Atmos. Chem. Phys. Discuss., 9, 22407–22458, 2009
www.atmos-chem-phys-discuss.net/9/22407/2009/

Atmospheric Chemistry and Physics Discussions

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

Regional-scale geostatistical inverse modeling of North American CO_2 fluxes: a synthetic data study

S. M. Gourdji[1], A. I. Hirsch[2], K. L. Mueller[1], A. E. Andrews[3], and A. M. Michalak[1,4]

[1]Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, USA
[2]National Renewable Energy Laboratory, Golden, CO 80401, USA
[3]Global Monitoring Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305, USA
[4]Department of Atmospheric, Oceanic & Space Sciences, University of Michigan, Ann Arbor, MI 48109, USA

Received: 27 August 2009 – Accepted: 5 October 2009 – Published: 23 October 2009

Correspondence to: S. M. Gourdji (sgourdji@umich.edu)

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

Abstract

Using synthetic continuous CO_2 measurements from the nine sampling locations operational across North America in 2004, this paper investigates the optimal setup for, and constraint on fluxes achieved by, a regional geostatistical atmospheric CO_2 inversion over the continent. The geostatistical framework does not require explicit prior flux estimates, nor any other process-based information, and is therefore particularly well suited for investigating the information content of the atmospheric CO_2 measurements from a limited network. The atmospheric data are first used with the Restricted Maximum Likelihood (RML) algorithm to infer the model-data mismatch and a priori spatial covariance parameters applied in the inversion. The implemented RML algorithm is found to infer robust spatial covariance parameters from the atmospheric data, as compared to the "true" solution, for cases where the flux and measurement timescales match, while model-data mismatch variances are inferred correctly across all examined cases. A series of analyses is also performed investigating the impact of the temporal scale of concentration measurements and fluxes on inversion results. Inversions using measurement data at sub-daily resolution are found to yield fluxes with a lower Root Mean Square Error (RMSE) relative to inversions using coarser-scale observations, whereas the flux resolution appears to have a lesser impact on the inversion quality. In addition, night-time data for the tall and marine boundary layer towers are found to help constrain fluxes across the continent, although they can potentially bias near-field fluxes. These general conclusions are likely to also be applicable to inversions using a synthesis Bayesian inversion approach. Overall, despite the relatively sparse and unevenly distributed network of nine towers across the North American continent, a geostatistical inversion using an optimal setup and relying solely on the atmospheric data constraint is found to estimate the North American sink for June 2004 to within approximately 10%.

ACPD

9, 22407–22458, 2009

1 Introduction

Improved estimates of regional-scale CO_2 land-atmosphere exchange are needed for designing and verifying carbon management and climate change mitigation policies, as well as for validating process-based models used to predict future CO_2 fluxes. Large-scale fluxes cannot be directly measured and have instead been inferred from atmospheric concentration patterns using inverse modeling techniques. While previous global inversion studies have used atmospheric CO_2 concentration measurements and transport models to infer continental-scale CO_2 fluxes (e.g. Gurney et al., 2002; Baker et al., 2006), the recent convergence of several factors has made it feasible to estimate higher-resolution CO_2 fluxes in a regional inverse modeling framework (e.g. Peylin et al., 2005; Lauvaux et al., 2008). First, continuous boundary layer measurements of atmospheric CO_2 taken at several North American and Eurasian sites (e.g. Bakwin et al., 1998; Haszpra, 1999) provide data with high temporal (and increasingly, high spatial) resolution to constrain carbon fluxes at finer scales. In addition, the continuous measurement locations tend to be sited in continental, low-altitude areas with strong biospheric activity, providing more information about flux variability at sub-continental scales relative to measurements taken in remote areas. Finally, recent advances in regional atmospheric transport modeling and the use of analyzed wind fields with high spatial resolution make it feasible to take advantage of continuous data from continental locations in regional inversions.

