Canada, M3H5T4, USA
[*]now at: CSIRO Marine and Atmospheric Research Aspendale, VIC, Australia

Received: 9 July 2009 – Accepted: 18 September 2009 – Published: 2 November 2009

Correspondence to: A. E. Schuh (aschuh@atmos.colostate.edu)

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

Abstract

Resolving the discrepancies between NEE estimates based upon (1) ground studies and (2) atmospheric inversion results, demands increasingly sophisticated techniques. In this paper we present a high-resolution inversion based upon a regional meteorology model (RAMS) and an underlying biosphere (SiB3) model, both running on an identical 40 km grid over most of North America. Previous papers have utilized inversion regions formed by collapsing biome-similar grid cells into large aggregated regions. The effect of this is that the NEE *correction* imposed on forested regions on the east coast of the United States might be the same as that imposed on forests on the west coast of the United States while, in reality, there likely exist subtle differences in the two areas, both natural and anthropogenic. Our current inversion framework utilizes a combination of previously employed inversion techniques while allowing carbon flux corrections to be biome independent. Temporally and spatially high-resolution results utilizing biome-independent corrections provide insight into carbon dynamics in North America. In particular, we analyze hourly CO_2 mixing ratio data from a sparse network of eight towers in North America for 2004. A prior estimate of carbon fluxes due to gross primary productivity (GPP) and ecosystem respiration (ER) is constructed from the SiB3 biosphere model on a 40 km grid. A combination of transport from the RAMS and the parameterized chemical transport model (PCTM) models is used to forge a connection between upwind biosphere fluxes and downwind observed CO_2 mixing ratio data. A Kalman filter procedure is used to estimate weekly corrections to biosphere fluxes based upon observed CO_2. RMSE-weighted annual NEE estimates, over an ensemble of potential inversion parameter sets, show a mean estimate 0.57 Pg/yr sink in North America. We perform the inversion with two independently derived boundary inflow conditions and calculate jackknife-based statistics to test the robustness of the model results. We then compare final results to estimates obtained from the CarbonTracker inversion system and the Ameriflux network. Results are promising, showing the ability to correct carbon fluxes from the biosphere models over annual and seasonal

time scales, as well as over the different GPP and ER components, and also providing interesting hypotheses for future work.

1 Introduction

Carbon dioxide inversion studies have generally been focused on improved estimation of terrestrial carbon fluxes such as ecosystem respiration (ER), gross primary production (GPP), and net ecosystem exchange (NEE) as a means to better understand the carbon cycle of the earth. Researchers have progressively increased the resolution, in both time and space, and accuracy of the carbon flux estimates over the past decade. Early inversion studies were focused primarily with finding an explanation for the missing sink of carbon that can be easily identified from calculating a budget from annual fossil fuel emissions to the atmosphere, the effect of land use changes, and the oceanic carbon sink and comparing it to annual records of increasing atmospheric carbon dioxide concentrations. Given that it often represents a third of the annual fossil fuel emissions, it is of great interest to scientists and policy makers alike. Inversion results have been very effective at identifying large defining features of the terrestrial portion of the carbon sink (Fan et al., 1998; Gurney et al., 2002) although much debate remains even at extremely large scales (Stephens et al., 2007). However, the debate on a global scale has not deterred researchers from focusing these techniques on finer scale problems. In fact, criticism has been aimed at large scale global inversions because of the fact that their estimates can be biased on finer regional scales (Kaminski et al., 2001). The data available for regional inversion studies is increasing rapidly year after year, primarily within the developed industrial nations of the Northern Hemisphere. This provides researchers with some of the first opportunities to perform inversion studies in a very high-resolution setting.

Gerbig et al. (2003) provided the first major regional inversion paper. They used a receptor-oriented inversion approach to investigate a series of flights from the CO_2 Budget and Rectification Airborne (COBRA) study conducted in 2000. Results showed

that the effect of biosphere carbon fluxes could be seen at altitude in mixed layer CO_2 observed by aircraft. The paper pointed out several areas for future improvements in regional inverse modeling including improving biosphere-atmosphere exchange and convective transport modeling. Peylin et al. (2005) followed this with a regional inversion based on western Europe in which he estimated daily fluxes for a month using relatively continuous measurements of CO_2 from towers in the inversion domain. The most similar effort made for North America comes from the ongoing CarbonTracker project (Peters et al., 2007). Peters et al. used a nested transport structure (TM5) with a relatively high-resolution 1-degree inner grid over North America. A priori carbon fluxes were estimated by modifying 1-degree by 1-degree monthly output from the Carnegie Ames Stanford Approach (CASA) model to provide diurnal variability by incorporating a Q_{10} temperature relationship for respiration and a linear scaling of GPP with solar radiance. NEE estimates were optimized by estimating linear correction factors for NEE for each of up to 19 ecoregion-based (Olsen et al., 1992) sub-areas of North America based upon a 5-week smoothing window. The coarseness of the inversion over North America is required in order to be able to solve biases simultaneously across the globe on the coarser nested grids.

Our inversion framework has drawn upon certain techniques from previous inversions while including some new features. The aim of the inversion is to provide fine scale inversion results over North America for 2004. A novel feature of this inversion is the distinct estimation of GPP and ER instead of just NEE, which to our knowledge has not previously been performed, at least in the regional framework. We have drawn upon the spatial correlation constraints used by Rödenbeck et al. (2003) and Michalak et al. (2004), largely in order to regularize the inversion problem. Large matrix inversions, required of nearly all inversion techniques, limited the inversion grid resolution to approximately 10 000 km^2 (60×36 grid composed of 100 km by 100 km grid cells). For sensitivity studies involving numerous inversion runs, a 40 000 km^2 grid (30×18 grid composed of 200 km by 200 km grid cells) is used. Most previous global inversions have been performed upon grid areas of around 5 to 10 times that size. In order

to provide some contrast, CarbonTracker optimizes 17 bias correction factors for NEE while this inversion typically optimizes 30×18=540 each (30×18 grid mentioned above) for ER and GPP. However, this does not come without a cost since we can't simultaneously optimize fluxes outside of North America. Therefore we used offline-derived boundary conditions and provided these as fixed contributions to the tower CO_2 budget.

Schuh et al. (2009) showed that considerable success could be achieved in estimating large spatial scale ER and GPP signals in the midst of small spatial scale variability in fluxes. We leveraged this result and put the problem in a Kalman filter framework in order to allow higher resolution spatial estimation. This filter is of a somewhat simple variety and allowed us to work with all portions of the inversion, such as complete prior and posterior covariance matrices, explicitly. We then tested sensitivity to a number of pieces of the inversion considered uncertain, including parameters in the actual inversion as well as fixed contributions to the modeled CO_2 such as fossil fuel and boundary inflow. As far as we know this is also the first paper providing a comparison of inversion results derived by using two independent boundary inflow estimates. Additionally, the effect of including recently available high-resolution fossil fuel inventory data is quantified.

2 Methods

2.1 Prior flux model and transport

The Simple Biosphere model (SiB) is based on a land-surface parameterization scheme originally used to compute biophysical exchanges in climate models (Sellers et al., 1986), but later adapted to include ecosystem metabolism (Sellers et al., 1996a; Denning et al., 1996a). SiB has been coupled to the Brazilian version of the Regional Atmospheric Modeling System (RAMS, Pielke et al., 1992; Frietas et al., 2006) and used to study PBL-scale interactions among carbon fluxes, turbulence, and CO_2 mixing

ratio (Denning et al., 2003) and regional-scale controls on CO_2 variations (Nicholls et al., 2004; Wang et al., 2006). This latest version of SiB is termed SiB3.

In SiB3, net ecosystem exchange (NEE) is composed of two component fluxes, gross primary productivity (GPP) and ecosystem respiration (ER), which includes autotrophic (canopy respiration and root respiration) and heterotrophic respiration terms (due to decomposition of dead organic matter),

$$\mathrm{NEE}(x,y,t) = \mathrm{ER}(x,y,t) - \mathrm{GPP}(x,y,t) \tag{1}$$

where x and y represent grid coordinates and t represents time. High-frequency time variations of photosynthesis and respiration are assumed to be well understood and easily modeled processes, i.e. due to diurnally varying quantities such as radiation, temperature, or longer term variations in modeled quantities such as soil moisture etc. Photosynthesis and assimilation are derived using a coupling of equations based upon the work of Farquhar, Collatz, and Ball (Farquhar et al., 1980; Collatz et al., 1992; Ball et al., 1987) while soil respiration is based upon a rather simple function of temperature and soil moisture and constrained in such a way that annual NEE is equal to zero (Raich et al., 1991; Denning et al., 1996).

Several papers have provided comparisons of models to observations, largely by using eddy flux towers to estimate true fluxes of water, carbon, and energy (Baker et al., 2003, 2008; Hanan et al., 2005). Longer-term, more persistent biases are estimated by solving for unknown multiplicative biases in each component flux after smoothing in space and time. While these biases could result from incorrectly modeled short term processes, such as errors in the daily development of the planetary boundary layer, or short-term processes not in the model such as seasonal fertilization and irrigation, the main purpose is to capture longer-term processes not explicitly modeled such as land use change (Robertson et al., 2000; Peterson et al., 1998), disturbances, anthropogenic fertilization effects (Oren et al., 2001), managed forestry (Tillman et al., 2000), and large scale carbon removal (Ciais et al., 2007). This modeling is accomplished by convolving the *influence* functions generated from a lagrangian particle dispersion

model, LPDM (Uliasz and Pielke, 1991; Uliasz, 1993, 1994; Uliasz et al., 1996; Zupanski, 2007), with gridded gross primary productivity (GPP) and total ecosystem respiration (ER) at each time step in SiB3-RAMS. The LPDM transport scheme reverses advection derived from RAMS at very fine time scales and parameterizes vertical turbulent diffusion according to a Gaussian process. A large advantage of this model is the ability to simulate transport of atmospheric constituents at sub grid scales, reducing representation error that might be caused by associating an observing tower with a 40 km grid cell in the model. By tracking particles upwind, backward in time, from the towers, one may make inferences about the contribution of upstream GPP and ER sources.

In particular, we have estimated regional fluxes from atmospheric mixing ratios by assuming that the model of the component fluxes is biased, and that the biases are smoother in time and space than the fluxes themselves:

$$\mathrm{NEE}(x,y,t) = (1+\beta_{\mathrm{RESP}}(x,y))\mathrm{ER}(x,y,t) - (1+\beta_{\mathrm{GPP}}(x,y))\mathrm{GPP}(x,y,t) \quad (2)$$

The model domain, shown in Figs. 1 and 2, consists of most of the United States as well as a large portion of Canada and the northern portions of Mexico. Both SiB3 and RAMS were run on a single 150×90 grid of 40 km cells, with SiB3 utilizing 3 patches per cell to capture subgrid-scale variability in land cover. RAMS meteorology was nudged with 40 km forecast meteorology from the National Center for Environmental Protection's Eta model throughout the domain using a 4 dimensional data assimilation (4DDA) scheme to produce more reliable wind fields. Soil classes were calculated from 5 min "% clay/% sand /% silt" soil data from the International Geosphere-Biosphere Programme (IGBP) (Fig. 1). Biomes were extracted from the UMD classification scheme of the MODIS 12 Landcover 1 km product and mapped to the most similar SiB biome class for all cells and for each of the three patches used (Fig. 2). An exception are the C_4 vegetation classes, grasses and crops, which were projected onto the MODIS biomes from (Wang et al., 2006). The crop characterization is admittedly simple and more work is currently being done to incorporate more accurate crop maps and more

realistic crop modeling into SiB (Lokupitiya et al., 2009). SiB has traditionally calculated fPAR, which defines the fraction of photosynthetically available radiation that is absorbed by the plant canopy, and leaf area index (LAI) using satellite derived NDVI fields. The code was changed to use fPAR and LAI fields derived by the Moderate Resolution Imaging Spectroradiometer (MODIS) (Mu et al., 2007) and averaged over appropriate biome-areas based upon the three patch scheme. SiB3 was run with these 8-day fPAR and LAI products that were provided by the Numerical Terradynamics Simulation Group at the University of Montana who generated it for use in constructing the official Moderate Resolution Imaging Spectroradiometer GPP product.

