



## Seeing the forest through the trees: Recovering large-scale carbon flux biases in the midst of small-scale variability

A. E. Schuh,<sup>1</sup> A. S. Denning,<sup>1</sup> M. Uliasz,<sup>1</sup> and K. D. Corbin<sup>1</sup>

Received 8 August 2008; revised 31 March 2009; accepted 5 May 2009; published 15 August 2009.

[1] This paper investigates the effect of fine-scale spatial variability in carbon fluxes upon regional carbon flux inversion estimates in North America using simulated data from 1 May through 31 August 2004 and a hypothetical sparse network of eight towers in North America. A suite of random smooth regional carbon flux patterns are created and then obscured with random fine-scale spatial flux “noise” to mimic the effect of fine-scale heterogeneity in carbon fluxes found in nature. Five hundred and forty grid-scale atmospheric inversions are run using the synthetic data. We find that, regardless of the particular fine spatial scale carbon fluxes used (noise), the inversions can improve a priori carbon flux estimates significantly by capturing the large-scale regional flux patterns. We also find significant improvement in the root-mean-square error of the model are possible across a wide range of spatial decorrelation length scales. Errors associated with the inversion decrease as estimates are sought for larger and larger areas. Results show dramatic differences between postaggregated fine-scale inversion results and preaggregated coarse-scale inversion results confirming recent warnings about the “preaggregation” of inversion regions.

**Citation:** Schuh, A. E., A. S. Denning, M. Uliasz, and K. D. Corbin (2009), 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.

### 1. Introduction

[2] During the last decade, Bayesian-based atmospheric inversion techniques have emerged as a viable tool to investigate the spatiotemporal pattern of terrestrial carbon fluxes [Enting *et al.*, 1994; Fan *et al.*, 1998; Gurney *et al.*, 2002; Rodenbeck *et al.*, 2003; Baker *et al.*, 2006; Peters *et al.*, 2007]. Research has largely been focused on continental-sized regions of the earth, using coupled general circulation models (GCM). Lately, researchers have begun applying these techniques to regional flux domains with increasingly finer resolution inversion domains [Gerbig *et al.*, 2003; Carouge *et al.*, 2008a, 2008b; Lauvaux *et al.*, 2008].

[3] In general, regional scale inversions focusing on temporal biases that are of a seasonal length, or longer, are possible because biosphere models have become adept at capturing the majority of carbon exchange that occurs on diurnal and seasonal time scales. The effects of the temperature, available soil water, and sunlight have been modeled extensively and predictions have become reasonably accurate over a variety of conditions and scales [Baker *et al.*, 2003, Hanan *et al.*, 2005; Vidale and Stockli, 2005]. However, the necessary components to model longer-term processes such as nitrogen deposition, land management, and other biogeochemical dynamics are often missing from these advanced biophysical models and thus lead to errors

in the model. These effects may be unrecognizable at the diurnal scale but may dominate over longer temporal scales. Thus, researchers can begin to estimate these unknown processes by effectively removing the high-frequency diurnal signals at fine scales and estimating the residuals over longer time and space scales.

[4] The biggest hurdle to these inversions is insufficient carbon dioxide concentration data to constrain the flux inversion problem. Therefore, various additional constraints must be added. Two major methodologies have been employed to deal with this problem. The first of these two methods, which was employed in many inversion papers [Enting *et al.*, 1994; Fan *et al.*, 1998; Gurney *et al.*, 2002; Peters *et al.*, 2007] involved the preaggregation of large flux regions, generally according to prior guesses of flux patterns based upon global spatial net primary production (NPP) estimates. Largely in response to criticisms of this method [Kaminski *et al.*, 2001; Engelen *et al.*, 2002], geostatistical techniques were employed [Michalak *et al.*, 2004] to constrain the inversion problem. Michalak *et al.* [2004] used maximum likelihood techniques to estimate spatial covariance parameters (of the carbon flux error component) and then applied the resulting smooth covariance matrices to the differences between the underlying fluxes and the a priori fluxes. As a consequence of these additional constraints, inversion resolutions could be used that were much closer to that of the underlying forward transport and carbon flux models. Zupanski *et al.* [2007] used techniques similar to Michalak *et al.* [2004], with the exception that they used a maximum likelihood ensemble filter (MLEF) to track the covariance structure dynamically

<sup>1</sup>Graduate Degree Program in Ecology, Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA.

instead of using more traditional geostatistical point-based estimates of spatial covariance parameters. *Peylin et al.* [2005] explored the effect of two different error correlation length-scale assumptions when estimating daily fluxes over a large portion of Europe. *Carouge et al.* [2008a, 2008b] used a 10-tower network in Europe in 2001, combined with synthetic data, to explore the sensitivity of inversion-based net ecosystem CO<sub>2</sub> exchange (NEE) estimates to various parameters of the inversion including temporal and spatial correlation.

[5] It seems reasonable to hypothesize that large-scale spatial patterns may exist in the errors for many models. For example, assume that one is modeling a large continental region such as North America. If the underlying flux model consistently underpredicts gross primary productivity (GPP) for forested regions and overpredicts for grassland regions over a given time interval such as a day or a year, then a map of the errors will likely show small positive errors in GPP over the grasslands and larger negative errors over the forested regions. Since grasslands and forested regions tend to exist in “clumps” on larger scales, this has the effect of inducing a spatially correlated structure to the errors. However, large-scale biases need not exist simply as a function of vegetation type. Persistent long-term droughts might affect large spatially connected regions of the continent over several different vegetation types. Fertilization effects from nitrogen deposition might also impact NEE over broad regions containing many vegetation types. It is difficult to exactly predict the structure in any of these cases, but it is reasonable to believe that correlations might exist on the order of several hundred kilometers or more. It is important to realize that this does not imply that the structure will be simple to recover. For instance, along ecotones such as the transition from the western to eastern slope of the Rocky Mountains and into the Great Plains of the central United States, one might not expect errors in fluxes to be strongly correlated. It is also reasonable to assume that the covariance function may not simply be a function of distance and may involve some kind of structure around covariates such as biome classification.

[6] Small-scale spatial variability has been a recurrent theme of eddy flux measurements. For instance, data from the Chequamegon Ecosystem Atmospheric Study (<http://cheas.psu.edu>) showed significant variability in annual NEE between mature hardwood forests and old growth hardwood forests [*Desai et al.*, 2005]. Disturbance histories and the associated age structure has also been shown to be important to carbon dynamics in ponderosa pines of the Western United States [*Thornton et al.*, 2002; *Law et al.*, 2003]. Important factors explored in these papers, such as stand age and land management, are generally only coarsely modeled, or not modeled at all in larger-scale inversion studies. Of course the sampling footprints of the towers that generate these estimates of variability are generally on the order of a square kilometer or two and thus aggregated flux results at, for instance, 1600 km<sup>2</sup> (40 km by 40 km) might be expected to show less variability than that because of the averaging effect of aggregation. Regional inversions provide corrections to a priori NEE estimates and these corrections exhibit features on much larger scales than 40 km [*Gerbig et al.*, 2003; *Peylin et al.*, 2005]. The effect this has on fluxes is to introduce a layer of “noise” relative

to potentially larger spatial scale error signals, such as continental scale sinks or large-scale agricultural expansion.

[7] Suppose that the flux model providing the prior estimates underpredicts GPP, on average, for a large forested area of North America. It is reasonable that this bias would vary spatially over this area on fine scales as a function of local land management practices, natural fire regimes, climate, and anthropogenic fertilization effects. These types of effects have different magnitudes and can be persistent at different temporal scales. Small-scale spatial variability has not typically been included as part of the prior error covariance structure [*Michalak et al.*, 2004; *Peylin et al.*, 2005; *Peters et al.*, 2005, 2007; *Zupanski et al.*, 2007], where it would be represented by an independent variance component that is typically termed the “nugget” in geostatistical literature [*Cressie*, 1993]. In general, it is unclear how the existence and/or exclusion of this error term in the inversion will affect inversion results.