The use of continuous data in grid-scale CO_2 inversions is relatively new, and therefore many questions remain as to the optimal inversion setup for taking advantage of these large data streams. Synthetic data (a.k.a. "pseudo-data") experiments are useful in the design of inversions because they include a set of specified baseline fluxes with which results can be compared, making it easier to diagnose biases in inferred fluxes under a number of different scenarios. Synthetic data studies also help to isolate the impact of inversion setup choices by simplifying the analysis relative to real data inversions in two important ways. First, synthetic measurements are only influenced by

ACPD
9, 22407–22458, 2009

fluxes occurring within the domain of study, and therefore there is no need to specify boundary conditions for the selected domain. Secondly, the effect of atmospheric transport model errors can be removed by using the same transport model to create the synthetic measurements as is used to estimate fluxes in the inversion. Alternately, the effect of transport model error can be investigated by adding errors with known statistics, and observing their impact on the inversion results. Finally, synthetic data experiments provide the opportunity to characterize the information content of the current atmospheric monitoring network at various spatiotemporal scales and distances from the measurement locations.

This paper uses a series of synthetic data inversion experiments to evaluate a regional grid-scale inversion over North America using the geostatistical approach to CO_2 flux estimation (Michalak et al., 2004). The purpose of this study is to evaluate the optimal flux and concentration temporal resolutions for taking advantage of continuous measurements from nine tower CO_2 observing sites that were operational in the United States and Canada in 2004. In addition, using one or more "optimal" inversion setups, the information content of this atmospheric network is assessed by comparing the results to the "true" prescribed fluxes at both the grid and aggregated scales.

A series of synthetic data inversion studies have previously been published that focus on the use of continuous atmospheric CO_2 measurements in regional inversions over Australia and Europe. In addition to focusing on the inversion setup, these studies analyze the constraint on inferred fluxes achieved by using continuous data from realistic measurement networks in their respective regions. Law et al. (2002) first investigated the reduction in bias for a global large region inversion made possible by using 4-hourly measurements instead of monthly-averaged data, as had been the case in earlier global inversion studies (e.g. Kaminski et al., 1999). This study concluded that high-resolution measurements help to reduce uncertainties, but that biases are only reduced proportionately when fluxes are estimated at a sufficiently fine spatial resolution, as shown in a case study over the Australasian continent. Law et al. (2003, 2004) then incorporated transport error and diurnally-varying fluxes into the same inversion

ACPD
9, 22407–22458, 2009

setup to investigate how to best handle these increasingly realistic inversion conditions. By solving for both a daily average and a daytime only flux over a given monthly period, Law et al. (2004) showed that it was possible to avoid the diurnal rectifier effect (Denning et al., 1996), which can lead to overestimated sources in inversions.

In a two-part paper, Carouge et al. (2008a,b) investigated the constraint on flux estimates achieved using a 10-tower continuous measurement network in Europe. The first part of the study found that even under perfect transport conditions, the recovery of fluxes at the grid-scale was poor, but that reasonable results could be recovered by aggregating in space and time to a 10-day average flux over the Western European region, where measurements are most dense. In the second paper, Carouge et al. (2008b) found that including both spatial and temporal correlation in the prior flux error covariance matrix was crucial for spreading the information content of the atmospheric data over larger spatial scales. They also performed a series of sensitivity tests varying prior flux and measurement errors; not surprisingly, the constraint achieved at finer scales was found to be more sensitive to inversion setup choices (including random transport error) than that at aggregated scales.

The study presented here also uses a series of synthetic data inversion experiments with continuous data to evaluate the choice of the temporal averaging of observations, the best approach for handling diurnally-varying fluxes, the effect of prior error and model-data mismatch covariance assumptions, and the effect of transport error on estimated results. However, the current work is distinct from these previous studies in several key respects. This is the first synthetic data inversion study to focus on the flux constraint achieved by using continuous measurements over the North American continent from the 2004 tower network. In addition, in contrast to previous studies, the current work uses the geostatistical inverse modeling (GIM) approach to CO_2 flux estimation (Michalak et al., 2004) rather than the synthesis Bayesian approach, and employs a Lagrangian rather than a Eulerian transport model, as will be discussed further below. This work also assesses the appropriate flux resolution for minimizing bias with continuous measurements, and analyzes temporal aggregation errors associated

with various combinations of flux and measurement resolutions. Finally, this work assesses the potential for using the atmospheric CO_2 data to correctly infer the full set of covariance parameters used in the inversion.