Modeled carbon dioxide at the tower is calculated as the sum of 3 component fluxes convoluted by *time* and *tower* dependent transport.

$$CO_2(\text{time},\text{tower}) = \text{Transport}_{\text{time},\text{tower}} \begin{pmatrix} \text{Boundary Inflow } (x,y,\text{time}) \\ +\text{Fossil Fuel } (x,y,\text{time}) \\ +\text{Domain Biogenic Fluxes } (x,y,\text{time}) \end{pmatrix} \quad (3)$$

The boundary inflow component was calculated by convolving the influence functions from the LPDM model over boundary CO_2 fields derived using a global biosphere-transport model. At any point in time, the boundary inflow is the average of all upstream particles located in a 3 dimensional 40 km thick rectangular "ring" around the domain. CO_2 resulting from the transport of fossil fuels to the towers is calculated by convolving the influence functions from the LPDM model with surface fossil fuel flux estimates. In particular, the boundary CO_2 fields were calculated by combining transport from the parameterized chemistry transport model (PCTM) (Kawa et al., 2004; Parazoo, 2007) and pre-calculated archived hourly SiB3 fluxes (Baker et al., 2007) on a 1.25-degree by 1-degree global grid. The model was spun up for 2000–2004 and the CO_2 was centered around the Northern Hemispheric mean CO_2 for 2004. In addition to this, results from the CarbonTracker project, which provide globally optimized CO_2 concentration fields, are used for comparison purposes.

Fossil fuel fields were constructed using recently available high resolution Vulcan fossil fuel inventory fields (Gurney et al., 2008), at a 10 km horizontal spatial scale

and hourly temporal scale. Previously available fossil fuel flux fields were derived by distributing country-level fossil fuel sources spatially as a function of population at a 1-degree resolution (Andres et al., 1995). The Vulcan fields provide many improvements including the incorporation of mobile emission sources and power plants, often located in areas distant from high density population centers, increased temporal resolution allowing the modeling of diurnal variability, and increased spatial resolution allowing better delineation of high density population centers. The sensitivity to the new fossil fuel fields is tested by running inversions using both the Vulcan fields as well as the Andres et al. (1995) fields.

SiB3 balances carbon annually by assuming that ER is in approximate balance with GPP on an annual time frame for each surface location and therefore annual NEE is zero for each surface location or grid cell (Raich et al., 1991; Denning et al., 1996). While this is accurate to a large degree and provides reasonable approximations of respiration on diurnal time scales, it ignores annual imbalances in carbon due to a number of external factors such as land use, fertilization effects, disturbance history, etc. For example, aerial photos and satellite images of the coastal mountains of Oregon show a patchwork of forest ages largely dictated by forest management practices. Under standard models of forest regrowth, a regenerating forest will eventually enter a long period in which carbon is being drawn from the atmosphere and stored in wood and roots, thereby providing a sink of carbon from the atmosphere to the biosphere. Conversely, recently burned forests usually enter a short-term period in which they represent a significant carbon source to the atmosphere. Processes such as these that are largely responsible for annual imbalances in NEE are not characterized in SiB3.

The effect of this on boundary inflow estimates is that the PCTM-SiB3 calculated boundary CO_2 fields lacks the effect of sources or sinks in 2004. Given the consensus opinion of an annual mean sink for carbon resulting from the biosphere, this means that the CO_2 fields used will be biased somewhat by the effect of not including this expected global sink. We investigate the effect of this by including a comparison of the inversion using CarbonTracker optimized CO_2 concentration fields for boundary inflow,

which provides an estimate of sources/sinks. As of this time, carbon dioxide resulting from forest fires is not included in the global PCTM-SiB3 inflow or domain SiB3 runs, but is included in the CarbonTracker inflow providing one more contrast between the two fields.

2.2 Observational data

Calibrated CO_2 observations were provided half-hourly at eight measuring sites (WLEF, Harvard Forest (Urbanski et al., 2007), ARM, BERMS, Fraserdale, Western Peatland, WKWT, and Argyle (ME)) for 2004 (Parazoo, 2007). Gerbig et al. (2003) found mean standard deviations on the order of 0.6 to 1 ppm when viewing morning and afternoon vertical profiles of CO_2 in the mixed layer. As a consequence, robust afternoon snapshot observations, at 12, 2, 4, and 6 p.m. LT, are used to lessen the impact of low quality modeled measurements made during times of extremely stable and stratified nocturnal atmospheric conditions near the ground. One exception is the WKWT tower in Moody, TX. Data at this tower consistently showed high diurnally-influenced CO_2 concentrations in the 12 p.m. LT records for most days. It is uncertain exactly what the cause of this is but it appears that it may be due to some kind of systematic late venting of nocturnal respiration-based CO_2 buildup. For this tower, mixed boundary layer conditions appeared to be better represented by snapshot observations shifted by 2 h: 2, 4, 6, and 8 p.m. LT. The first 10 days of the year are not comparable due to a lack of transport preceding 2004. In all there were 2433 missing observations, resulting in 4(observations/day)×8(towers)×355(days)−2433(missing)=8927 observations.

In a previous pseudo-data inversion using a very similar model (Zupanski et al., 2007), the errors on the observations were assumed to be 1 ppm for afternoon observations. Nevertheless, relative to the inversion techniques presented in the next section, the errors on these observations should include errors due to calibration error, mapping error, transport error, and representation error. For this inversion, transport error and representation error are likely the largest components which are notoriously tricky to quantify. Investigations into the sensitivity of inversion test results combined

with initial maximum likelihood estimation results suggest errors in the range of 5–6 ppm are appropriate for this particular inversion. For the remaining inversions, the errors are assumed to be identical and independently distributed (i.i.d.) mean zero errors with standard deviation set to 5.5 ppm. It should be noted that while it is possible to run inversions with artificially low prescribed "observation" errors, this will generally manifest itself in a need to "over tighten" the a priori covariance structure.

2.3 Climatic conditions for 2004

The 2004 year was the 6th wettest in the contiguous United States over the preceeding 110 years (1894–2004). It was also warmer than on average. Nevertheless, there was a great amount of variability in precipitation and temperature as a function of location and season. Drought continued in the west through the summer of 2004, essentially prolonging a multi-year period of drought conditions. The spring was also very dry for the southeast, extending a period of dry conditions from late in 2003. However, summer brought increased precipitation to the east and southeast, culminating in enormous amounts of rain in late summer and early fall due to an extremely active hurricane season. The south (Texas, Louisiana, Mississippi, Arkansas, Oklahoma, and Kansas) had the wettest summer on record and was much cooler than average. These conditions were important as they provided initial conditions for the inversion that involved soil moisture induced plant stress over large areas of the United States.

2.4 Inversion technique

Standard multivariate Gaussian assumptions are made and data are assimilated using a modified Kalman Filter algorithm (Kalman, 1960). In particular, for an initial length n CO_2 measurement vector $\boldsymbol{y}$ representing the first set of measurements, length m unknown CO_2 flux bias vector $\boldsymbol{\beta}$ (dimensionless), $n \times n$ observation error covariance matrix $\boldsymbol{\Sigma}$ (ppm^2), $n \times m$ Jacobian flux-transport matrix $\mathbf{G}$ (ppm), length m prior flux bias estimate $\boldsymbol{\beta}_0$ (dimensionless), and $m \times m$ model-prior mismatch covariance matrix $\boldsymbol{\Sigma}_0$

(dimensionless), the Bayesian statistical assumptions are:

$$\begin{aligned} \boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{\Sigma} &\sim N(\mathbf{G}\boldsymbol{\beta},\boldsymbol{\Sigma}) \\ \boldsymbol{\beta} &\sim N(\boldsymbol{\beta}_0,\boldsymbol{\Sigma}_0) \end{aligned} \tag{4}$$

The posterior distribution of the flux bias vector can be solved for analytically and is:

$$\begin{aligned} p(\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{\Sigma}) &\propto -\frac{1}{2}\left[(\mathbf{G}\boldsymbol{\beta}-\boldsymbol{y})^T\boldsymbol{\Sigma}^{-1}(\mathbf{G}\boldsymbol{\beta}-\boldsymbol{y})+(\boldsymbol{\beta}-\boldsymbol{\beta}_0)^T\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\right] \\ &\sim N\left(\left(\boldsymbol{\Sigma}_0^{-1}+\mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G}\right)^{-1}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\beta}_0+\mathbf{G}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y}\right),\left(\left(\boldsymbol{\Sigma}_0^{-1}+\mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G}\right)\right)^{-1}\right) \end{aligned} \tag{5}$$

With a little bit of algebra, one can rewrite the mean of the posterior distribution of the mean, giving the Kalman-filter updating equation for the mean.

$$E[\boldsymbol{\beta}] = \boldsymbol{\beta}_0 + \left(\mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G}+\boldsymbol{\Sigma}_0^{-1}\right)\mathbf{G}^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\mathbf{G}\boldsymbol{\beta}_0) \tag{6}$$

The posterior mean and variance of $\boldsymbol{x}$ are then fed into the next filter step with a new set of measurements. This particular inversion estimates biases over 7-day periods using available data from that 7-day period of time. Therefore, bias estimates for both ecosystem respiration and GPP as well as corresponding variance estimates are available for all of 2004 with the bias estimates changing with a weekly resolution.

Two difficulties often arise when using filter-style correction schemes. The filter estimates can drift away from realistic values if the data are not plentiful or precise enough to constrain it. Secondly, the nature of the Kalman filter at each step is to create posterior variance estimates that are in general smaller than the prior estimates. This can essentially cause the filter to get "stuck", when an explicit dynamical model of the biases is not available, and thus produce unrealistically small posterior variance estimates around the biases. There is generally no easy solution to this problem. Artificially inflating the posterior variance at each filter step is one method in which one can try to circumvent (Zupanski et al., 2007). This accommodates the fact the biases are likely to change in reality and it allows the filter to consider a wider range of possibilities

for the bias factors. However, it does not necessarily constrain the biases to any particular "reasonable" region of values allowing the bias estimates to drift into unrealistic parameter space. Therefore, we have chosen to weight the filter at each step with a "grand" prior. This effectively handles both of the preceding problems. With respect to our inversion, there will be three pieces of information at each step, the grand prior which is derived from the forward SiB3-RAMS model with an error assumption, the local prior which is derived from the previous filter step's posterior flux bias distribution, and the data which forms the statistical likelihood function. In some sense, this new piece of the covariance structure provides a bound upon how much the inversion can "learn" about the bias structure.