[8] For instance, assume one is tasked with building and maintaining several towers to collect CO<sub>2</sub> observations which will be used to provide regional scale NEE estimates for a reasonably large managed forest region of North America (the Pacific Northwest United States for instance). Upon getting into the field, the researcher sees that the land is a patchwork of old growth, new growth, and recently clear-cut forest areas, essentially a myriad of fine scale ecosystems. Where might one locate their tower? If one puts their tower in a clear-cut location, will it “bias” his observations? Or what about putting it in an old growth stand surrounded by very vigorous young tree stands? Will the location have an effect and what will it be? These are obviously very important questions considering the work and cost involved in obtaining carbon dioxide measurements. We certainly know that the precise location of an eddy covariance tower has a huge effect upon the NEE measurements and any inferences that a researcher might want to make from them. The question is then: is the precise location important for a tower providing CO<sub>2</sub> observations to an atmospheric inversion?

[9] In this paper we investigate the effect of fine-scale spatial variability upon large spatial scale improvements in estimated NEE and use synthetic data and experiments to show that regional inversions are robust to fine-scale spatially independent variance in the flux errors. These inversions are performed in a manner in which assumptions need not be made about a fixed “pattern” of fluxes across large regions. In particular, we vary both the level of small-scale-independent variance (noise) as well as the decorrelation length scale of the spatially correlated portion of the bias which has a covarying effect upon the success of the inversion. A hypothetical sparse network of 8 towers in North America is used and the effects of varying these two quantities are tested using simulated fluxes and corresponding simulated measurements from a biosphere-meteorological model.

## 2. Methods

### 2.1. Model

[10] 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.*, 1996; Denning *et al.*, 1996]. The parameterization of photosynthetic carbon assimilation is based on enzyme kinetics originally developed by Farquhar *et al.* [1980], and is linked to stomatal conductance and hence to the surface energy budget and atmosphere [Collatz *et al.*, 1991, 1992; Sellers *et al.*, 1996; Randall *et al.*, 1996]. The model has been updated to include prognostic calculation of temperature, moisture, and trace gases in the canopy air space, and the model has been evaluated against eddy covariance measurements at a number of sites [Baker *et al.*, 2003; Hanan *et al.*, 2005; Vidale and Stockli, 2005]. SiB has been coupled to the Regional Atmospheric Modeling System (RAMS) and used to study PBL-scale interactions among carbon fluxes, turbulence, and CO<sub>2</sub> mixing ratio [Denning *et al.*, 2003] and regional scale controls on CO<sub>2</sub> variations [Nicholls *et al.*, 2004; Wang *et al.*, 2006]. Other recent improvements include biogeochemical fractionation and recycling of stable carbon isotopes [Suits *et al.*, 2005], improved treatment of soil hydrology and thermodynamics, and the introduction of a multilayer snow model based on the Community Land Model [Dai *et al.*, 2003]. This latest version of SiB is termed SiB3.

[11] In SiB3, the net ecosystem exchange (NEE) is composed of two component fluxes, gross primary productivity (GPP) and ecosystem respiration (RESP), which includes autotrophic and heterotrophic respiration terms where  $x$  and  $y$  represent grid coordinates and  $t$  represents time

$$\text{NEE}(x, y, t) = \text{RESP}(x, y, t) - \text{GPP}(x, y, t). \quad (1)$$

[12] High-frequency time variations of photosynthesis and respiration are assumed to be well understood and easily modeled processes, i.e., because of changes in radiation, temperature, soil moisture, etc. Long-term, more persistent biases are estimated (equation (2)) by solving for unknown multiplicative biases in each component flux after smoothing in space and time. This is accomplished by convolving the observation-specific “influence” functions generated from a Lagrangian particle dispersion model, LPDM [Uliasz and Pielke, 1991; Zupanski *et al.*, 2007; Lauvaux *et al.*, 2008], with GPP and RESP at each time step in SiB-RAMS. Figure 1 shows examples of daily mean influence functions for the WLEF tower for ecosystem respiration. One can see that the influence function is weaker for Figure 1 (top), 10 May, mainly because of faster transport from the northwest as well as weaker carbon fluxes due to late spring/early summer conditions in the northern regions of North America.

[13] To summarize, we estimate 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:

$$\begin{aligned} \text{NEE}(x, y, t) = & (1 + \beta_{\text{RESP}}(x, y))\text{RESP}(x, y, t) \\ & - (1 + \beta_{\text{GPP}}(x, y))\text{GPP}(x, y, t). \end{aligned} \quad (2)$$

The model domain, shown in Figure 1, consists of most of the United States as well as a large portion of Canada and

the northern portions of Mexico. SiB3-RAMS was run on a single  $150 \times 90$  grid of 40 km cells. RAMS meteorology was nudged with NCEP ETA 40 km analysis data throughout the domain using the 4DDA scheme to produce more reliable wind fields. The fine-scale RAMS output was then used to drive the backward in time LPDM model. SiB3 was run with 8-day fractional photosynthetically available radiation (FPAR) and leaf area index (LAI) fields derived from the MODIS MOD15 product. This was provided from the Numerical Terradynamics Simulation Group at the University of Montana who generated it for use in constructing the official MOD17 GPP product [Mu *et al.*, 2007]. The focus of this study was on the regional domain and therefore boundary inflow of CO<sub>2</sub> was not optimized or investigated. Given the simulated nature of the experiments, no actual estimate of inflow was needed. An inversion of North America using real data could follow a nested coarse-inversion concept, similar to that presented by Peylin *et al.* [2005].