The choice of the GIM approach to CO_2 flux estimation is based on the work of Michalak et al. (2004), Mueller et al. (2008) and Gourdji et al. (2008). The approach is Bayesian, but rather than prescribing a prior estimate of the flux distribution, the GIM approach uses: 1) a deterministic model of the trend that estimates the relationship of CO_2 flux to key covariates, and 2) a prior covariance function that describes the expected variability in flux departures from the trend as a function of the separation distance in space and time between individual estimated fluxes. The model of the trend can be as simple as a single mean unknown flux across the domain, as used in portions of this study, or can include more complex components, such as a linear combination of auxiliary variables related to CO_2 flux processes (Gourdji et al., 2008).

The GIM technique is a strongly data-driven approach that relies on the atmospheric measurements to inform a number of different aspects of the inversion. First, the atmospheric data can be used in an initial step to estimate the a priori and model-data mismatch covariance parameters that will be used in the inversion (Michalak et al., 2004). Second, the atmospheric data are used to simultaneously estimate the coefficients in the model of the trend and the fine-scale fluxes, as well as their associated uncertainties. An additional feature of the GIM approach is the ability to directly estimate fluxes at fine resolutions, thereby minimizing spatial aggregation errors (Kaminski et al., 2001), which can occur when fixed flux patterns are imposed for large regions (e.g. Law et al., 2002). For example, in this study, fluxes are directly estimated at a $1^{\circ}\times1^{\circ}$ resolution.

Regardless of the choice of the statistical inversion technique, transport model error is of particular concern when using atmospheric measurements from inland continental towers that are sensitive to strong, diurnally and synoptically-varying flux signals (Geels et al., 2004). Regional-scale Lagrangian particle dispersion models, as employed in this study, are well suited for making use of such data in inversions, because they allow

ACPD
9, 22407–22458, 2009

the representation of CO_2 measurements at a point in space (Lin et al., 2003), thereby reducing model-data mismatch errors, and making fuller use of the information content of the measurements relative to grid-based Eulerian models (Patra et al., 2008). For the current study, this advantage is only relevant when transport model error is considered; however, the use of a Lagrangian transport model is implemented in anticipation of future real data inversions applying the inversion setup developed here.

The synthetic data experiments performed in this study were designed to identify the optimal setup for a GIM inversion using continuous measurements over North America. The first set of experiments focuses on the feasibility of using atmospheric data with the Restricted Maximum Likelihood algorithm for recovering the spatial covariance structure of CO_2 fluxes across the continent, as well as model-data mismatch variances. In the second set of experiments, a series of inversions is performed (both with and without transport error) to identify the optimal temporal scales for flux estimates and for binning the atmospheric observations (ranging from 3-hourly to 8-day averages) in order to minimize bias in the inferred fluxes. Model-data mismatch errors are also characterized as a function of temporal scale, observation location, and time of day. The two best performing inversion setups are chosen for further development and analysis in the third set of experiments. For these two cases, a series of refinements are made to the covariance matrices to further reduce errors and create realistic real data inversion conditions. Finally, the results from the final setup for these two cases are analyzed in comparison to the true fluxes at both grid and aggregated ecoregion scales to assess inversion quality.

2 Methods and experimental setup

This section describes the experimental setup common to all analyses performed in this study.

2.1 Inversion cases

Six inversion setups are considered (Table 1), where the temporal resolution of both flux and synthetic concentration data are varied between 3-hourly, daily, and 8-day averages. The naming convention for the six cases represents the temporal resolution of retrieved fluxes followed by that of the concentration data, such that, for example, the F8d/C3h case uses 3-hourly average observations to estimate 8-day average fluxes. For all experiments, fluxes are estimated at a 1°×1° resolution, with the domain including all land grid cells within the range of 10°–70° N and 50°–170° W, yielding 2641 estimation locations.

2.2 Geostatistical inverse modeling (GIM)

In the GIM approach, the spatial and/or temporal variability of fluxes is described through a model of the trend and a covariance function that describes the correlation of grid-scale flux deviations from this trend. For the remainder of this paper, "grid-scale" refers to the 1°×1° estimation resolution. The GIM approach is summarized here, and readers are referred to Michalak et al. (2004), Mueller et al. (2008), and Gourdji et al. (2008) for additional details.