In order to quantify, we denote the grand prior as a multivariate Gaussian distribution around $\boldsymbol{\beta}_{\text{grand}}$ with covariance matrix $\sigma^2_{\text{grand}}\boldsymbol{\Sigma}_{\text{grand}}$, and additional weight factor w, and we rewrite the expression given in Eq. (4) as:

$$p(\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{\Sigma}) \propto -\frac{1}{2}\Big[(\mathbf{G}\beta-\boldsymbol{y})^T\sigma^{-2}_{\text{obs}}I(\mathbf{G}\beta-\boldsymbol{y})+(\boldsymbol{\beta}-\boldsymbol{\beta}_0)^T\sigma_0^{-2}\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)+(\boldsymbol{\beta}-\boldsymbol{\beta}_{\text{grand}})^T \times w\sigma^{-2}_{\text{grand}}\boldsymbol{\Sigma}^{-1}_{\text{grand}}(\boldsymbol{\beta}-\boldsymbol{\beta}_{\text{grand}})\Big] \tag{7}$$

Thus $\boldsymbol{\beta}$ is distributed as a multivariate Gaussian with parameters:

$$\begin{aligned}\text{Mean}(\boldsymbol{\beta}) = E[\boldsymbol{\beta}] &= \left(w^{-1}\sigma^{-2}_{\text{grand}}\boldsymbol{\Sigma}^{-1}_{\text{grand}}+\sigma_0^{-2}\boldsymbol{\Sigma}_0^{-1}+\mathbf{G}^T\sigma^{-2}_{\text{obs}}\mathbf{IG}\right)^{-1} \\ &\quad\times\left(w\sigma^{-2}_{\text{grand}}\boldsymbol{\Sigma}^{-1}_{\text{grand}}\boldsymbol{\beta}_{\text{grand}}+\sigma_0^{-2}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\beta}_0+\mathbf{G}^T\sigma^{-2}_{\text{obs}}\mathbf{I}\boldsymbol{y}\right) \end{aligned} \tag{8}$$

$$\text{Variance}(\boldsymbol{\beta}) = E[\boldsymbol{\beta}^2]-(E[\boldsymbol{\beta}])^2 = \left(\left(w\sigma^{-2}_{\text{grand}}\boldsymbol{\Sigma}^{-1}_{\text{grand}}+\sigma_0^{-2}\boldsymbol{\Sigma}_0^{-1}+\mathbf{G}^T\sigma^{-2}_{\text{obs}}\mathbf{IG}\right)\right)^{-1} \tag{9}$$

Equation (7) specifically separates out the variance scalars, σ^2_{grand}, σ^2_0, and σ^2_{obs} from the covariance matrices, leaving the covariance matrices essentially scaled to 1. The w weight is a redundant factor and is simply included to facilitate easier interpretation of tightening/loosening of the grand prior covariance (around the SiB3 derived a priori

carbon fluxes). Unless otherwise specified, this weight, w, on the grand covariance matrix is set to 2. This means that the initial variance around the grand prior is increased, thus providing a weaker constraint. For the initial filter step, only the grand prior is used. After that point, there exist both a grand prior and a prior (from the posterior of the previous filter step). The inversion is further constrained by the assumption of spatially correlated errors in the grand prior, i.e. the covariance matrix $\mathbf{\Sigma}_{\text{grand}}$ will take on the following form.

$$\mathbf{\Sigma}_{\text{grand}} = \begin{bmatrix} \mathbf{\Sigma}_{\text{Respg,prior}} & 0 \\ 0 & \mathbf{\Sigma}_{\text{Assimn,prior}} \end{bmatrix} \tag{10}$$

The respiration and GPP covariance matrices are each formed from the exponential covariance function, where $t_{i,j}$ is the distance between points β_i and $\boldsymbol{\beta}_j$.

$$\text{Cov}(\beta_i, \boldsymbol{\beta}_j) = \begin{cases} \sigma_0^2(1-\alpha_0)\exp\left(\frac{-t_{i,j}}{h_0}\right), i \neq j \\ \alpha_0\sigma_0^2, i = j \end{cases} \tag{11}$$

The h_0 parameter is the range, or decorrelation length scale parameter, giving the distance at which the covariance between two points is equal to $\sigma_0^2(1-\alpha_0)e^{-1}$.The σ^2 parameter is the scalar variance parameter and determines the variance of the marginal distribution of the particular flux component. The parameter α_0 controls what percentage of the covariance can be attributed to spatial covariance, as opposed to spatially independent errors.

Inversion techniques can be extremely sensitive to assumptions. It was shown in (Schuh et al., 2009) that this inversion model is robust to small spatial scale random deviations in flux bias and that post-aggregated (in space) estimates can be very good even when using a fairly sparse network of towers observing CO_2. Nevertheless, given the unconstrained nature of the inversion problem, it is always important to assess the impact of varying certain unknown parameters in the inversion, such as spatial decorrelation length scales, the weight given to the "grand" prior, and the fixed CO_2 contributions from both the boundary inflow and fossil fuel sources.

2.5 Sensitivity

The inversion essentially guarantees some improvement in prediction of observed CO_2 (Eq. 5). However, when using a regression style approach in a heavily unconstrained environment, this improvement can often be overstated because of the great freedom the inversion has to fit the data. Therefore, it is often desirable to go beyond simply comparing observed carbon dioxide at the towers to model-based predicted carbon dioxide. Comparing model observations to independent observations not used in the inversion, comparing models which predict similar quantities, as well as testing the sensitivity of the model to variations in unknown parameters are all methods of generating more confidence in estimates.

We used a variety of different procedures to test the sensitivity of the inversion. Regional inversions have been shown to be very sensitive to boundary inflow variations. Therefore, we first test the sensitivity of the inversion to varying the inflow of CO_2 at the boundaries. To do this, we derive boundary inflow to the 8 towers using the LPDM model and optimized carbon dioxide concentration fields from the CarbonTracker project (Peters et al., 2007). Inversion results are then compared with the results derived from the LPDM model and the PCTM inflow. Secondly, we vary several different variance parameters and derive annual domain-summed NEE and tower observation based RMSE based upon the varied parameters. Thirdly, we use a resampling procedure in which we create 100 different observation data subsets by holding out a randomly selected 50% of the observation data for each. Each set of data is run through the weekly inversion scheme and the sensitivity of the predicted CO_2 at the towers and the estimated flux biases is explored. This provides estimates of the variability of the flux correction factors and can be used to assess the sensitivity of the source/sink to the constraint provided by the data. Using the held out data as independent evaluation data and the complementing data as training data for the inversion, one may also derive a more accurate estimate of root mean-squared error (RMSE) of the inversion-optimized fluxes. We test the impact of the high resolution Vulcan fossil fuel

inventory on the inversion results by comparing inversion results relying upon Vulcan to those results utilizing the Andres et al. (1995) fossil fuel inventory.

SiB3 has been evaluated at many sites and over many time periods, nevertheless, the particular model run used for the a priori flux estimates was not optimized to fit the flux data at any site in particular. Even though there is a mismatch in representation, with the flux towers representing footprints of less than a square kilometer and the inversion results representing flux estimates on the scale of thousands of square kilometers, we believe that these comparisons are of value, especially in locations that are more spatially homogeneous than others, such as grasslands and large forest reaches. This is then the fourth comparison we make.

3 Results

As was indicated in the previous section, there are a number of variables that the inversion will likely be sensitive to and therefore the results are expected to be quite variable. For results, we choose to present one particular case with a fixed set of inversion inputs as an initial case study and then use it to compare the effect of varying the boundary inflow and the source of the domain fossil fuel fluxes. With reference to the preceding section and Eq. (7) in particular, the following values are used for these inversions: σ_{grand}=0.25, σ_0=0.25, σ_{obs}=5.5ppm, w=2, h_0=1000km. In particular, a value of σ_{grand}=0.25 would mean that we expect that approximately 68% of the GPP and ER biases are within ±25% of the original SiB3 estimated fluxes, with 95% within ±50%. This variation when combined with positive spatial correlations was shown to provide a reasonable a priori range of annual domain-summed NEE. These deviations must generally be kept to less than 30–40% to ensure that posterior ER and GPP fluxes are not reduced by more than 100%, which makes no conceptual sense. We then test the sensitivity of the results over a number of varying inversion inputs using the PCTM boundary conditions and the Vulcan fossil fuel flux field.

3.1 General structure of results

CO_2 can be predicted by invoking the relationship shown in Eq. (3). The predicted mean observed CO_2 is derived as $\mathbf{G}\hat{x}$ where $\hat{x}$ represents one (for the prior fluxes) plus the inversion-optimized flux biases. Using the PCTM boundary conditions and the Vulcan fossil fuel inventory, a comparison of the inversion-corrected posterior predictions at the towers to the observations is shown in Fig. 3. For domain-summed temporal plots, NEE is calculated via Eq. (2) while ER and GPP are calculated via the two respective summands on right hand side of that equation. These domain-summed temporal results are shown in Fig. 4.

The observed carbon dioxide concentrations contain information that infers a dampening of the a priori annual GPP cycle, and hence the a priori annual ER cycle (due to the strong correlation of the annual sums of each). Since both GPP and ER are significantly dampened, it is not surprising that the NEE signal is dampened as well. Furthermore, the data suggest a weak temporal shift in the prior NEE signal. This manifests itself as a stronger, but more gradual onset of spring, followed by a weaker overall carbon sink over the middle and late summer periods.

We use a resampling procedure to account for variability that might be associated with over fitting the model and which provides additional variability to the standard covariance estimates of the biases given in Eq. (6). Forty-five different inversions are run, each based upon a different subsample of the observations. Assuming temporal independence of the errors in the filter, one may simulate properties of the annual NEE probability density functions (pdf) for each of these 45 inversions by using the posterior covariance provided at each step of the Kalman Filter for each inversion. A 95% confidence interval (CI) for the entire domain can be calculated at each step of the filter for each of the 45 inversions. The CI shown in Fig. 4 then characterizes variability in the NEE by selecting the 95% CI of each set of 95% CIs for each weekly time step.

The ensemble mean of the domain summed annual NEE flux is approximately −0.68 Pg/yr while the standard deviation of this estimate is about 0.11 Pg/yr. It is important to note that this standard deviation estimate does appear to be too small, giving tighter bounds on the flux than found in other inversion papers (Gurney et al., 2002; Peters et al., 2007). An additional source of variability in the estimate is discussed later (Sect. 3.4) and likely provides another 0.1–0.15 Pg/yr to this standard deviation estimate. The spatial representation of these sources and sinks can be seen in the first panel of Fig. 7. Depictions of this variability in a spatial framework are shown in Fig. 5. This variability is partitioned into two pieces, variability associated with the spread of mean estimates over the 45 inversions (measure of over fitting) and variability associated with summing up the posterior variances at each filter step (regular KF variance) evaluated over all 45 inversions. Besides the spatial display of posterior variance information for NEE, which roughly tracks the convolution of the sampling footprint of the network and the prior ER/GPP signals, the results show that over fitting the model may provide a significant source of variability comparable to that which is normally constructed from each filter step's posterior covariance matrix.

3.2 Sensitivity and robustness of results to inflow

Inflow of CO_2 from the boundaries has typically been a large concern of regional models (Gerbig et al., 2003; Peylin et al., 2005). In extremely limited domain problems, the variance of the CO_2 coming in from the boundary can easily dwarf the changes inside the domain due to local biotic uptake and release. Therefore it is of interest to gauge the sensitivity of the inversion to varying boundary inflows. The boundary conditions included in this model were constructed from a global simulation using SiB3 and PCTM (Parazoo et al., 2007). The CarbonTracker project has provided CO_2 mixing ratio data based upon globally optimized fluxes (Peters et al., 2007). SiB3 has no annual source/sinks whereas CarbonTracker includes an annual source/sink estimated from observations of CO_2. A plot of the difference between the two inflows is shown in Fig. 6. The inflow annual mean and temporal pattern is very similar for PCTM

and CarbonTracker with the main difference being a seasonally stronger cycle in the PCTM-SiB3 results, likely a result of the underlying biosphere model, SiB3, providing a stronger seasonal GPP/NEE signal than the corresponding CASA model used in CarbonTracker. In addition to running comparison inversions between these two CO_2 inflow estimates, we also run the inversion with a fixed inflow estimate of 378 ppm representing the annually averaged PCTM inflow over the period of the simulation in order to show the necessity of reasonable boundary inflow values in calculating reasonable source/sink estimates.