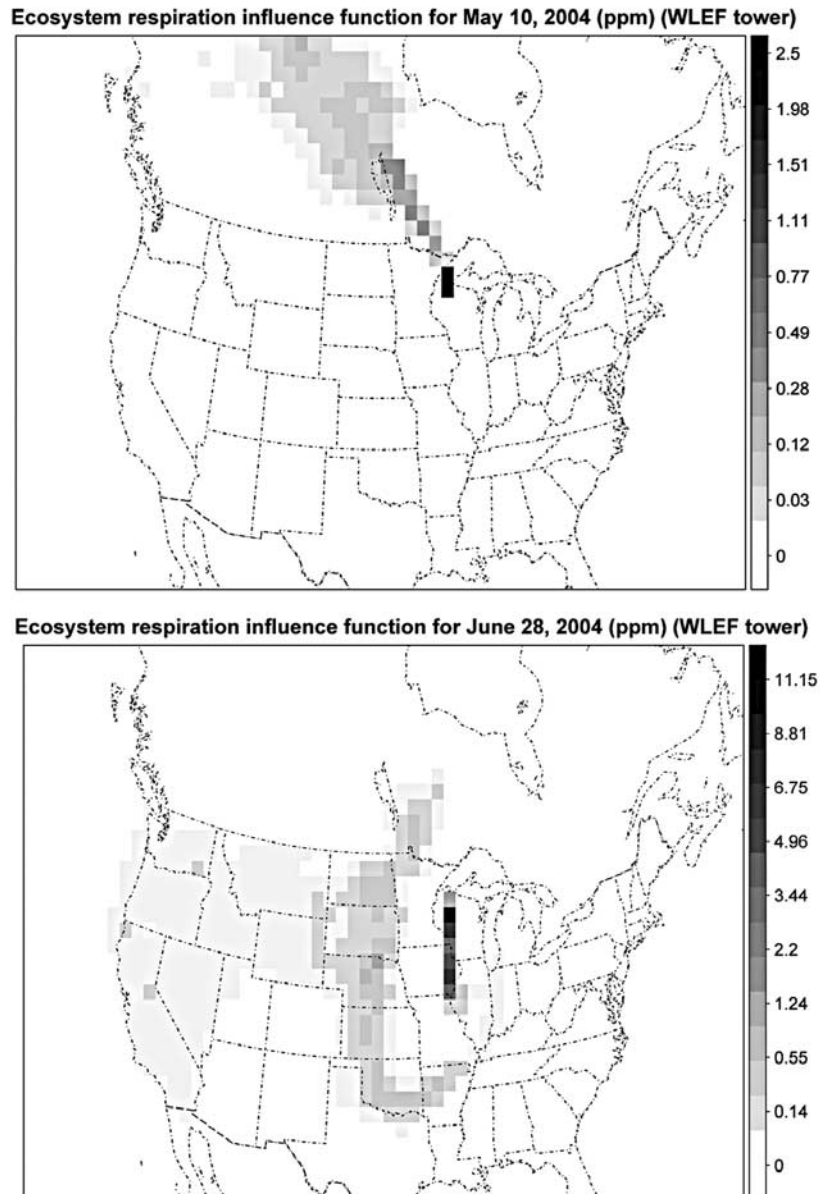
## 2.2. Synthetic Data

[14] CO<sub>2</sub> mixing ratio observations are simulated hourly at eight measuring sites (WLEF, Harvard Forest, ARM, BERMS, Fraserdale, Western Peatland, WKWT, and Argyle (ME), see Figure 2 for locations) over a 113-day period. These were produced by first running SiB for the period and domain of interest to serve as our a priori biosphere flux model. Then we convolved simulated flux bias fields for GPP and RESP, shown as coefficients to RESP and GPP in equation (2), with LPDM derived influence functions representing contributions to an observation from upwind flux areas. 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<sub>2</sub>. Afternoon hourly average observations, at 1200, 1400, 1600, and 1800 LT, are used to lessen the impact of low-quality modeling of transport during times of extremely stable and stratified nocturnal atmospheric conditions near the ground. In total, there are 3616 synthetic observations covering the period 1 May to 20 August 2004. An independent mean zero 2 ppm standard deviation Gaussian error term is added to the CO<sub>2</sub> observations to provide a crude estimate of transport errors.

[15] In summary, we used a continental scale model run of SiB, based upon a 113-day period in the summer of 2004, to provide realistic GPP and respiration fluxes. We also used a model run of RAMS during the same period to provide transport fields. We then assume ‘truth’ is actually represented by these biosphere fluxes multiplied by some synthetic, simulated, bias fields (as shown in Figure 2). We then simulated what the carbon dioxide concentrations would be at the observing towers give these biases. Finally, we performed the inversion to see how well we can estimate the biases from the carbon dioxide concentration observations.

## 2.3. Inversion Procedure

[16] Standard multivariate normal assumptions are made and data are assimilated using a Bayesian synthesis inversion, or equivalently, a single standard Kalman filter updating step. The resolution of the inversion domain ( $36 \times 60$ , 100 km grid spacing) and the number of measurements (3616) were selected such that the needed matrix inversions

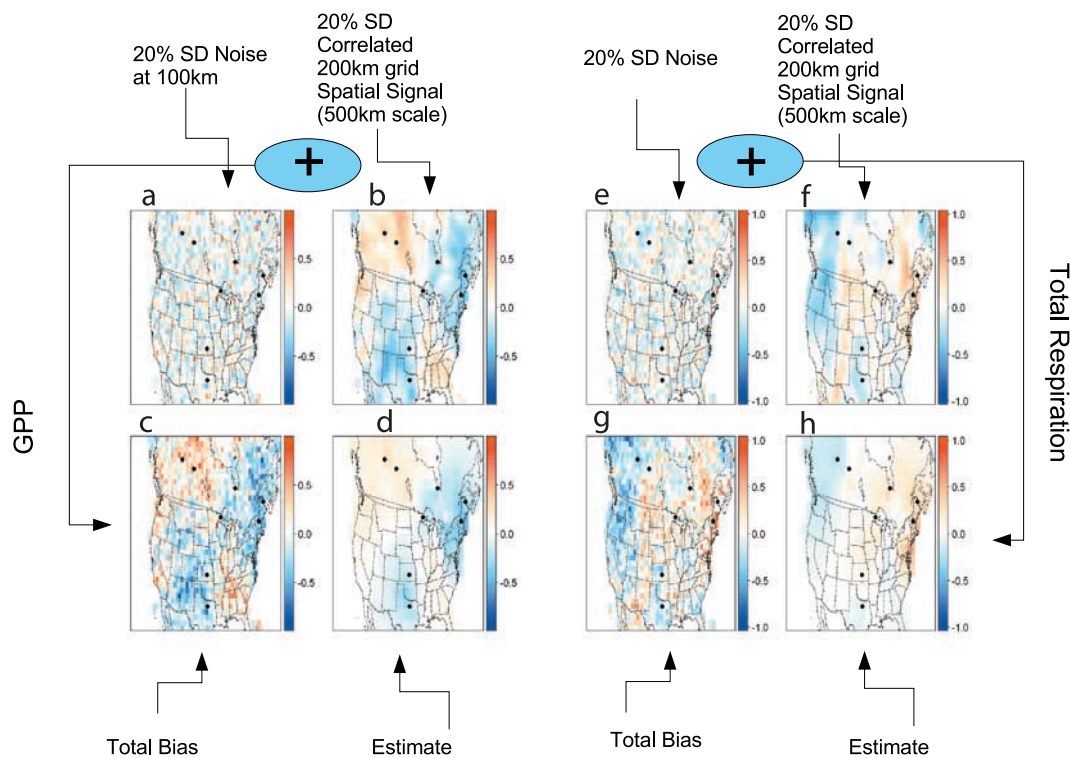


**Figure 1.** (top) An influence function for the modeled upwind respiration contributions to the WLEF tower for the mean of observations recorded at 1200, 1400, 1600, and 1800 LT on 20 May 2004. For example, if a cell value is 0.29 ppm then it contributed 0.29 ppm to the mean afternoon observation (estimated by the four hourly observations above) at the WLEF tower for that day. (bottom) A similarly derived respiration influence function for 28 June 2004.

could be calculated relatively quickly and without the aid of additional covariance subsampling procedures such as the Ensemble Kalman Filter methods [Evensen, 1994; Zupanski *et al.*, 2007] employ. While sufficient for theoretical exercises, it is noted that additional measurements and increased inversion domain resolution would require more involved subsampling procedures such as those used in the ensemble methods as well as a filter mechanism to propagate information forward. In particular, for a length  $n$  CO<sub>2</sub> measurement vector  $y$ , length  $m$  CO<sub>2</sub> flux bias vector  $\beta$ ,  $n \times n$

observation error covariance matrix  $\Sigma$ ,  $n \times m$  Jacobian transport matrix  $G$ , length  $m$  prior flux estimate  $\beta_0$ , and  $m \times m$  model prior mismatch covariance matrix  $\Sigma_0$ , the Bayesian statistical assumptions are  $N(\mu, \Sigma)$  represents a multivariate Gaussian/Normal distribution with mean vector  $\mu$  and covariance matrix  $\Sigma$

$$y|\beta, \Sigma \sim N(G\beta, \Sigma) \beta \sim N(\beta_0, \Sigma_0). \quad (3)$$



**Figure 2.** Example correction of (a–d) GPP and (e–h) total respiration signal. (a and e) Small spatial scale bias applied over model domain, (b and f) large-scale bias over model domain which we would like to recover, (c and g) total signal (sum of small and large), and (d and h) posterior estimate of mean bias. The eight CO<sub>2</sub> observing towers are shown as black dots on the images. Since biases are plotted as deviations from the prior (as in equation (2)), the mean a priori fluxes can be visualized as solid white plots with all grid cells equal to zero.