The GIM approach entails minimizing the following Bayesian objective function:

$$L_{s,\beta} = \frac{1}{2}(\mathbf{z} - \mathbf{Hs})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{z} - \mathbf{Hs}) + \frac{1}{2}(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}\mathbf{Q}^{-1}(\mathbf{s} - \mathbf{X}\boldsymbol{\beta}) \quad (1)$$

where the vector $\mathbf{z}$ ($n\times 1$) represents the atmospheric CO_2 measurements, and $\mathbf{s}$ ($m\times 1$) is the vector of estimated fluxes. $\mathbf{H}$ ($n\times m$) contains the sensitivity of CO_2 measurements to surface fluxes as derived from an atmospheric transport model, with units of ppm/(μmol/(m^2×s)), and $\mathbf{Hs}$ therefore represents a vector of modeled CO_2 observations. $\mathbf{X}$ is a known ($m\times p$) matrix containing the flux covariates in the model of the trend, $\boldsymbol{\beta}$ are ($p\times 1$) unknown drift coefficients, and $\mathbf{X}\boldsymbol{\beta}$ is the model of the trend. The two covariance matrices in the objective function, $\mathbf{R}$ ($n\times n$) and $\mathbf{Q}$ ($m\times m$), balance the relative weight of the atmospheric data and the model of the trend in estimating the fluxes.

ACPD
9, 22407–22458, 2009

R is the model-data mismatch covariance matrix describing the expected magnitude of discrepancies between observed (**z**) and modeled (**Hs**) CO_2 concentrations (due to measurement, transport model, representation and aggregation errors). **Q** ($m \times m$) is the a priori flux covariance matrix, containing information on how flux deviations from the model of the trend (i.e. $\mathbf{s}-\mathbf{X}\boldsymbol{\beta}$) are correlated in time and space. The setup of each component of the GIM objective function will be further discussed in Sects. 2.3–2.7.

The geostatistical inverse problem involves estimating both $\boldsymbol{\beta}$ and **s** (e.g. Michalak et al., 2004), whereas the synthesis Bayesian approach estimates only **s**. Minimizing Eq. (1) with respect to the fluxes, **s**, and the drift coefficients, $\boldsymbol{\beta}$, yields a system of linear equations:

$$\begin{bmatrix} \mathbf{HQH}^{\mathrm{T}}+\mathbf{R} & \mathbf{HX} \\ (\mathbf{HX})^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}^{\mathrm{T}} \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{HQ} \\ \mathbf{X}^{\mathrm{T}} \end{bmatrix} \tag{2}$$

A posteriori best estimates of $\boldsymbol{s}$ are then calculated as:

$$\hat{\mathbf{s}} = \boldsymbol{\Lambda}\mathbf{z} \tag{3}$$

The best estimates of flux can alternately be expressed as the sum of a deterministic model of the trend ($\mathbf{X}\hat{\boldsymbol{\beta}}$) and a spatiotemporally correlated stochastic component:

$$\hat{\mathbf{s}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{QH}^{\mathrm{T}}\left(\mathbf{HQH}^{\mathrm{T}}+\mathbf{R}\right)^{-1}\left(\mathbf{z}-\mathbf{HX}\hat{\boldsymbol{\beta}}\right) \tag{4}$$

The a posteriori uncertainties of the grid-scale fluxes are given by:

$$\mathbf{V}_{\hat{\mathbf{s}}} = -\mathbf{XM} + \mathbf{Q} - \mathbf{QH}^{\mathrm{T}}\boldsymbol{\Lambda}^{\mathrm{T}} \tag{5}$$

The a posteriori uncertainties of the flux estimates are a composite of uncertainty associated with the estimation of unknown drift coefficients ($\boldsymbol{\beta}$) in the model of the trend, the spatiotemporal variability of fluxes as represented in **Q**, and the overall constraint on fluxes as determined by the concentration footprints, the model of the trend and the covariance matrices.

2.3 Atmospheric transport (H)

The Stochastic Time-Inverted Lagrangian Transport Model (STILT) model (Lin et al., 2003) is used to quantify the sensitivity of measurements to fluxes, or the concentration footprints which populate the atmospheric transport matrix **H**. STILT, which has already been applied in several pilot studies aimed at constraining CO_2 sources and sinks in the United States (Gerbig et al., 2003, 2006; Lin et al., 2004; Matross et al., 2006), represents air arriving at observation locations as an ensemble of particles that are transported backward in time.