Figure 7 shows a comparison plot of maps of the annual mean NEE estimate based upon CarbonTracker (w/CASA), PCTM (w/SiB3), and the fixed inflow condition. The results are similar for the CarbonTracker and PCTM inflows. Both results have similar spatial and temporal characteristics but differ mainly in magnitude. The PCTM-based inversion results in a sink of 0.1–0.2 Pg/yr more than that of the CarbonTracker-based result. The PCTM-based boundary conditions do not account for the expected global carbon sink outside of the inversion domain, which forces the inversion to increase the North American sink to compensate. This results in the PCTM-inflow based inversion having a larger annual sink estimate than the CT-inflow based inversion. The sink estimated with the PCTM inflow was 0.65 Pg/yr while the sink estimated with the CarbonTracker inflow was estimated at 0.48 Pg/yr. It does seem somewhat surprising that the results from the two inflows are still close, within approximately 30% of one another. This indicates that local observations may be affected significantly more by local fluxes than by larger scale fluxes in distant locations outside of the model boundary.

3.3 Sensitivity of results to fossil fuel inventory

Until the release of the Vulcan fossil fuel inventory in 2008, most researchers were reliant upon the Andres et al. (1995) fossil fuel inventory, which was released at annual time scales and at a 1-degree resolution over the globe. For many large-scale inversion applications, this inventory is adequate. However, for higher resolution studies within the United States, the Vulcan fossil fuel inventory provides a dramatic improvement in

both space and time accounting of fossil fuel fluxes. The main difference between these inventories is the redistribution of some fossil fuel sources from population centers to more distant locations representing mobile sources and power plants. The Vulcan fossil fuel flux estimates are at a much higher resolution in both time and space. Previous inversions had to grapple with the fact that some observing stations are located within enormous fossil fuel flux regions. For example, a semi-rural location like Harvard Forest would very likely be located in the same grid cell as the large metropolitan city of Boston. Given no sub-annual temporal resolution to the fossil fuel fluxes, an observing tower located at Harvard Forest was often seeing a 24 h continuous stream of fossil fuel fluxes arising from a city over 100 km away. However, the 10 km horizontal resolution of the Vulcan inventory allows these to be separated and additionally provides a diurnal and seasonal estimate of these fluxes, which is important for inversions based upon hourly observations.

In order to gauge the impact of incorporating the Vulcan data, we first contrasted the contributions to each of the 8 towers from each of the inventories. For many of the stations, the afternoon differences between the two were very small. Differences at the ARM site in Oklahoma, the WLEF site in Wisconsin, the Canadian sites, and the Argyle, Maine site were on the order of a few ppm. Differences at the Moody, Texas tower were in the range of −5 to 5 ppm. While the differences across most towers were relatively small, the differences at Harvard Forest were between −25 and 30 ppm!

The difference in the annual NEE estimate is shown in Fig. 8. The effect on the inversion is far from trivial with differences of up to 150 g/m^2 per year recorded along the northeast coast of the United States, similar in magnitude to the maximum annual sinks estimated by the inversion. These differences are a result of coarse fossil fuel flux fields providing artificially high sources of CO_2 to the Harvard Forest tower which must be neutralized via a large local sink.

3.4 Sensitivity and robustness of results to prior variance structure

A test of the sensitivity and effect of the prior upon results is important because of the use of an informative Bayesian prior, that is, a prior flux estimate in which the inversion will likely be sensitive. With reference to Eqs. (5) and (7), the w, σ_0^2, and h_0 parameters are varied and results are shown in Fig. 9. These figures show that results are sensitive to nearly all of these parameters, providing different degrees of RMSE and sink strength depending upon the particular combination. In particular, sink estimates range between 0 and 1 Pg/yr. The ensemble of estimates, over the various possible a priori variance parameters, has a standard deviation of approximately 0.2 PgC/yr. This likely contributes another 0.1 to 0.15 PgC/yr (depending upon the existence of correlation between the variance shown here and earlier variance estimates due to jackknife resampling and the Kalman filter posterior variances) to the initial standard deviation estimate of 0.11 Pg/yr given earlier. This would give an adjusted standard deviation estimate of approximately 0.2–0.25 PgC/yr to the posterior annual NEE estimate shown in Fig. 4.

An RMSE-weighted average of the sink estimates show a sink of 0.57 PgC/yr, 20% higher than our single case scenario that we have followed throughout these results. Values very near the lower left of the plot are somewhat unrealistic since low spatial correlation (h_0) and a low variance on the prior (σ_0^2) will not provide a reasonable enough range around the prior to provide a realistic posterior sink estimate which generally is thought to range between 0 and 1.5 PgC/yr (Schimel et al., 2000; Gurney et al., 2002) inter-annually. Increasing either the variance multiplier (along x-axis) or the spatial decorrelation length scale (along y-axis), or both jointly, increases the error variance around the a priori mean allowing more realistic domain-wide summed posterior flux estimates. Therefore if one "de-weights" these sink estimates occurring in the lower left hand portions of the panels in Fig. 9, the RMSE-weighted sink will likely increase to more than 0.57 PgC/yr.

The weight of the grand prior (w) has two effects. First, it constrains solutions back towards the prior, essentially anchoring the Kalman filter so that, over time, it does not drift too far from the prior. Given the fact that this grand prior is fixed in time, it also provides a degree of variance inflation (over the regular KF) by providing a lower bound on the prior variance for each filtering step. It is interesting to note that, for cases in which the global prior is weaker (bottom two panels), the maximum sink estimate occurs on the inside of the plot bounds and not at the boundary. The Kalman filter becomes more entrenched without the grand prior since there is no lower limit on the prior variability at each inversion filter step and there is no inflation. Therefore it is likely that the initial reduction in respiration and associated "sink" of carbon in the early months of the year becomes entrenched and leaves a strong sink signature on the rest of the year resulting in the largest sink estimates. We did not test any additional forms of variance inflation on the model and acknowledge that additional efforts are needed to construct more robust filter techniques.

3.5 Comparison to CarbonTracker flux estimates

Given the fact that the majority of the underlying observations supporting the inversion were also used in the CarbonTracker project, one would expect posterior flux estimates to be somewhat similar. One of the most important differences between these inversions and CarbonTracker is the optimization of encompassing global fluxes, which affect CO_2 concentrations within our domain. However, this can be mitigated somewhat by the use of optimized CO_2 concentrations from CarbonTracker in the inversion. Under this scenario, one would expect the inversion results to be similar to CarbonTracker but there are still many differences. As can be seen in Fig. 10, the carbon fluxes in the priors, CASA and SiB3, play an important role in the posterior estimates. The posterior estimates of both inversion models display the signature of the a priori fluxes prominently. These results would lead one to believe that either the data does not provide sufficient constraint or the covariance structure is specified too tightly around the prior.

3.6 Comparison to filled level 4 Ameriflux data

Posterior respiration and GPP estimates from the model can also be compared to Ameriflux level 4 data. As indicated earlier, there is a spatial representation mismatch in doing so due to the fact that the model estimate is an average over approximately 1600 km^2 and the associated flux tower estimate is over a much smaller footprint, likely less than 1 km^2. Nevertheless, some useful comparisons and observations can be made. Figure 11 shows comparisons of the model to the observations for weekly ER and GPP at three Ameriflux sites, which appear in the more observation constrained portion of the model domain. The ARM site is one of the more constrained sites in the domain and lies in a relatively homogenous landscape making it an excellent candidate for analysis. The prior site NEE estimate appears to be improved on average by the posterior flux estimates. In particular, the prior model is corrected significantly in the summer when it predicts significant respiration occurring. Clearly one can see an early spring winter wheat signal in the observations, forming a significant amount of carbon drawdown over an 8–10 week period. SiB3 necessarily balances GPP and ER annually and is thus forced to redistribute this carbon into respiration in other portions of the year. This is the likely reason for displacement of the prior estimate in the summer. The posterior corrects for a large portion of this but the large distance between the prior and observed fluxes make a complete correction difficult. Just as important, but perhaps more subtle, is the fact that the inversion is able to provide significant corrections to ER and GPP separately. SiB3 appears to significantly overestimate GPP. However, due to the annual NEE balance constraint, SiB3 will overestimate ER as well, providing an NEE signal that appears very reasonable. If the forward model is only compared to NEE estimates at various sites then this fact can be easily overlooked but is likely very important to biosphere dynamics on certain time scales.

3.7 Evaluation of annual NEE source/sinks against ancillary data and hypotheses

Using two sets of boundary conditions, we arrived at a final sink estimate of approximately 0.5–0.7 PgC/yr±0.25 PgC/yr. This is significantly less than CarbonTracker's sink estimate of 0.9 PgC per year and other estimates currently emerging from an ongoing top-down synthesis project. It is clearly possible that other globally based inversions provide more constraint on certain areas of North America, such as the Pacific Northwest forest regions of North America, the Southeastern United States, or extreme Northeast Canada. Both of these areas have large annual GPP signals and are thus capable of being a strong source/sink of CO_2. However, our inversion results show a generous sink in the coastal N.W. forests while CarbonTracker shows little sink there. Furthermore, CarbonTracker's sink is largely located in the agricultural Midwest of the United States (and a portion of Canada), an area reasonably constrained by the observation network we've used.

On the other hand, perhaps the globally based sink estimates are too high. The recently completed State of the Carbon Cycle Report (SOCCR, 2007) provides an inventory-based sink estimate for North America of approximately 0.66 PgC per year (land sink) using a variety of data sources collected over the last ten to fifteen years. Uncertainty is presented as a 95% confidence interval, 0 to 1.32 PgC. This is similar to what we've recovered in these inversions. However, this is a mean sink estimate over many years and 2004 is believed to be a year in which the sink in North America was very strong, likely putting the SOCCR estimate closer to 0.8–0.9 PgC/yr, the upper range of their annual estimates. Stephens et al. 2007 called into question the magnitude of the Northern Hemispheric (and North American) global annual NEE sink which has been a cornerstone of inversion results for the last 10 years (Fan et al., 1998: Gurney et al., 2002; Peters et al., 2007) indicating that it may be much smaller than previously assumed. In any case, the rapid expansion of the calibrated CO_2 tower network (currently over 30 towers in North America) should soon provide significant additional

data constraints to researchers performing atmospheric CO_2 based inversions.

The spatial character of the annual NEE estimate has several distinctive features. The most definitive feature of the annual NEE estimate shown in Fig. 7 is the large sink located over Texas, Louisiana, Arkansas, and portions of Oklahoma. This sink is located largely between, and to the east of, the ARM and WKWT sites in south central portion of the domain. At first glance this may appear to be an artifact of incorrect transport, poor boundary conditions, or incorrect fossil fuel emissions specifications. However, summing the ARM NEE observations for the year provides a sink estimate of approximately 275 g/m^2, similar to the estimates the inversion produces to the south of the ARM site (Fig. 7). A likely hypothesis for this sink is the lateral export of crops, primarily winter wheat that draws most of its carbon from the atmosphere in the spring and then is harvested and exported in early summer. The WKWT tower concentrations have proven to be somewhat difficult to model given its late diurnal venting of nocturnally built up carbon dioxide, its close proximity to both the model boundary and the ocean, and its proximity to fossil fuel sources of major metropolitan areas and oil refining facilities. Given the negative correlation often exhibited in annual NEE from the north of WKWT to the south of WKWT, it is possible that errors resulting in sources to the south are in part responsible for the strength of the sink to the north.