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

$$p(\beta|y, \Sigma) \propto -\frac{1}{2} \left[ (G\beta - y)^T \Sigma^{-1} (G\beta - y) + (\beta - \beta_0)^T \Sigma_0^{-1} (\beta - \beta_0) \right] \\ \sim N \left( (\Sigma_0^{-1} + G^T \Sigma^{-1} G)^{-1} (\Sigma_0^{-1} \beta_0 + G^T \Sigma^{-1} y), ((\Sigma_0^{-1} + G^T \Sigma^{-1} G))^{-1} \right). \quad (4)$$

With a little bit of algebra, one can rewrite the mean/expectation of the posterior distribution of the mean, giving the familiar Kalman filter updating equation

$$E[\beta_{\text{posterior}}] = \beta_0 + (G^T \Sigma^{-1} G + \Sigma_0^{-1})^{-1} G^T \Sigma^{-1} (y - G\beta_0). \quad (5)$$

With respect to constraining the problem with spatially correlated errors, the covariance matrix  $\Sigma_0$  is partitioned into RESP and GPP components,  $\Sigma_{\text{RESP,prior}}$  and  $\Sigma_{\text{GPP,prior}}$ , and will take on the following form:

$$\Sigma_0 = \begin{bmatrix} \Sigma_{\text{RESP,prior}} & 0 \\ 0 & \Sigma_{\text{GPP,prior}} \end{bmatrix}. \quad (6)$$

For the case of correlated errors in the prior flux, the respiration and GPP covariance matrices are each formed

from the exponential covariance function, where  $t_{i,j}$  is the distance between points  $x_i$  and  $x_j$

$$\text{Cov}(\beta_i, \beta_j) = \begin{cases} \sigma_0^2 (1 - \alpha_0) \exp\left(\frac{-t_{i,j}}{h_0}\right), & i \neq j \\ \alpha_0 \sigma_0^2, & i = j. \end{cases} \quad (7)$$

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  $\sigma^2$  parameter is the scalar variance parameter and determines the variance of the marginal distribution of the particular flux component. The parameter  $\alpha_0$  controls what percentage of the covariance can be attributed to spatial covariance, as opposed to spatially independent errors.

[17] Given a posterior mean NEE  $\mathbf{x}_{\text{posterior}}$  of length  $n$ , a posterior mean NEE variance estimate  $\Sigma_{\text{posterior}}$  of dimension  $n \times n$ , and a scalar vector  $\mathbf{b}$  of length  $n$  that maps higher-resolution fluxes to coarser resolution fluxes, the following result from multivariate Gaussian statistics

**Table 1.** Summary Statistics<sup>a</sup>

Flux-Based Statistics	200 km	400 km	600 km	1200 km	Domain
Prior mean RMSE (g/m <sup>2</sup> )	45.3	38.4	35.3	26.8	2.6
Posterior mean RMSE (g/m <sup>2</sup> )	28.2	20.5	16.4	8.0	1.7
Improvement over prior <sup>b</sup> (%)	39.2	49.1	56	72.3	57.1
Improvement in mean SD for grid cell mean over prior <sup>c</sup> (%)	32.5	40	45.8	59.1	77.9

<sup>a</sup>For example, inversion shown in Figure 1. This result is based upon a simulated observation error and thus changes slightly with different realizations.

<sup>b</sup>Presented in Figure 4 for multiple inversion study.

<sup>c</sup>Crude measure of tightening of posterior.

[Johnson and Wichern, 1988] can be employed to compare mean NEE at larger postaggregated scales

$$\overline{NEE}_b = b'x_{\text{posterior}} \sim N\left(b'x_{\text{posterior}}, b' \sum_{\text{posterior}} b\right). \quad (8)$$

The scalar vector  $b$  can be chosen as a sequence of  $1/k$ 's and 0s where one is estimating the mean of a block of  $k$  cells together. In essence, this is mapping the higher-resolution posterior mean fluxes to coarser resolution mean fluxes. Given that we are considering NEE as the sum of GPP and RESP, the above result can first be employed to sum GPP and RESP correctly and then employed again to aggregate up resulting NEE. In this example, our finest resolution was 100 km, a grid of 60 by 36. Values of  $k$  were chosen to be 4, 9, 16, 36, 144, and 2160, which represent aggregations to 400 km, 900 km, 1600 km, 3600 km, 14,400 km, and the entire domain. In order to compare to the prior, this

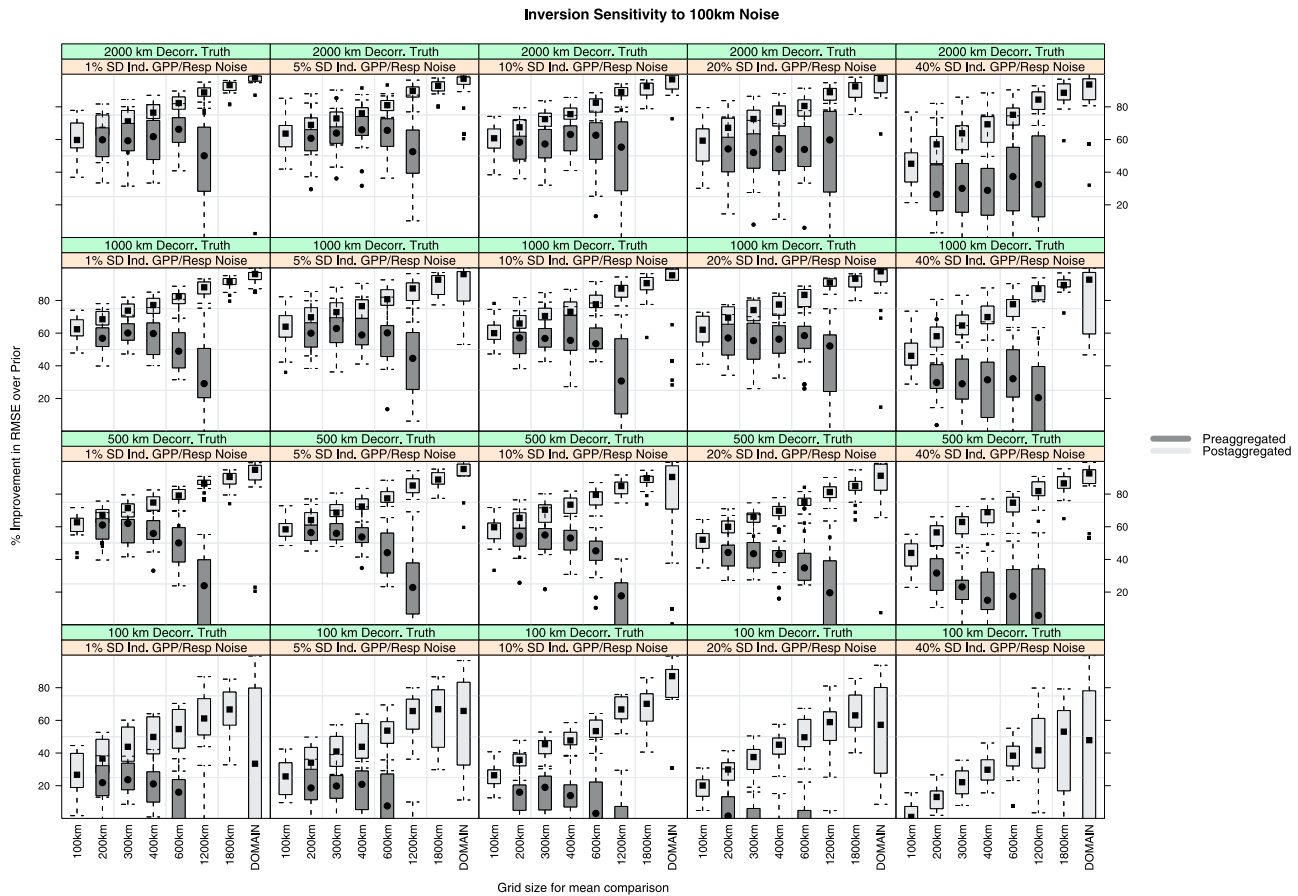
calculation was performed on both the distribution of the mean of the posterior fluxes as well as the assumed distribution of the mean of the prior fluxes.