The particle velocities in STILT are derived from meteorological fields generated by gridded numerical weather prediction models, in this case from the Weather Research and Forecasting (WRF) model (Skamarock et al., 2005), version 2.2. For this study, WRF v2.2 was configured to use three levels of high resolution nesting around the three tallest measurement towers (LEF, AMT and WKT, see Table 2), which are maintained by the United States National Oceanic and Atmospheric Administration. The 2-km resolution grid around these towers was embedded in a 10-km resolution grid over the northern Midwest, Gulf Region, and New England, extending to approximately 105° W, and then an outermost 40-km resolution grid covering the rest of the North American continent (Fig. 1). These nested WRF wind fields contain more fine-scale information than the products used in typical global-scale inversions, or other readily available regional analysis fields (e.g. from the Eta Data Assimilation System, or EDAS, available at http://www.ncdc.noaa.gov). In addition, nesting the WRF model down to high spatial resolution around the tallest towers allows the model to properly resolve mesoscale circulations (Pielke et al., 1991), cloud venting (Stull, 1984) and other detailed atmospheric phenomena that affect CO_2 transport in these regions (Nehrkorn et al., 2009).

At each measurement location, 10-day back-trajectories of 500 particles were generated every hour from 1 June to 8 July 2004 using WRF/STILT. Concentration footprints, or sensitivities, are then calculated at 3-hourly intervals back in time by integrating par-

ACPD
9, 22407–22458, 2009

ticle trajectories over the North American 1°×1° grid using Eq. (8) from Lin et al. (2003). Finally, these high-resolution **H** matrices are aggregated to the flux and concentration temporal resolutions corresponding to each examined case (Table 1) for use in the inversions.

2.4 Synthetic concentration time series (z)

One goal of this study is to assess the projected accuracy of North American estimates of CO_2 flux using a contemporary observation network. Therefore, synthetic data were generated at the highest sampling elevation of the nine towers that were collecting continuous high-precision calibrated CO_2 measurements in North America in June of 2004 (Fig. 1, Table 2). For the remainder of the paper we refer to two groups of towers which have different characteristics: a group of tall (in this case $\geq\approx$400 m) or marine boundary layer towers with little diurnal cycle in the measurements (the "Tall/MBL" group, which includes LEF, WKT, BRW, and SBL) and a group of shorter continental towers that are strongly influenced by diurnal planetary boundary layer (PBL) dynamics (the "Short" group, which includes AMT, ARM, CDL, FRD, and HF). A primary criterion for determining whether a tower is "tall" is whether or not it normally penetrates the night-time stable boundary layer.

A full set of synthetic measurements without data gaps from 1 June to 8 July are generated by multiplying CO_2 surface flux estimates (**s**) from a biospheric model by the atmospheric transport matrices (**H**). This vector of modeled CO_2 (i.e. **Hs**) is first generated at the hourly resolution corresponding to the STILT particle releases for the nine tower locations. Then, the synthetic observation vectors are averaged to 3-hourly, daily, or 8-day average concentrations corresponding to the cases in Table 1. Finally, for those inversion experiments simulating transport error, random noise with a variance corresponding to the expected magnitude of errors at each tower are added to the synthetic measurements.

The biospheric fluxes used in this study are from the Carnegie Ames Stanford Approach terrestrial carbon cycle model, as configured for the Global Fire Emissions

ACPD
9, 22407–22458, 2009

Database v2 project (henceforth referred to as CASA-GFEDv2; Randerson et al., 1997; Van der Werf et al., 2006). For this study, the monthly-average CASA-GFEDv2 Net Ecosystem Exchange (NEE) was temporally downscaled to a 3-hourly resolution before generating the synthetic measurements, in order to test the ability of the inversion setup to accurately recover diurnally-varying fluxes. This downscaling was accomplished using the method of Olsen and Randerson (2004), which is based on net shortwave radiation and near-surface temperature data from the NASA Global Land Data Assimilation System (GLDAS; Rodell et al., 2004).

CASA-GFEDv2 was chosen because it is a well-accepted model and has also been used for specifying prior flux estimates in previous synthesis Bayesian inversion studies (e.g. Peters et al., 2007). However, the choice of biospheric model for a pseudo-data study is somewhat arbitrary, given that the aim of a synthetic data inversion is to assess the accuracy of the setup relative to a given set of prescribed fluxes. Regardless, the CASA-GFEDv2 fluxes represent the "truth" to which inversion results will be compared for the remainder of the paper.