The aforementioned sink also extends to the east and northeast of the ARM tower. This is an area of significant crop production, with corn and soybeans being grown extensively in the northern portions while soybeans, rice, and other crops are grown to the south in the Arkansas/Mississippi region. This area is also covered by heavily managed forest regions, which produce large annual harvests of wood primarily for paper pulp. These managed forests are largely composed of very young productive loblolly pine trees providing a major source of carbon sequestration. This area is known for quite variable precipitation patterns and it would seem to reasonable to assume that young productive forests in this area would be very productive under the unusually wet and cool conditions of 2004.

It is interesting to note that the most intensely cultivated portion of the Midwestern United States, centered on the state of Iowa, shows little to no sink. This is an area typically planted extensively with corn, which has been shown to be an extremely effective consumer of atmospheric CO_2. The a priori estimate of NEE based upon SiB3 included a very strong summer time sink of carbon over the Iowa region using a C_4 photosynthesis scheme from Collatz et al. (1992). Whether the CO_2 flux is reasonably close to the truth is difficult to determine although the increased amplitude of the seasonal cycle due to the increased summer time GPP of the corn seems to have put a slight signature upon annual NEE in the inversion results (top left panel of Fig. 7).

One other hypothesis for this disparity in sink strength concerns the lateral transport of crop harvest. Significant annual sinks can only occur if carbon is added or removed from the system. In the case of croplands, the mechanism for this is usually the export of harvested crops, which should provide the appearance of a carbon sink in the area. Annual NEE estimates from the corn-planted Bondville, IL Ameriflux site indicate a sink on the order of 500–600 g/m^2. Soybeans can be expected to provide sinks of about half of this. Assuming steady state conditions over several years, these types of sinks can be attributed directly to the harvest. Approximately 20% of the corn harvest and 35% of the soy harvest is exported overseas, mostly for animal feed, while half of the corn and soy retained in the United States is used to feed livestock domestically (National Corn Growers Association website: http://www.ncga.com/files/pdf/2009WOC.pdf, Soy Stats, http://www.soystats.com). Most of the carbon in this livestock feed is then returned to the atmosphere as CO_2 and CH4 at locations where it is consumed by livestock. Almost 70% of the feedlots in the United States are located in just 3 states: Texas, Kansas, and Nebraska (http://www.cattlenetwork.com). This may provide a partial explanation for the lack of an agriculturally-induced sink over Nebraska and Kansas, states with very high crop production and intense livestock operations, and the existence of sinks over portions of Arkansas, Mississippi, Missouri and Illinois, states with relatively high crop production but with significantly less livestock operations.

Forested regions in the northwestern United States and boreal forests of Canada show slight sinks. However, variability estimates surrounding these sink estimates are typically much smaller than the variability estimates of similar sink magnitudes in the Midwest or southeastern United States showing relatively more confidence in the sink despite the lack of proximity to the observing towers. The sink estimate in the northwestern United States is not surprising since the northwestern coastal mountains of California, Oregon and Washington have been intensely managed over the last 50 years and are expected to provide a sink of carbon for many decades into the future (Alig et al., 2006). The estimate for the boreal forest regions appears much harder to objectively evaluate. Most studies have indicated that Canadian sources should currently be a weak sink, although the projection of this weak sink into the future is highly uncertain. The inversion results show a fairly carbon neutral Canada on average, but shows the boreal forests of central Canada and the boreal and coastal forests of western Canada as slight sinks while the agricultural plains of Canada and the forests of eastern Canada provide slight sources. It is interesting to note that areas to the south of the two Canadian towers show an annual source of carbon in an area just to the east of large expansive forest ecosystems of British Columbia that have recently experienced unprecedented bark beetle invasions and tree mortality. It is important to note that forest fires were not included in the SiB3 domain run for the regional inversion. Average carbon emissions from Canadian forest fires were estimated at 27±6Tg/yr (Amiro et al., 2001), a non-trivial amount that could increase the strength of the boreal forest sink predicted by the inversion.

4 Conclusions

GPP, ER, and NEE flux corrections implied by this inversion provide posterior annual NEE estimates similar to those provided by a number of independently derived models including CASA (via CarbonTracker optimized) and the MODIS 17 GPP product. NEE estimates for the entire domain appear on the low side of estimates derived from global

models, which is understandable given the lack of constraint on some key regions of high annual GPP, and hence potentially high annual NEE. This was corroborated by a comparison to INTEX aircraft data which shows the existence of a deficit in GPP over the southeast which would, when all other things are considered equal, inflate the domain-wide sink closer to levels estimated from global models such as CarbonTracker. Results are relatively sensitive to a number of parameters in the inversion setup, which is also to be expected with an inversion constrained by such a sparse observing network. Using a temporally uniform boundary condition seems to produce a very unrealistic annual sink on the order of 0.38 Pg per year, supporting the notion that regional inversions require realistic boundary inflow of CO_2. However, much to our surprise, we find that two completely independent boundary inflow estimates provide similar results with the main difference being an approximately 30% difference in magnitude. This leads us to believe that, while probably not preferable to optimized global CO_2 fields, the inclusion of annual NEE balanced models in global models used to provide boundary inflow estimation does not significantly damage inversions based upon it.

In the course of trying to improve NEE estimates, we were able to find that the inversion was able to provide some degree of correction to the individual summands of NEE, ER and GPP, which are generally highly correlated at many different scales in time and space. Considering that SiB3 currently calculates ER as a relatively simple function of soil moisture and temperature such that annual ER equals annual GPP, the significant adjustment inferred upon GPP may prove to be valuable estimation of other quantities of interest in the biosphere. For example, while photosynthesizing, plants must generally release water to compensate, meaning that artificially high GPP may infer artificially high water exchange with the atmosphere and possibly associated latent heat fluxes.

The agricultural Midwestern United States appears to play a large role in the inversion results, providing a large sink. However, the sink does not correlate exactly with crop productivity, when compared to crop production maps from the United States De-

partment of Agriculture, and several states with significant crop production such as Nebraska, Kansas, and Iowa, appear to be in approximate annual carbon balance. While the magnitude of this difference between carbon neutral states with crops and carbon sink states with crops is likely influenced by the lack of data in the inversion and the general unconstrained nature of the solution at fine scales, the discrimination between them seems likely to stay. One hypothesis proposed is the lateral movement of crops which has been shown to be a major portion of the carbon budget globally (Ciais et al., 2007). The main crops of interest in the domain are wheat, soy and corn. Soy and corn are grown across large expanses of the north-central Midwest and are primarily used to feed livestock. These livestock are typically fed in feedlots in the states of Iowa, Colorado, Nebraska, Kansas, and Texas, generally located to the west and south of the areas of growth and harvest. The end result would be that eastern states within the Midwest would be a sink because of the near complete export of crops grown there. However, states in the western portion of the Midwest would receive the majority of these crops where they would be fed to cattle and other animals, returned to the atmosphere as CO_2 and CH_4 and largely balance any local sinks due to crop production.

Technical considerations concerning the inversion could also affect these results. In particular, a large amount of missing data for the WKWT (Moody, TX) tower leaves the southern boundary inflow unconstrained beyond the normal PCTM inflow. This could result in the inflation of an Oklahoma/Texas sink to account for a positive bias in the inflow at the southern boundary, particularly after 1 July 2004 when the Midwest receives its heaviest influence from the Gulf of Mexico. The WLEF tower was also missing most of its observations for June, a time of intense drawdown for croplands to the south of the site.

In 2004, the southern states of Texas, Oklahoma, Kansas, Louisiana, Arkansas, and Mississippi had the wettest summer ever potentially mitigating some degree of drought and providing an increase in GPP for the region which includes managed forests, a large percentage of the United States' exported wheat crop, and soybeans

and other crops along the lower Mississippi river valley. Additional research is needed to determine if any of these could represent a plausible hypothesis that would result in the net carbon neutrality of large crop growing states in the western portions of the Great Plains and the expansive southern and Mississippi river valley sink predicted by the inversion.

Acknowledgements. I would like to acknowledge several people who provided their calibrated carbon dioxide data to me and without which this paper would not have been possible. A. Andrews (ESRL/National Oceanic and Atmospheric Administration) provided data for the WKWT (Texas, US) and Argyle (Maine, US) towers, M. Fisher/M. Torn provided data for the Southern Great Plains site (Oklahoma, US), S. Wofsy/B. Munger provided data for Harvard Forest site (Massachusetts, US), and D. Worthy provided data for BERMS, Western Peatland, and Fraserdale sites (Canada). I also would like to acknowledge K. Gurney and B. Andres for their contributions of fossil fuel inventory data as well as A. Jacobsen and W. Peters for the Carbon Tracker boundary inflow CO_2 data and N. Parazoo for the PCTM boundary inflow CO_2 data. Again, these were very generous contributions without which, little of this research could have been done.

References

Alig, R. J., Krankina, O., Yost, A., and Kuzminykh, J.: Forest carbon dynamics in the Pacific Northwest (USA) and the St. Petersburg region of Russia: Comparisons and policy implications, Clim. Change, 79, 335–360, 2006.

Amiro, B. D., Todd, J. B., Wotton, B. M., Logan, K. A., Flannigan, M. D., Stocks, B. J., Mason, J. A., Martell, D. L., and Hirsch, K. G.: Direct carbon emissions from Canadian forest fires, 1959–1999, Can. J. Forest Res., 31(3), 512–525, 2001.

Andres, R. J., Marland, G., Fung, I., and Matthews, E.: A 1°×1° distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990, Global Biogeochem. Cy., 10, 419–429, 1996.

Baker, I., Denning, A. S., Hanan, N., Prihodko, L., Uliasz, M., Vidale, P., Davis, K., and Bakwin, P.: Simulated and observed fluxes of sensible and latent heat and CO_2 at the WLEF-TW tower using Sib 2.5, Global Change Biol., 9, 1262–1277, 2003.

Baker, I., Denning, A., Prihodko, L., Schaefer, K., Berry, J., Collatz, G., Suits, N., Stockli, R., Philpott, A., and Leonard, O.: Global Net Ecosystem Exchange (NEE) of CO_2, Available on-line (http://www.daac.ornl.gov) from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, 2007.

Baker, I. T., Prihodko, L., Denning, A. S., Goulden, M., Miller, S., and da Rocha, H. R.: Seasonal drought stress in the Amazon: Reconciling 3 models and observations, J. Geophys. Res., 113, G00B01, doi:10.1029/2007JG000644, 2008.

Case, J. L., Manobianco, J., Dianic, A. V., and Wheeler, M. M.: Verification of high-resolution RAMS forecasts over east-central Florida during the 1999 and 2000 Summer Months, Weather Forecast., 17(6), 1133–1151, 2002.

CCSP: The First State of the Carbon Cycle Report (SOCCR): The North American Carbon Budget and Implications for the Global Carbon Cycle. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research, edited by: King, A. W., Dilling, L., Zimmerman, G. P., Fairman, D. M., Houghton, R. A., Marland, G., Rose, A. Z., and Wilbanks, T. J., National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, NC, USA, 242 pp., 2007.

Choi, Y., Vay, S., Vadrevu, K., Soja, A., Woo, J., Nolf, S., Sachse, G., Diskin, G., Blake, D., Blake, N., Singh, H., Avery, M., Fried, A., Pfister, L., and Fuelberg, H.: Characteristics of the atmospheric CO_2 signal as observed over the conterminous United States during INTEX-NA, J. Geophys. Res., 113, D07301, doi:10.10129/2007JD008899, 2008.