## 2.4. Experiments

[18] In order to test the sensitivity of the inversion to fine-scale spatial noise, we introduce a set of Monte Carlo inversion experiments. Recall that the principle motivation of this paper is to investigate the recovery, or estimation, of large-scale flux patterns through a “veil” of small-scale spatial noise. Given the difficulty in estimating decorrelation length scale from the data and the uncertainty surrounding the effect of one's choice of prior decorrelation scale length for the flux errors, we choose to include it in the sensitivity tests. In traditional Bayesian statistics, one is working with ‘fixed’ observations and so one typically perturbs a priori distributions to test the sensitivity of the estimation procedures to them. We take a different approach

- Loop over (100 km, 500 km, 1000 km, 2000 km) spatial decorrelation length scale
- Loop over (1%, 5%, 10%, 20%, 40%) standard deviation of Gaussian fine scale flux bias
  - Loop over 1, ..., 18 (18 random samples)
    1. generate random “observation” error vector (2 ppm standard deviation.) at 100 km
    2. generate random large scale spatial patterns for GPP/RESP flux bias using spatial decorrelation length scale selected in first loop and 20% standard deviation multiplier
    3. generate random small scale spatial pattern for flux bias based upon standard deviation selected in second loop
    4. create synthetic data by combining transport and fluxes (at 100 km) and adding “observation” error vector
    5. run inversion at 100 km with a priori decorrelation length scale set to 500 km and a priori scalar standard deviation set to 20%
    6. statistically aggregate up results to 100 km, 200 km, 300 km, 400 km, 600 km, 1200 km, 1800 km, entire domain
    7. compare results from (6) with ‘truth’ given by (2) + (3) using RMSE of NEE estimate (RESP-GPP)

**Figure 3.** A summary of the algorithm used to generate postaggregated experiments in section 3.



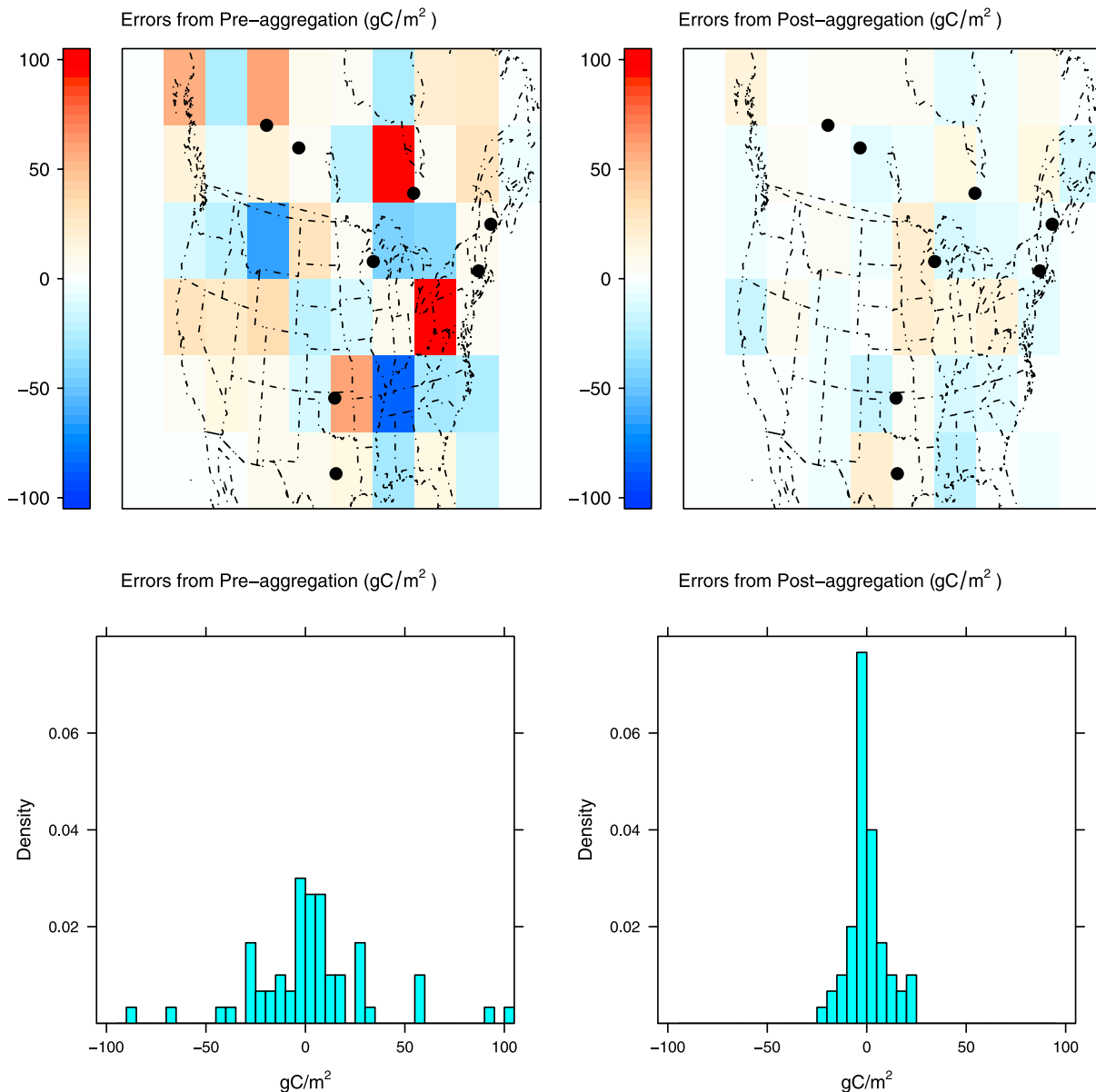
**Figure 4.** The improvement of posterior with respect to prior for preaggregated (dark gray) and postaggregated (light gray) inversion grid, factored over noise level and decorrelation length scale of true pattern used. Preaggregated (dark gray) inversions are only performed for grid sizes between 200 and 1200 km. An individual box plot shows the median (square for postagg and circle for preagg), the interquartile range (25th to 75th quantile) as the solid gray (light or dark) box, and whiskers indicating the most extreme data point which is no more than 1.5 times the interquartile range from the box (an ad hoc indicator of outliers).

in this paper because of the fact that our observations/data are changing with different simulations. We choose reasonably broad a priori specifications that should apply across many different models and then test how well the estimation procedure can reproduce a variety of simulated ‘true’ flux situations, each with a corresponding set of simulated  $\text{CO}_2$  observations from the eight towers. Recall from section 2 that the forward model of both fluxes (SiB3) and transport (RAMS) operate on a 40 km grid and is then postaggregated to a 100 km grid for computational reasons.