2.5 Use of night-time measurements

Due to stable conditions, night-time measurements taken from shorter towers within the nocturnal boundary layer provide little information about flux over large aggregated spatial and temporal scales (Haszpra, 1999). In addition, meteorological fields used in transport models cannot reliably simulate the height of the night-time planetary boundary layer (PBL) or the sharp gradient across it, which can lead to biased flux estimates from inversions using night-time measurements from within the PBL (Geels et al., 2007). For example, Gerbig et al. (2008) found biases of up to 50% in night-time PBL height in a study using high-resolution winds from the European Centre for Medium-Range Weather Forecasts (i.e. ECMWF, available from http://data.ecmwf.int/data/); comparison of the high resolution WRF wind fields (T. Nehrkorn, personal communication) used in this study with wind profiler PBL-height measurements yielded a similar conclusion.

ACPD
9, 22407–22458, 2009

By modeling transport error as unbiased and uncorrelated noise, the current study does not consider systematic biases associated with PBL height. However, such transport biases would affect future real-data inversions, and therefore in the current study, only afternoon measurements are included for Short towers in those inversion cases estimating daily and 8-day average fluxes. For the five Short towers, which are all in the Eastern or Central Standard Time zones, "afternoon" is considered to be 18:00–24:00 UTC. For the one inversion case estimating 3-hourly fluxes, all 24 hours of measurements are included for the Short towers, as it would have been difficult to recover a diurnal cycle in the fluxes without a diurnal cycle in the measurement data.

For the Tall/MBL towers, all 24 hours of atmospheric data are included in the inversions for all cases (shown in Table 1). At these towers, observations are collected higher in the atmosphere or in the marine boundary layer, thereby sampling relatively well-mixed air, and night-time measurements are therefore assumed to be easier for the WRF/STILT model to represent relative to the Short towers. This was qualitatively confirmed by comparing the diurnal cycle in actual observations at these towers at the tallest sampling levels to that from transported CASA-GFEDv2 fluxes.

A number of sensitivity tests are performed to evaluate the choice of including night-time data for Tall/MBL towers vs. excluding it for Short towers. One test with the F8d/C3h case includes night-time data for the Short towers, so that 24 hours of measurement data are used for all measurement locations (similarly to the F3h/C3h case but instead estimating 8-day average fluxes). Next, the use of afternoon-only measurements for all nine towers is tested for the F8d/C3h, as well as the F8d/Cd and Fd/C3h cases. Here, afternoon values are selected to reflect shifts in local time zones for SBL (i.e. 15:00–21:00 UTC) and BRW (21:00–03:00 UTC). Overall, such experiments help to assess the additional constraint on fluxes provided by night-time measurements, relative to potential biases associated with their use.

ACPD
9, 22407–22458, 2009

2.6 Model of the trend (Xβ)

When estimating daily and 8-day average fluxes, a very simple model of the trend is applied, analogous to that used in Michalak et al. (2004) and Mueller et al. (2008). In this trend, one drift coefficient ($\boldsymbol{\beta}$) represents the mean value in space and time of flux across all grid-cells, and **X** is therefore a vector of ones. For the case estimating 3-hourly fluxes, the **X** matrix is instead structured to allow for both a mean continental flux and a longitudinal gradient, which are allowed to vary for each 3-hourly flux period throughout the day (but remain constant across days). The addition of the longitudinal gradient, which helps to account for the shifting day/ night boundary across the continent within a 24-hour period, was implemented based on an examination of CASA-GFEDv2 fluxes as a function of longitude.

In the GIM approach, the model of the trend can also include auxiliary variables which help to better represent the spatiotemporal variability of flux processes at the grid-scale (Gourdji et al., 2008). However, this pseudo-data study did not include any such variables in order to focus on the fluxes inferred using solely the atmospheric data constraint. In addition, including some of the variables that are themselves used as input into the CASA-GFEDv2 model may have provided an unrealistically large amount of information for a pseudo-data experiment.

2.7 Covariance matrices (Q and R)

The covariance matrix **Q** contains information on the spatial and/or temporal correlation of the flux deviations from the model of the trend **X**$\boldsymbol{\beta}$. Therefore **Q** is typically a block diagonal or full matrix in a geostatistical inversion. Because a simple model of the trend is used in this study (see Sect. 2.6), **Q** describes the covariance of the fluxes themselves, whereas with a more complex trend containing auxiliary variables, **Q** would describe the covariance of de-trended flux residuals.