Ciais, P., Bousquet, P., Freibauer, A., and Naegler, T.: Horizontal displacement of carbon associated with agriculture and its impacts on atmospheric CO_2, Global Biogeochem. Cy., 21, GB2014, doi:10.1029/206GB002741, 2007.

Collatz, G. J., Ribas-Carbo, M., and Berry, J. A.: Coupled photosythesis-stomatal conductance model for leaves of C4 plants, Aust. J. Plant Physiol. 19, 519–538, 1992.

Denning, A. S., Collatz, J. G., Zhang, C., Randall, D. A., Berry, J. A., Sellers, P. J., Colello, G. D., and Dazlich, D. A.: Simulations of terrestrial carbon metabolism and atmospheric CO_2 in a general circulation model. Part 1: Surface carbon fluxes, Tellus, 48B, 521–542, 1996.

Denning, A. S., Nicholls, M., Prihodko, L., Baker, I., Vidale, P., Davis, K., and Bakwin, P.: Simulated variations in atmospheric CO_2 over a Wisconsin forest using a couple ecosystem-atmosphere model, Global Change Biol., 9, 1241–1250, 2003.

Denning, A. S., Zhang, N., Yi, C., Branson, M., Davis, K., Kleist, J., and Bakwin, P.: Evaluation of modeled atmospheric boundary layer depth at the WLEF Tower, Agr. Forest Meteorol.,

148, 206–215, 2008.
Engelen, R. J., Denning, A. S., and Gurney, K. R.: On error estimation in atmospheric CO_2 inversions, J. Geophys. Res., 107(10), 1–13, 2002.
Fan, S., Gloor, M., Mahlman, J., Pacala, S., Sarmiento, J., Takahashi, T., and Tans, P.: A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models, Science, 282, 442–446, 1998.
Freitas, S. R., Longo, K., Silva Dias, M., Silva Dias, P., Chatfield, R., Fazenda, Á., and Rodrigues, L. F.: The coupled aerosol and tracer transport model to the brazilian developments on the Regional Atmospheric Modeling System: validation using direct and remote sensing observations, International Conference on Southern Hemisphere Meteorology and Oceanography (ICSHMO), 2006.
Fuelberg, H. E., Porter, M. J., Kiley, C. M., Halland, J. J., and Morse, D.: Meteorological conditions and anomalies during the Intercontinental Chemical Transport ExperimentNorth America, J. Geophys. Res., 112, D12S06, doi:10.1029/2006JD007734, 2007.
Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, J. Geophys. Res., 108(D24), 4756, doi:10.1029/2002JD003018, 2003.
Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y., Clals, P., Fan, S., Fung, I. Y., Gloor, M., Helmann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarle, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Tagucki, S., Takahashi, T., and Yuen, C.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, 2002.
Gurney, K. R., Seib, B., Ansley, W., Mendoza, D., Fischer, M., Miller, C., and Murtishaw, S.: The Vulcan Inventory, version 1.0, Purdue University, http://www.purdue.edu/eas/carbon/vulcan/research.html, 2008.
Hanan, N. P., Berry, J. A., Verma, S. B., Walter-Shea, E. A., Suyker, A. E., Burba, G. G., and Denning, A. S.: Model analyses of biosphere-atmosphere exchanges of CO_2, water and energy in Great Plains tallgrass prairie and wheat ecosystems, Agr. Forest Meteorol., 131, 162–179, 2004.
Kalman, R. E.: A new approach to linear filtering and prediction problems, Trans. ASME J. Basic Eng., 82, 35–45, 1960.
Kaminiski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmo-

spheric transport inversions, J. Geophys. Res., 106, 4703–4715, 2001.
Kawa III, S. R., D. J. E., Pawson, S., and Zhu, Z.: Global CO_2 transport simulations using meteorological data from the NASA data assimilation system, J. Geophys. Res., 109, D18312, doi:10.1029/2004JD004554, 2004.
Liang, Q., Jaeglé, L., Hudman, R. C., et al.: Summertime influence of Asian pollution in the free troposphere over North America, J. Geophys. Res., 112, D12S11, doi:10.1029/2006JD007919, 2007.
Lokupitiya, E., Denning, S., Paustian, K., Baker, I., Schaefer, K., Verma, S., Meyers, T., Bernacchi, C. J., Suyker, A., and Fischer, M.: Incorporation of crop phenology in Simple Biosphere Model (SiBcrop) to improve land-atmosphere carbon exchanges from croplands, Biogeosciences, 6, 969–986, 2009, http://www.biogeosciences.net/6/969/2009/.
McGrath-Spangler, E. L., Denning, A. S., Corbin, K. D., and Baker, I. T.: Implementation of a boundary layer heat flux parameterization into the Regional Atmospheric Modeling System (RAMS), Atmos. Chem. Phys. Discuss., 8, 14311–14346, 2008, http://www.atmos-chem-phys-discuss.net/8/14311/2008/.
Michalak, A. M., Bruhwiler, L., and Tans, P. P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, J. Geophys. Res. 109, 1–19, 2004.
Nicholls, M. E., Denning, A. S., Prihodko, L., Vidale, P., Davis, K., and Bakwin, P.: A multiple-scale simulation of variations in atmospheric carbon dioxide using a coupled biosphere-atmospheric model, J. Geophys. Res., 109, D18117, doi:10.1029/2003JD004482, 2004.
Olson, J. S., Watts, J. A., and Allison, L. J.: Major World Ecosystem Complexes Ranked by Carbon in Live Vegetation: A Database, ORNL/CDIAC-134, NDP-017. Carbon Dioxide Information Analysis Center, US Department of Energy, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, 2001.
Oren, R., Ellsworth, D. S., Johnsen, K. H., Phillips, N., Ewers, B. E., Maler, C., Schafer, K. V., McCarthy, H., Hendrey, G., McNulty, S. G., and Katul, G. G.: Soil fertility limits carbon seqestration by forest ecosystems in a CO_2-enriched atmosphere, Nature, 411, 469–471, 2001.
Parazoo, N.: Investigating synoptic variations in atmospheric CO_2 using continuousobservations and a global transport model, Master's thesis, Colorado State University, 2007.
Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective

on North American carbon dioxide exchange: CarbonTracker, Proceedings of the National Academy of Sciences of the United States of America 104, 18925–18930, 2007.
Prince, S.: NPP Cropland: Gridded Estimates for the Central, USA 1982–1996', online available at http://knb.ecoinformatics.org/knb/metacat?action=read&qformat=nceas&sessionid=&docid=nceas.184, 2000.
Peterson, G. A., Halvorson, A. D., Havlin, J. L., Jones, O. R., Lyon, D. J., and Tanaka, D. L.: Reduced tillage and increasing cropping intensity in the Great Plains conserves soil C, Soil Tillage Res., 47, 207–218, 1998.
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, 2005, http://www.atmos-chem-phys.net/5/3173/2005/.
Pielke, R. A., Cotton, W., Walko, R., Tremback, C., Lyons, W., Grasso, L., Nicholls, M., Moran, M., Wesley, D., Lee, T., and Copeland, J.: A comprehensive meteorological modeling system – RAMS, Meteorol. Atmos. Phys., 46, 69–91, 1992.
Raich, J. W., Rastetter, E. B., Melillo, J. M., Kicklighter, D. W., Steudler, P. A., Peterson, B. J., Grace, A. L., Moore III, B., and Vorosmarty, C. J.: Potential Net Primary Productivity in South America: Application of a Global Model, Ecol. Appl., 1(4), 399–429, 1991.
Robertson, G. P., Paul, E. A., and Harwood, R. R.: Greenhouse gases in intensive agriculture: contributions of individual gases to the radiative forcing of the atmosphere, Science, 289, 1922–1925, 2000.
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/.
Schimel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peylin, P., Braswell, B. H., Apps, M. J., Baker, D., Bondeau, A., Canadell, J., Churkina, G., Cramer, W., Denning, A. S., Field, C. B., Friedlingstein, P., Goodale, C., Heimann, M., Houghton, R. A., Melillo, J. M., B. Moore III, D. M., Noble, I., Pacala, S. W., Prentice, I. C., Raupach, M. R., Rayner, P. J., Scholes, R. J., Steffen, W. L., and Wirth, C.: Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems, Nature, 414, 169–172, 2001.
Schuh, A. E., Denning, A. S., Uliasz, M., and Corbin, K. D.: Seeing the forest through the trees: recovering large scale carbon flux biases in the midst of small scale variability, J. Geophys.

Res., 114, G03007, doi:10.1029/2008JG000842, 2009.
Sellers, P. J., Mintz, Y., Sud, Y. C., and Dalcher, A.: A simple biosphere model (SiB) for use within general circulation models, J. Atmos. Sci., 43, 505–531, 1986.
Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., Zhang, C., Collelo, G. D., and Bounoua, L.: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation, J. Climate, 9, 676–705, 1996.
Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.: Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2, Science, 316, 1732–1735, 2007.
Tilman, D., Reich, P., Phillips, H., Menton, M., Patel, A., Vos, E., Petersion, D., and Knops, J.: Fire suppression and ecosystem carbon storage, Ecology, 81(10), 2680–2685, 2000.
Tripoli, G. J. and Cotton, W. R.: The Colorado State University three-dimensional cloud/mesoscale model – Part I General theoretical framework and sensitivity experiments, J. Recherches Atmospheriques, 16, 185–219, 1982.
Ulliasz, M. and Pielke, R. A.: Application of the receptor oriented approach in mesoscale dispersion modeling', in: Air Pollution Modeling and its Applications VIII, edited by: Van Dop, H. and Steyn, D. G., Plenum Press, New York, 399–408, 1991.
Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., and Munger, J. W.: Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest, J. Geophys. Res.-Biogeosci., 112(G2), G02020, doi:10.1029/2006JG000293, 2007.
Wang, J. W., Denning, A. S., Lu, L., Baker, I. T., and Corbin, K. D.: Observations and simulations of synoptic, regional, and local variations in atmospheric CO_2, J. Geophys. Res., 112, D04108, doi:10.1029/2006JD007410, 2006.
Wise, T. A.: Identifying the Real Winners from, US Agricultural Policies, Global Development and Environmental Institute, Working Paper No. 05-07, 2005, 2006.
Zupanski, D., Denning, A. S., Marek, U., 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., 112, D17107, doi:10.1029/2006JD008371, 2007.

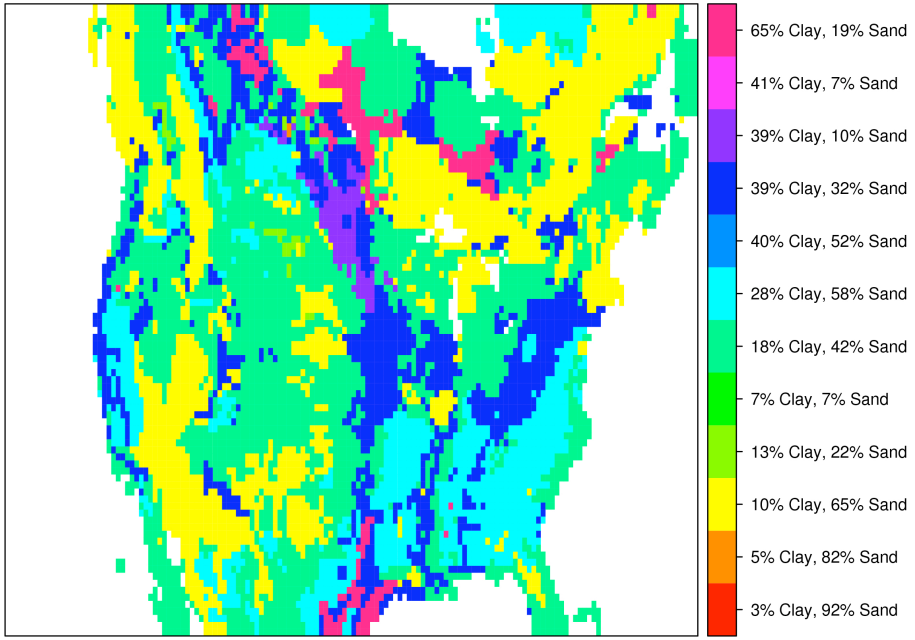

Fig. 1. Soil classes (IGBP) used for SiB3.