[19] A key component of atmospheric  $\text{CO}_2$  inversions is the specification of a priori error bounds for the different fluxes. An intercomparison of atmospheric  $\text{CO}_2$  inversion models (Transcom3 [Gurney *et al.*, 2002]) provided source/sink estimates on the order of a few tenths of a Pg of carbon per inversion region per year. When compared to the actual net photosynthesis or ground respiration fluxes for this region, this results in uncertainties on the order of 10–30% in either direction, on a cumulative basis. We chose to represent ensembles of potential ‘true’ flux scenarios with mean zero, spatially correlated, 20% marginal standard

deviation, Gaussian-based biases for individual 100 km grid cell GPP and respiration. These biases also seem to be a reasonably conservative a priori specification for the scalar multiplier on the spatial portion of the prior Gaussian covariance. In other words, we do not expect GPP and RESP biases to be outside of  $\pm 40\%$  of the a priori estimates. Simulated flux bias realizations (examples shown in Figures 2b and 2f) are drawn from this range and we assume this is known to set the a priori covariance matrix. Small-scale spatial noise of the same order also seems reasonable, and in combination with the spatial component generates a suitably wide range of potential biases, on the order of 40% standard deviation for the individual 100 km grid cells for which they are applied.

[20] Decorrelation length scales are investigated at levels of 100 km, 500 km, 1000 km, and 2000 km. Small-scale Gaussian flux noise will be allowed to vary between standard deviation levels of 1%, 5%, 10%, 20%, and 40% of the a priori fluxes. The a priori scalar standard deviation on the spatial covariance term is set to 20% and the prior inversion decorrelation length scale will be set to 500 km, a



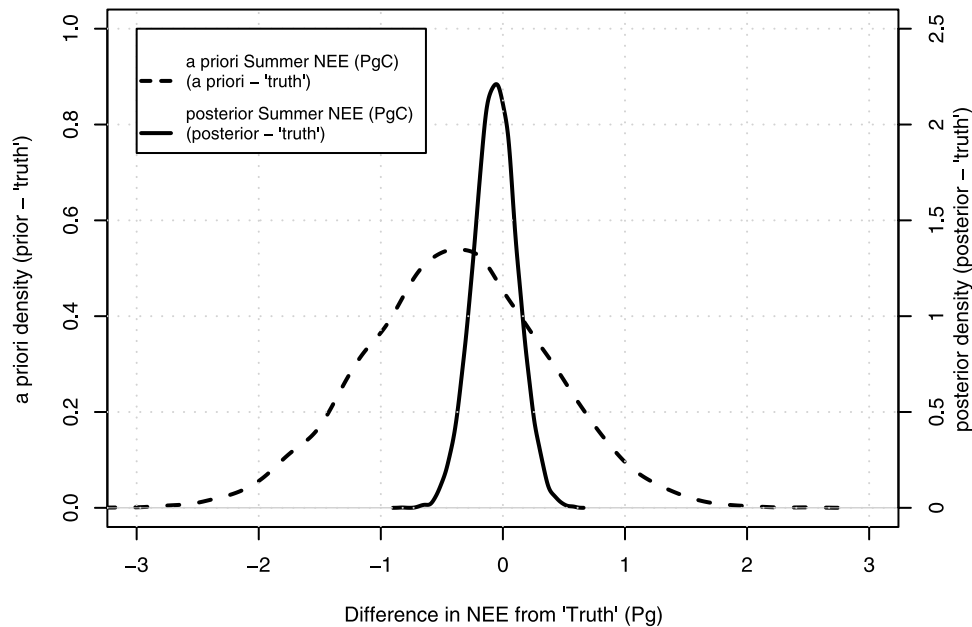
**Figure 5.** A comparison of errors from preaggregation versus postaggregation. Underlying spatial pattern comes from Figure 2. Preaggregation is a result of summing carbon fluxes on 1200 by 1200 km grid and running inversion on that grid. Postaggregation is a result of summing carbon fluxes on a 100 by 100 km grid and running the inversion on that grid and then summing up fluxes to a 1200 by 1200 km grid. Distinctly better results for postaggregation of fine-scale (100 km grid) inversion shown.

reasonably conservative prior compromise between similar parameters used in some recent papers [Michalak *et al.*, 2004; Peylin *et al.*, 2005]. For each combination of these two levels, 18 realizations of each scenario were run using randomly generated pseudo data corresponding to the levels used. Each realization introduces random “observation” error (mostly transport error) and random flux bias spatial patterns, both large and small scale. Since the temporally varying sampling pattern of the 8 towers is stationary, we must ensure that many different potential flux patterns are realized by the experiments so that the results are not dependent upon the sampling footprint of the towers.

[21] A specific example is presented to show the methodology of one realization. Figure 2 shows the spatial noise

pattern, the longer-scale spatially correlated signal, as well as the summed bias and the inversion estimate for both GPP and respiration fluxes. This particular example employed a noise level of 20%, equivalent to the scalar variability of the spatially correlated signal. The spatial decorrelation length scale used to create the correlated flux errors was 500 km, equal to that used as the a priori estimate. Table 1 shows summary statistics for the mean flux estimates of upscaled, increasingly coarse, gridded flux regions for this example. These statistics will be used as the measure of fit for inversions based upon the complete set of levels mentioned above. In section 3 we present inversion results across a variety of ‘noise’ levels and decorrelation length scales. A





**Figure 6.** Prior and posterior cumulative NEE over period of 11 May to 31 August 2004, for example, shown in Figure 2 and Table 1.

summary of this procedure, for postaggregated experiments, is shown in Figure 3.

### 3. Results and Discussion

[22] Results from the sample realization, shown in Figure 2, indicate that the posterior improves fluxes considerably over the a priori estimates. Improvement in the spatial average RMSE over the prior fluxes is from 40% to 90% depending upon the postaggregation level. For example, Table 1 shows that when the inversion is run on a 100 km by 100 km grid and the results are postaggregated to 1200 km by 1200 km grid, the average root mean squared error (RMSE) over all of the 1200 km by 1200 km grid cells is reduced from 26.8 g/m<sup>2</sup> to 8.0 g/m<sup>2</sup>. This is promising, considering that the level of small-scale noise (20% at 100 km) is equivalent to that of the spatially correlated portion of the flux errors (20%) for this example.

[23] Figure 4 shows these results over the entire range of small-scale variability and decorrelation length-scale parameters given in the algorithm above. The aggregated results, based upon 100 km resolution inversions, are shown in light gray. Variability within each panel of the image is due to the fact that the underlying bias field is not known and therefore has to be sampled over the set of all possible bias fields. The improvement in the spatial average RMSE over the prior is generally in the range of 20% to 90% over all combinations. The results show that the inversion is robust to small-scale spatial noise over a wide range of noise levels and decorrelation length scales. Although it may seem at first glance that these results contradict findings of others, such as *Peylin et al.* [2005] who found that changing a priori covariance assumptions impacts the strength and location of corrections, spatially, it must be understood that these results are presented as large-scale spatial averages. The degree and location of correction is likely to change with varying a priori spatial assumptions on the errors but as one

postaggregates results to larger scales, corrections are more robust. This is likely a result of varying a priori spatial assumptions driving correlated posterior flux estimates.