In geostatistical applications, the prior covariance is modeled using a covariance function (e.g. Cressie, 1991; Kitanidis, 1997). Here, as in Michalak et al. (2004), we

ACPD
9, 22407–22458, 2009

use an isotropic exponential decay model for **Q**:

$$\mathbf{Q}\left(h\middle|\sigma^2_{\mathbf{Q}}, l\right) = \sigma^2_{\mathbf{Q}} \exp\left(-\frac{h}{l}\right) \quad (6)$$

where h is the separation distance between two grid cells, l is the correlation range parameter, and $\sigma^2_{\mathbf{Q}}$ is the asymptotic variance of fluxes at large separation distances. The correlation length for an exponential model is $\approx 3\,l$.

The same exponential model is used in this study for correlation in both space and time. When only spatial covariance is considered, **Q** is a block diagonal matrix, with the same block, which describes correlation between grid-cells, repeated for each time period in the inversion. When temporal covariance is additionally considered, the off-diagonal blocks in **Q** contain diagonal entries describing the correlation among grid-cells with themselves over time. Cross spatial-temporal covariance is not considered.

In contrast, model-data mismatch errors are assumed uncorrelated in space and time, yielding a diagonal matrix **R**, as is typical in most inversion studies. For all inversion experiments performed in this study, unless otherwise noted, two separate model-data mismatch variances are estimated for measurements from the Tall/MBL and Short towers, respectively.

All covariance parameters are estimated using the Restricted Maximum Likelihood (RML) approach, described in detail for atmospheric applications in Michalak et al. (2004) and Mueller et al. (2008). A single set of spatial covariance parameters is estimated for **Q** when solving for daily or 8-day average fluxes. However, when estimating 3-hourly fluxes, three sets of spatial covariance parameters are estimated corresponding to "day" (18:00–24:00 UTC), "night" (03:00–15:00 UTC), and "shelf" (00:00–03:00 and 15:00–18:00 UTC) periods, with "shelf" referring to periods when only portions of the continent are sunlit. This parameterization was chosen based on a qualitative examination of the spatial correlation of CASA-GFEDv2 fluxes across a 24-hour period, and represents the expected difference in the underlying spatial variability of photosynthesis and respiration, which dominate day- and night-time fluxes, respectively.

3 Analyses performed

Three main analyses are performed for this study. The first analysis focuses on the a priori covariance matrices for the inversion: first, the feasibility of using atmospheric data to correctly estimate covariance parameters in an inverse modeling setup, and second, the characterization of model-data mismatch variances by tower and time of day. The second analysis investigates the optimal measurement and flux temporal resolutions for minimizing errors in inferred fluxes under both "perfect transport" and simulated "transport error" conditions. Lastly, the third analysis focuses on the two best inversion setups identified in the second analysis, investigates various improvements in the parameterization of their covariance matrices, and examines resulting flux estimates at two different spatial scales.

3.1 Optimization of a priori covariance parameters

To infer unbiased fluxes with accurate uncertainty estimates, it is important to correctly specify the parameters associated with the spatial and temporal flux covariance (Gerbig et al., 2006), as well as the model-data mismatch variances. The ability of the RML approach to estimate these parameters directly from the atmospheric data in an inverse setup (referred to as RML-Inv) was previously demonstrated in Michalak et al. (2004) for the continental/global scale, and is investigated here for regional inversions. If RML-Inv is able to provide accurate covariance parameters, then this method provides a strong advantage relative to the use of other proxy methods implemented in previous studies to estimate covariance parameters in real data inversions (e.g. Rödenbeck et al., 2003; Baker et al., 2006; Peters et al., 2007).

3.1.1 Flux spatial covariance parameters (Q)

The "true" spatial covariance parameters for **Q** can be inferred for this synthetic data setup using RML in a kriging setup (referred to as RML-Krig, see Mueller et al., 2008;

ACPD
9, 22407–22458, 2009

and Gourdji et al., 2008), where the "true" fluxes in the RML objective function replace the atmospheric measurements transported backward to flux locations, as implemented in RML-Inv. In the current study, RML-Krig is run for each of the flux temporal resolutions (3-hour, daily and 8-day), while RML-Inv is run with the synthetic atmospheric measurements for each of the six inversion cases in Table 1. Finally, the inferred spatial covariance parameters in **Q** estimated by RML-Inv are compared to those from RML-Krig to assess the ability of the atmospheric data to correctly infer these parameters. For these tests, transport error was not considered.