BGD
6, 10195–10241, 2009

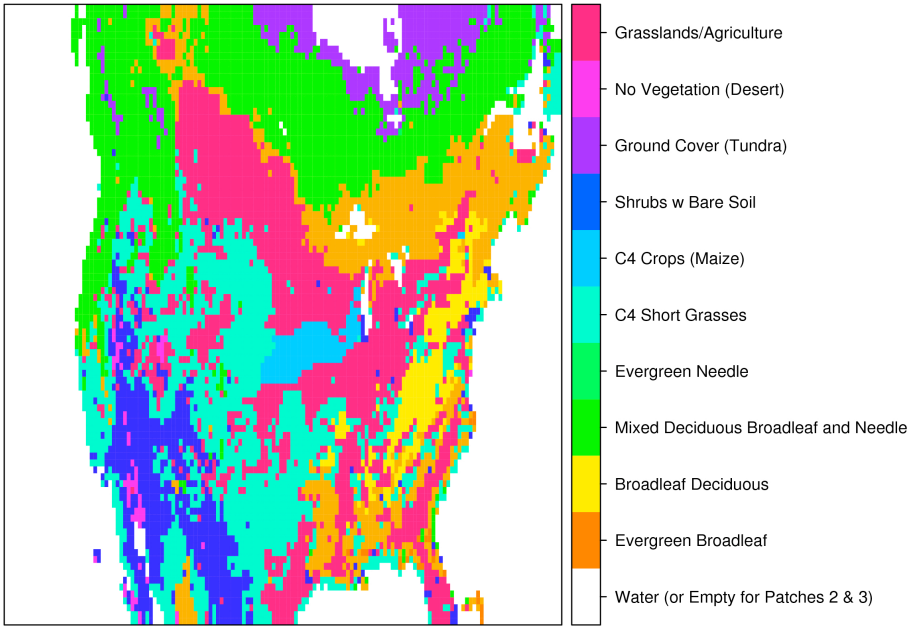

Fig. 2. Dominant SiB3 biome classes for the first biome patch derived from MODIS 12 Land-cover product.

BGD
6, 10195–10241, 2009

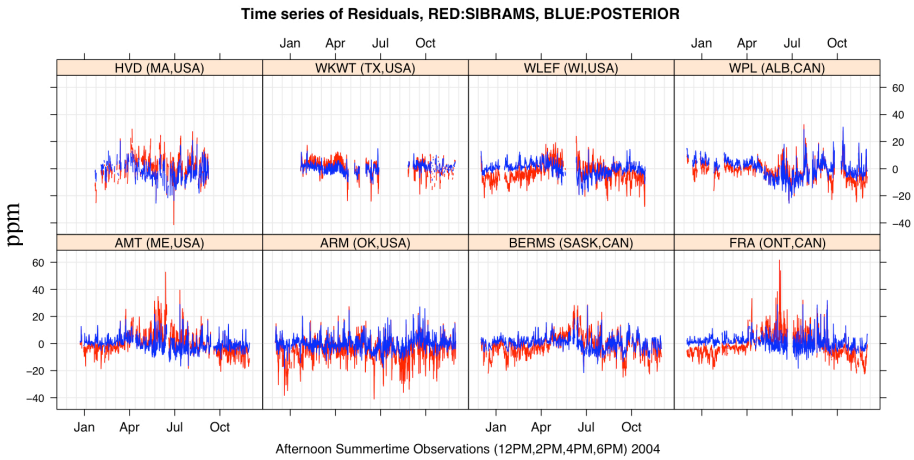

Fig. 3. Time series plots of carbon dioxide residuals based upon SiBRAMS prior (red) and inversion posterior (blue).

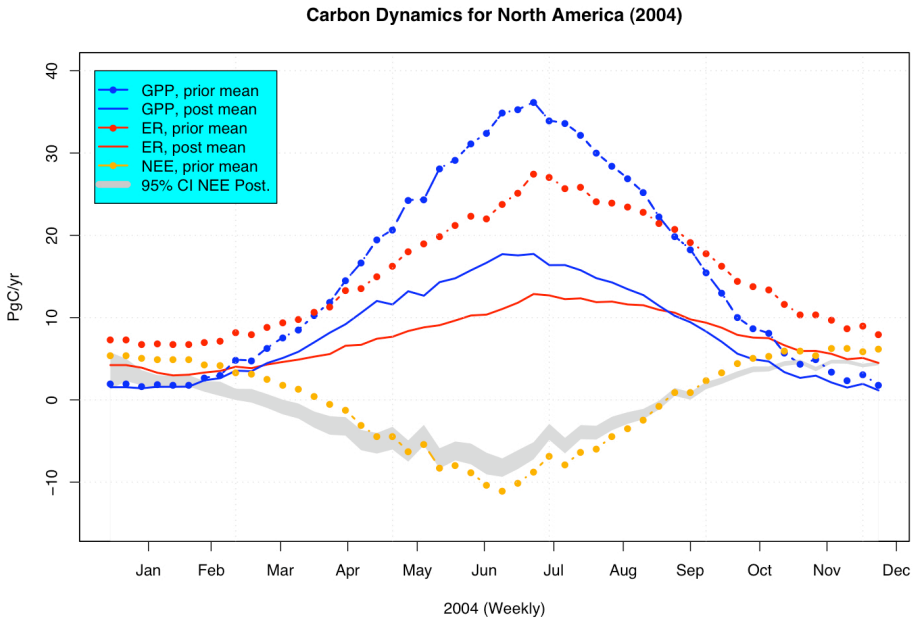

Fig. 4. Plots of prior and posterior estimates for GPP, ER, and NEE. Results are shown for a single inversion while the confidence intervals are derived from an ensemble of 100 inversions.

BGD
6, 10195–10241, 2009

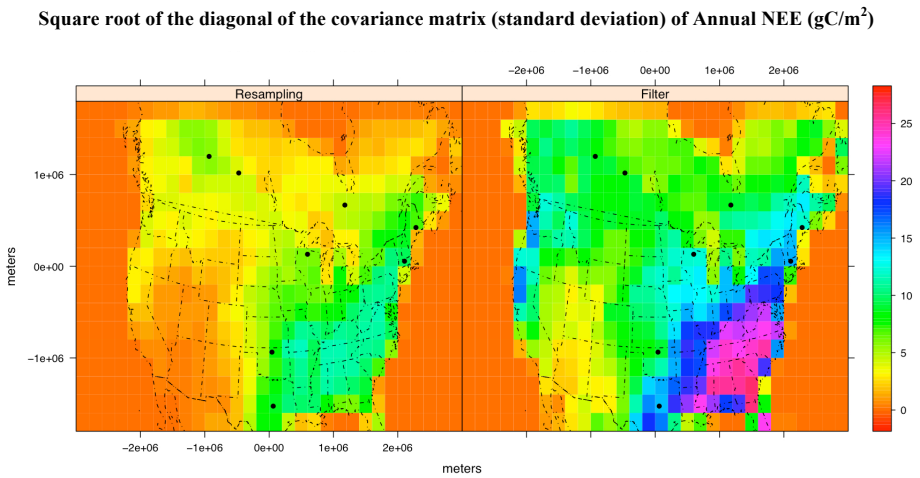

Fig. 5. Uncertainty in annual NEE. The left panel is the result of running 100 inversions each using a randomly selected 50% of the data and then calculating the variance of each cell's mean estimate, over the 100 inversions, and summing over each of the weekly filter cycles. Finally, the square root of this summed variance (standard deviation) is displayed and is a measure of the uncertainty of the mean estimate due to model over-fitting. For the right panel, the summed annual variance in NEE is calculated for each inversion, from the weekly filter estimates, and the the square root of this (standard deviation) is shown for each cell. These plots aim to provide a measure of the uncertainty of each cell's NEE estimate, incorporating the correlation between ER and GPP in each cell, but not incorporating the spatial correlation in the covariance matrices.

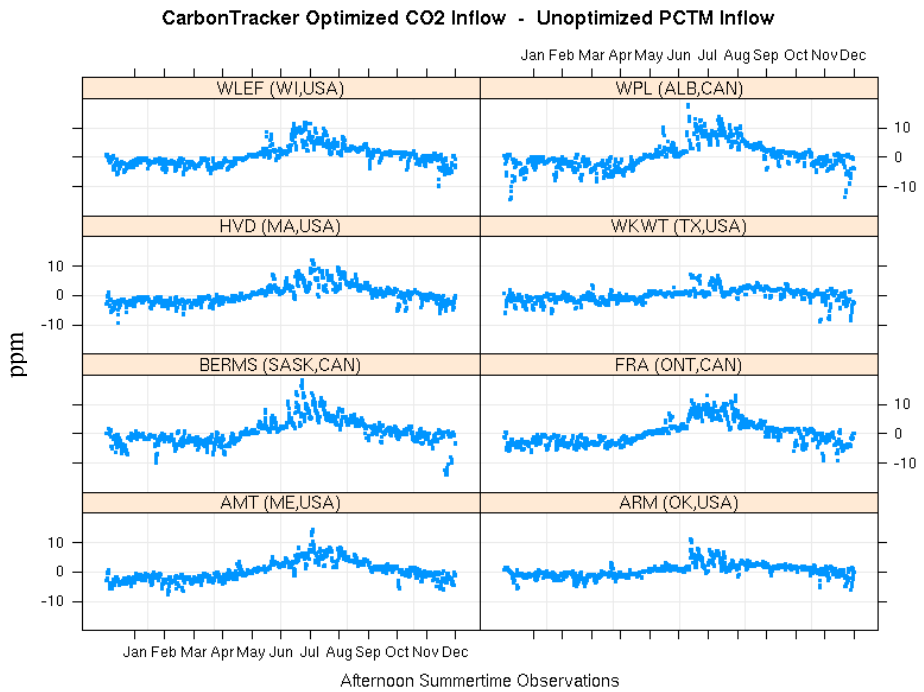

Fig. 6. Figure shows the effect of boundary inflow CO_2 upon tower CO_2 concentrations. In particular, this figure shows the "difference" between estimates of CO_2 arriving at tower due to two distinct boundary inflows (1420 sequential "12/2/4/6 p.m." observation sequences for each of 8 towers).