[24] The power of higher-resolution inversions versus lower-resolution “preaggregated” inversions is shown in Figure 4 as well. Inversions performed on the grid cell size shown in the  $x$  axis are shown in dark gray. For instance, at the point in an individual panel at which the  $x$  axis indicates 600 km, the light gray results give aggregated results based upon 100 km inversions while the dark gray results give results based upon 600 km inversions. The difference is clearly most sensitive to the spatial correlation length scale of the bias pattern while much less sensitive to the layer of noise added to the flux biases. This is as one would expect: very smooth bias fields require less precise spatial estimates of the biases while less smooth bias fields require more precise spatial estimates.

[25] Preaggregated and postaggregated inversion results both provide significant NEE corrections but postaggregated results provide larger improvements in estimation than preaggregated results. This is investigated by plotting an example based upon the spatial patterns shown in Figure 2. First, an inversion is run at a 100 km resolution and the results are statistically combined to 1200 km resolution. Then the various carbon fluxes are summed up across a 100 km grid to a 1200 km grid and the inversion is run at the 1200 km grid resolution. The results are shown in Figure 5. It is clear that postaggregation is preferable [*Kaminski et al.*, 2001; *Engelen et al.*, 2002]. If one reviews the differences in the estimates, it becomes clear that they often do not appear in the grid cell that contains the CO<sub>2</sub> observing tower, or necessarily in completely unconstrained grid cells. The largest errors appear to coincide with locations where steep sampling gradients (i.e., the upwind sampling crossing primarily a corner of the grid cell) intersect with fairly significant and heterogeneous fluxes at the 100 km scale, the scale of the fine-scale inversion. This manifests itself as

a type of “halo” effect around the combined sampling footprint of the towers.

[26] Figure 6 shows the “contraction” of the cumulative NEE integrated over the entire domain from the a priori cumulative flux to the posterior cumulative flux, centered around the assumed true cumulative NEE. The a priori NEE is the same for all the inversions while the posterior NEE distribution is based upon the example inversion given previously in Figure 2. The posterior cumulative flux estimates are much closer to the truth, displaying significantly less variability. Furthermore, the a priori spatially integrated cumulative fluxes appear to show a reasonable range of possible deviations,  $\pm 3\text{PgC}$  per year, from the a priori assumed mean zero annual NEE balance of SiB3, representing the potential to encompass many realistic source/sink scenarios.

#### 4. Conclusions

[27] The results of this paper show that NEE predictions can be significantly improved when large-scale spatial bias patterns exist in the GPP and RESP estimates. Predictions are improved across a range of possible spatial decorrelation length scales. Furthermore, and most importantly, these relatively large-scale postaggregated fluxes are robust to significant small-scale spatial noise that may exist in the flux biases at resolutions that are commonly used for regional inversion studies.

[28] One might have predicted that the inversion would be influenced heavily by small-scale variability in a few grid cells surrounding the towers where the  $\text{CO}_2$  observations were made. However, even when only 33% of the overall variability is on the larger scales, improvements of greater than 40% (RMSE) can be made. Furthermore, the estimates get more accurate as the region of interest gets larger. In general, this is not true of eddy-covariance-based flux tower measurements which often capture the effect of a small flux footprint (a few km). These measurements may not be very representative of surrounding fluxes, even those in close proximity to the tower and shows the value of collecting and analyzing  $\text{CO}_2$  mixing ratio measurements.

[29] The results also show the continued importance of running inversions at the finest scale available and this confirms the analysis made by [Kaminski et al., 2001]. Preaggregated and postaggregated inversion styles both show robustness to small-scale spatial variability in the flux biases. However, it is clear that preaggregation severely diminishes the quality of the corrections to NEE. In particular, there should be a focus on improving the accuracy of inversions in areas with steep sampling gradients and heterogeneous fluxes.

[30] There are several components of a standard regional inversion which are not addressed in this paper because of the nature of the hypothesis and result. For example, the choice of temporal averaging time for observations is not necessarily needed for this paper but needs investigation in an applied regional inversion. Boundary inflow of  $\text{CO}_2$  also plays a critical role in regional inversions but is not needed for this paper. These will be investigated and included in an upcoming paper which focuses upon GPP/RESP/NEE prediction in 2004 for North America.

[31] **Acknowledgments.** This research was funded by NOAA contract NA17RJ1228 and by Department of Energy grant DE-FG02-02ER63474. We wish to thank the North American Carbon Program and the Office of Science as well as the Numerical Terradynamics Simulation Group (NTSG) at University of Montana for providing the FPAR/LAI data that was used in this paper. I would like to thank the other authors as well as my Ph.D. committee members whose comments strengthened the paper, i.e., Scott Denning, Niall Hanan, Stephen Ogle, and Jennifer Hoeting, all of Colorado State University. Although not directly used in this paper, we thank the following tower PIs whose work and data provided motivation for this paper, WLEF (Arlyn Andrews, NOAA GMD), ARM at Great Plains (Sebastien Biraud), LBNL at Harvard Forest (William Munger, Harvard University, Argyle), ME (Arlyn Andrews, NOAA GMD), WKWT at Moody, Texas (Arlyn Andrews, NOAA GMD), Fraserdale (Douglas Worthy, MSC), Western Peatland (Larry Flanagan, University of Lethbridge), and BERMS at Candle Lake (Douglas Worthy, MSC). Additionally, I would like to thank the reviewers for many helpful comments and useful suggestions for improving this manuscript.

#### References

- Baker, D. F., S. C. Doney, and D. S. Schimel (2006), Variational data assimilation for atmospheric  $\text{CO}_2$ , *Tellus B*, 58, 359–365.
- Baker, I., A. S. Denning, N. Hanan, L. Prihodko, M. Uliasz, P. Vidale, K. Davis, and P. Bakwin (2003), Simulated and observed fluxes of sensible and latent heat and  $\text{CO}_2$  at the WLEF-TW tower using Sib 2.5, *Global Change Biol.*, 9, 1262–1277, doi:10.1046/j.1365-2486.2003.00671.x.
- Carouge, C., P. Bousquet, P. Peylin, P. J. Rayner, and P. Ciais (2008a), What can we learn from European continuous atmospheric  $\text{CO}_2$  measurements to quantify regional fluxes—Part 1: Potential of the network, *Atmos. Chem. Phys. Discuss.*, 8, 18,591–18,620.
- Carouge, C., P. Peylin, P. J. Rayner, P. Bousquet, F. Chevallier, and P. Ciais (2008b), What can we learn from European continuous atmospheric  $\text{CO}_2$  measurements to quantify regional fluxes—Part 2: Sensitivity of flux accuracy to inverse setup, *Atmos. Chem. Phys. Discuss.*, 8, 18,621–18,649.
- Collatz, G. J., J. T. Ball, C. Grivet, and J. A. Berry (1991), Physiological and environmental regulation of stomatal conductance, photosynthesis, and transpiration: A model that includes a laminar boundary layer, *Agric. For. Meteorol.*, 54, 107–136, doi:10.1016/0168-1923(91)90002-8.
- Collatz, G. J., M. Ribas-Carbo, and J. A. Berry (1992), Coupled photosynthesis-stomatal conductance model for leaves of C4 plants, *Aust. J. Plant Physiol.*, 19, 519–538.
- Cressie, N. (1993), *Statistics for Spatial Data*, Wiley, New York.
- Dai, Y., et al. (2003), The common land model (CLM), *Bull. Am. Meteorol. Soc.*, 84, 1013–1023, doi:10.1175/BAMS-84-8-1013.
- Denning, A. S., J. G. Collatz, C. Zhang, D. A. Randall, J. A. Berry, P. J. Sellers, G. D. Colello, and D. A. Dazlich (1996), Simulations of terrestrial carbon metabolism and atmospheric  $\text{CO}_2$  in a general circulation model. Part 1: Surface carbon fluxes, *Tellus B*, 48, 521–542.
- Denning, A. S., M. Nicholls, L. Prihodko, I. Baker, P. Vidale, K. Davis, and P. Bakwin (2003), Simulated variations in atmospheric  $\text{CO}_2$  over a Wisconsin forest using a couple ecosystem-atmosphere model, *Global Change Biol.*, 9, 1241–1250, doi:10.1046/j.1365-2486.2003.00613.x.
- Desai, A. R., P. V. Bolstad, B. D. Cook, K. J. Davis, and E. V. Carey (2005), Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the upper Midwest, USA, *Agric. For. Meteorol.*, 128, 33–55, doi:10.1016/j.agrformet.2004.09.005.
- Engelen, R. J., A. S. Denning, and K. R. Gurney (2002), On error estimation in atmospheric  $\text{CO}_2$  inversions, *J. Geophys. Res.*, 107(D22), 4635, doi:10.1029/2002JD002195.
- Enting, I., C. Trudinger, and R. Francey (1994), A synthesis inversion of the concentration and  $\delta^{13}\text{C}$  of atmospheric  $\text{CO}_2$ , *Tellus B*, 47, 35–52.
- Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99, 143–162.
- Fan, S., M. Gloor, J. Mahlman, S. Pacala, J. Sarmiento, T. Takahashi, and P. Tans (1998), A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models, *Science*, 282, 442–446.
- Farquhar, G. D., S. V. Caemmerer, and J. A. Berry (1980), A biochemical model of photosynthetic  $\text{CO}_2$  fixation in the leaves of C3 species, *Planta*, 149, 78–90, doi:10.1007/BF00386231.
- Gerbig, C., J. C. Lin, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger (2003), Toward constraining regional-scale fluxes of  $\text{CO}_2$  with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, *J. Geophys. Res.*, 108(D24), 4757, doi:10.1029/2003JD003770.
- Gurney, K. R., et al. (2002), Towards robust regional estimates of  $\text{CO}_2$  sources and sinks using atmospheric transport models, *Nature*, 415, 626–630, doi:10.1038/415626a.