BGD
6, 10195–10241, 2009

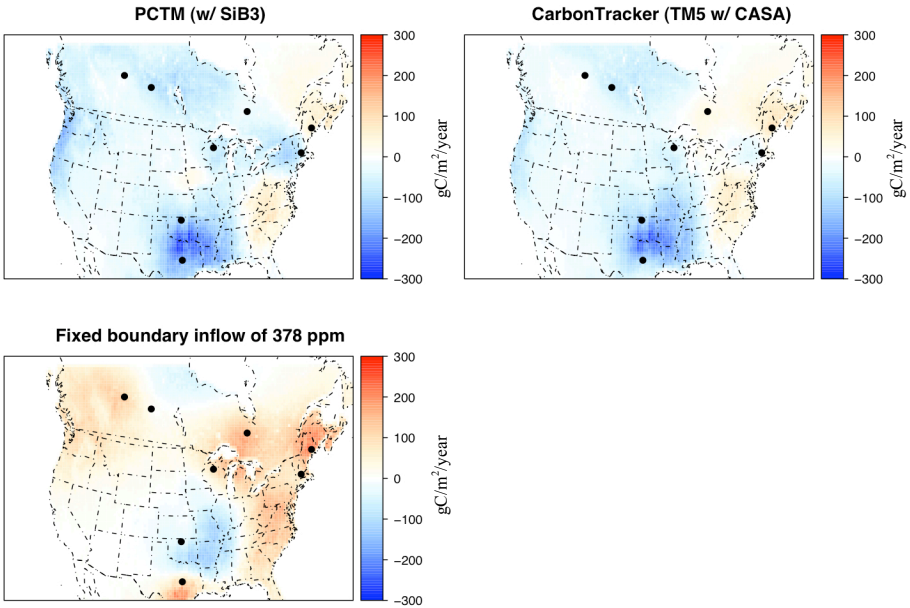

Fig. 7. Inversion estimates for three different inflow scenarios, one without modeled annual source/sink (PCTM w/SiB), one with modeled source/sink (CarbonTracker w/CASA), and a uniform fixed 378 ppm inflow. Negative values denote land uptake of carbon. Summed annual NEE is −0.65, −0.48, and 0.38 PgC/yr, respectively.

BGD
6, 10195–10241, 2009

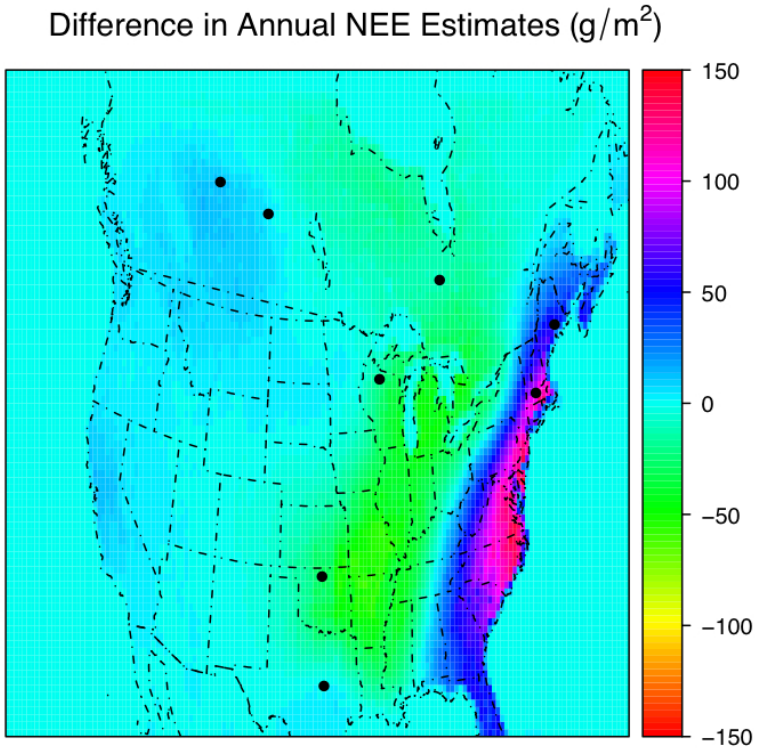

Fig. 8. Difference in annual sink inferred by inversions based upon the Vulcan fossil fuel inventory and the Andres et al. (1995) fossil fuel inventory. Positive values (purple) indicate carbon sinks were stronger using Andres inventory. Spatially-summed annual difference between Vulcan-based NEE estimate for 2004 and Andres (1995) based NEE estimate for 2004 is less than 0.01 PgC.

BGD
6, 10195–10241, 2009

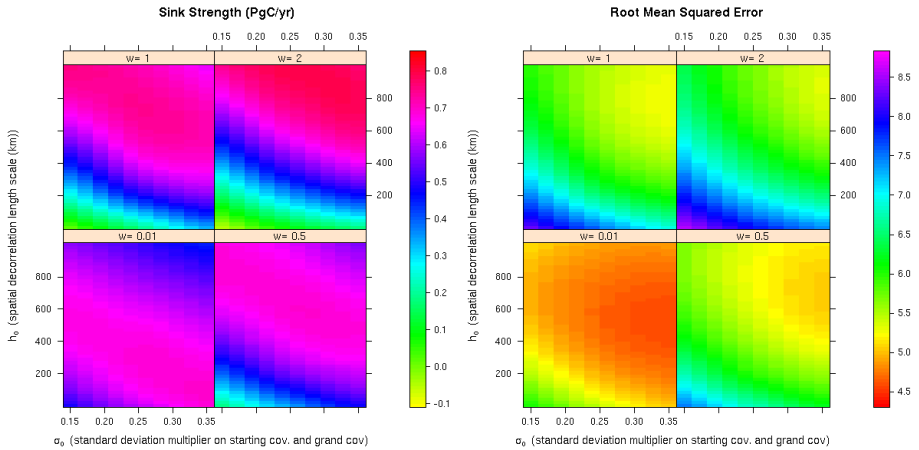

Fig. 9. Sensitivity of **(a)** sink estimate and **(b)** root mean squared error to varying covariance parameters in inversion. For example, a set of parameter values like w=0, h_0=600km, σ_0=0.35 provides the estimate w/ the lowest RMSE **(b)** and an estimated sink of approximately 0.55 PgC/yr **(a)**. Nevertheless, qualitatively, this represents a maximum departure from the prior and thus must be viewed with some skepticism due to the likelihood of overfitting the data.

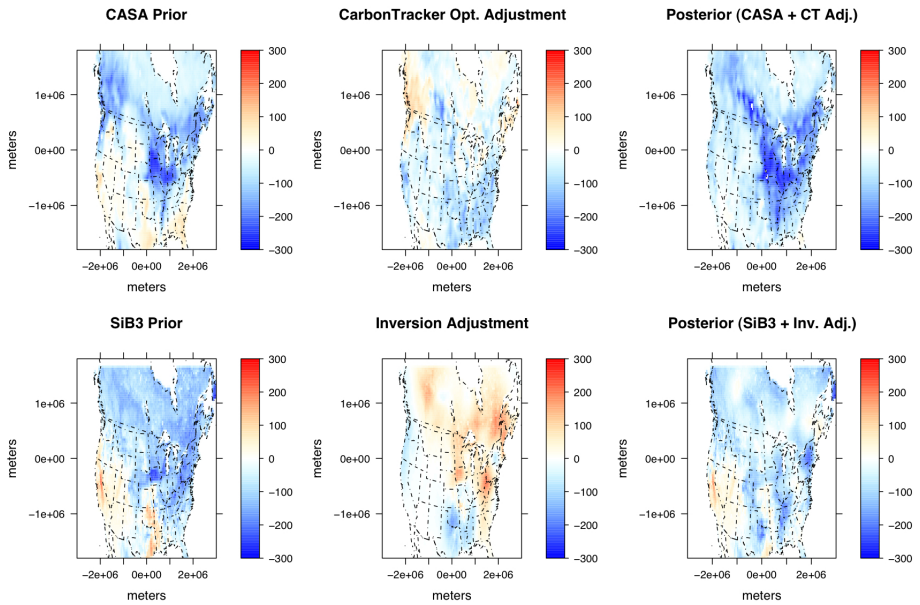

Fig. 10. July-August-September comparison. Top panels concern CarbonTracker and lower panels concern our inversion. Left panels show a priori NEE, middle panels show inversion adjustment, and right panels show a posteriori NEE.

BGD
6, 10195–10241, 2009

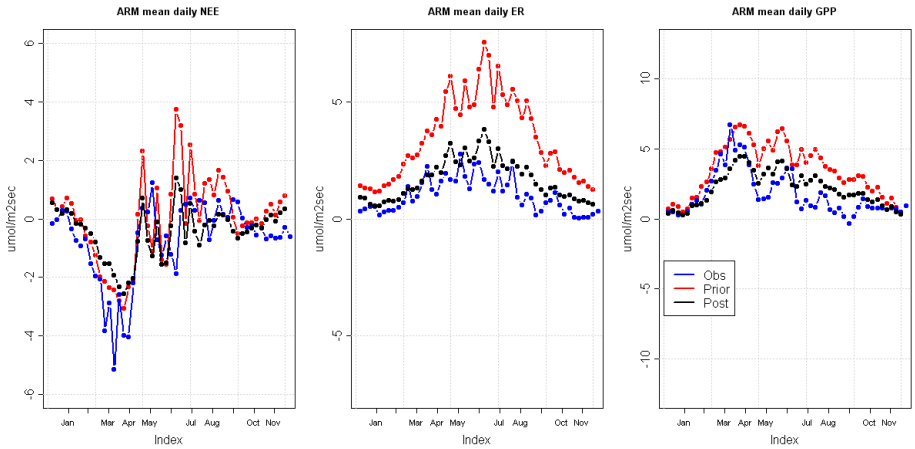

Fig. 11. Comparison of posterior fluxes of GPP, ER, and NEE with Ameriflux Level 4 flux tower data for ARM Site in OK. Pay particular attention to the fact that the *y*-axes are different scales.

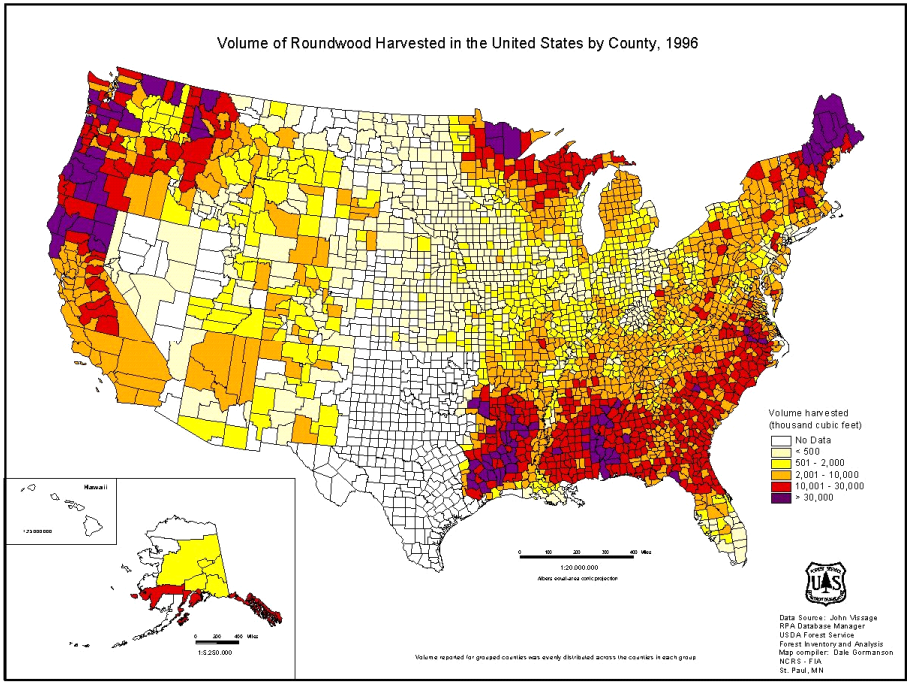

Fig. 12. Timber harvest plot from National Report on Forest Resources (1996).