- Hanan, N. P., J. A. Berry, S. B. Verma, E. A. Walter-Shea, A. E. Suyker, G. G. Burba, and A. S. Denning (2005), Testing a model of CO<sub>2</sub>, water and energy exchange in Great Plains tallgrass prairie and wheat ecosystems, *Agric. For. Meteorol.*, *131*, 162–179.
- Johnson, R. A., and D. W. Wichern (1988), *Applied Multivariate Statistical Analysis*, Prentice Hall, Englewood Cliffs, N. J.
- Kaminski, T., P. J. Rayner, M. Heimann, and I. G. Enting (2001), On aggregation errors in atmospheric transport inversions, *J. Geophys. Res.*, *106*, 4703–4715, doi:10.1029/2000JD900581.
- Lauvaux, T., M. Uliasz, C. Sarrat, F. Chevallier, P. Bousquet, C. Lac, K. J. Davis, P. Ciais, A. S. Denning, and P. J. Rayner (2008), Mesoscale inversion: First results from the CERES campaign with synthetic data, *Atmos. Chem. Phys.*, *8*, 3459–3471.
- Law, B. E., O. J. Sun, J. Campbell, S. V. Tuyl, and P. E. Thornton (2003), Changes in carbon storage and fluxes in a chronosequence of ponderosa pine, *Global Change Biol.*, *9*, 510–524, doi:10.1046/j.1365-2486.2003.00624.x.
- Michalak, A. M., L. Bruhwiler, and P. P. Tans (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, *J. Geophys. Res.*, *109*, D14109, doi:10.1029/2003JD004422.
- Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running (2007), Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, *Remote Sens. Environ.*, *111*, 519–536, doi:10.1016/j.rse.2007.04.015.
- Nicholls, M. E., A. S. Denning, L. Prihodko, P. Vidale, K. Davis, and P. Bakwin (2004), 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.
- Peters, W., J. B. Miller, J. Whitaker, A. S. Denning, A. Hirsch, M. C. Krol, D. Zupanski, L. Bruhwiler, and P. P. Tans (2005), An ensemble data assimilation system to estimate CO<sub>2</sub> surface fluxes from atmospheric trace gas observations, *J. Geophys. Res.*, *110*, D24304, doi:10.1029/2005JD006157.
- Peters, W., et al. (2007), An atmospheric perspective on North American carbon dioxide exchange: Carbon Tracker, *Proc. Natl. Acad. Sci. U. S. A.*, *104*, 18,925–18,930, doi:10.1073/pnas.0708986104.
- Peylin, P., et al. (2005), Daily CO<sub>2</sub> flux estimates over Europe from continuous atmospheric measurements: Part I, Inverse methodology, *Atmos. Chem. Phys.*, *5*, 3173–3186.
- Randall, D. A., et al. (1996), A revised land-surface parameterization (SiB2) for atmospheric GCMs. Part 3: The greening of the CSU General Circulation Model, *J. Clim.*, *9*, 738–763, doi:10.1175/1520-0442(1996)009<0738:ARLSPF>2.0.CO;2.
- Rodenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003), CO<sub>2</sub> flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem. Phys.*, *3*, 1919–1964.
- Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher (1986), A simple biosphere model (SiB) for use within general circulation models, *J. Atmos. Sci.*, *43*, 505–531, doi:10.1175/1520-0469(1986)043<0505:ASBMFU>2.0.CO;2.
- Sellers, P. J., D. A. Randall, G. J. Collatz, J. A. Berry, C. B. Field, D. A. Dazlich, C. Zhang, G. D. Collelo, and L. Bounoua (1996), A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation, *J. Clim.*, *9*, 676–705, doi:10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2.
- Suits, N. S., A. S. Denning, J. A. Berry, C. J. Still, J. Kaduk, J. B. Miller, and I. T. Baker (2005), Simulation of carbon isotope discrimination of the terrestrial biosphere, *Global Biogeochem. Cycles*, *19*, GB1017, doi:10.1029/2003GB002141.
- Thornton, P. E., et al. (2002), Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests, *Agric. For. Meteorol.*, *113*, 185–222, doi:10.1016/S0168-1923(02)00108-9.
- Uliasz, M., and R. A. Pielke (1991), Application of the receptor oriented approach in mesoscale dispersion modeling, in *Air Pollution Modeling and its Applications VIII*, edited by H. Van Dop and D. G. Steyn, pp. 399–408, Plenum, New York.
- Vidale, P. L., and R. Stockli (2005), Prognostic canopy air space solutions for land surface exchanges, *Theor. Appl. Climatol.*, *80*, 245–257, doi:10.1007/s00704-004-0103-2.
- Wang, J. W., A. S. Denning, L. Lu, I. T. Baker, and K. D. Corbin (2006), Observations and simulations of synoptic, regional, and local variations in atmospheric CO<sub>2</sub>, *J. Geophys. Res.*, *112*, D04108, doi:10.1029/2006JD007410.
- Zupanski, D., A. S. Denning, M. Uliasz, M. Zupanski, A. E. Schuh, P. J. Rayner, W. Peters, and K. D. Corbin (2007), Carbon flux bias estimation employing Maximum Likelihood Ensemble Filter (MLEF), *J. Geophys. Res.*, *112*, D17107, doi:10.1029/2006JD008371.

---

K. D. Corbin, A. S. Denning, A. E. Schuh, and M. Uliasz, Graduate Degree Program in Ecology, Department of Atmospheric Science, Colorado State University, 1371 Campus Delivery, Fort Collins, CO 80523-1371, USA. (aschuh@atmos.colostate.edu)