3.1.2 Model-data mismatch variances (R)

The "true" model-data mismatch variances in **R** in this synthetic data study consist of temporal aggregation errors for the "perfect transport" inversions, and temporal aggregation errors plus the noise added to the measurements for the "transport error" inversions. The ability of the atmospheric data to correctly recover temporal aggregation error for the inversion cases with daily and 8-day fluxes is tested by comparing the RML-Inv estimated variances in **R**, using synthetic data vectors without added noise, to the "true" variances as discussed below. An additional test using RML-Inv for **R** aims to determine whether the method can correctly recover the true amount of random noise added to the synthetic measurements.

Aggregation errors occur in inversions when fluxes are estimated over long temporal intervals or when flux patterns within large spatial regions are assumed known and prescribed a priori (e.g. Kaminski et al., 2001). These errors result because the inversion cannot adjust the flux patterns within the specified spatial or temporal blocks, even though atmospheric observations are sensitive to the sub-block variability. In this pseudo-data setup, spatial aggregation error is technically zero, because fluxes are estimated at the same spatial resolution as the prescribed CASA-GFEDv2 fluxes (i.e. $1^{\circ} \times 1^{\circ}$). However, temporal aggregation errors occur when fluxes are estimated at temporal resolutions coarser than the 3-hourly resolution of the CASA-GFEDv2 fluxes used to generate the synthetic observations. The "true" temporal aggregation error for this

study can be calculated as the variance of the difference between two synthetic data vectors: the observations used in the inversion (generated using 3-hourly fluxes), and observations generated with fluxes pre-averaged to the daily or 8-day scale. "True" temporal aggregation errors are quantified in several different ways to assess how these errors vary by inversion case, tower, and time of day.

While temporal aggregation error is likely to be smallest for inversions estimating high resolution fluxes using high resolution atmospheric data, transport model errors are also expected to increase at higher temporal resolutions, potentially canceling out the benefit of lower temporal aggregation error. In order to evaluate this trade-off, as well as to add noise to the measurement data with realistic variances for the "transport error" inversions, transport errors are characterized for different inversion cases and towers. However, in this case, the "true" transport error variances cannot be calculated directly from the synthetic measurements. Therefore, RML-Inv is used with actual CO_2 observations for each of the examined cases in June 2004, to quantify the total model-data mismatch corresponding to temporal aggregation, transport and measurement errors. (Note that this is the only portion of the study where real atmospheric observations are used.) Assuming that measurement errors are minimal, the difference between these total model-data mismatch variances inferred using actual measurements and the temporal aggregation errors described previously represents a rough estimate of the magnitude of the transport model errors. After rounding to the nearest 0.5 ppm, white noise with the inferred variances is added to the synthetic observations for the "transport error" inversions.

3.2 Identification of best inversion setups

In the second set of experiments, a series of geostatistical inversions is run for each case (Table 1) using synthetic data vectors without and with added noise to represent transport model error. For all cases, the RML-Krig "true" parameters are used in **Q** to isolate the impact of the varying temporal resolutions in each inversion case from the estimation of the flux covariance parameters. Given that RML-Inv for **R** performed

ACPD
9, 22407–22458, 2009

well for all examined cases using synthetic data both with and without noise, RML-Inv values are used in **R** for all inversions throughout this study. Although the F3h/C3h case technically has no temporal aggregation error, a minimum variance of 0.01 ppm^2 is used for this case in the "perfect transport" inversion.

The flux estimates from all inversions, as well as the "true" fluxes, are aggregated a posteriori to 8-day temporal resolution in order to provide a common basis for comparison across cases. The inferred fluxes are then compared using two metrics. First, the root mean square error (RMSE) (e.g. Law et al., 2002) between the "true" and estimated fluxes is calculated at the native $1^\circ \times 1^\circ$ spatial resolution. The RMSE is calculated across all grid-cells in North America, as well as for a subset of grid-cells where the measurements exhibit high sensitivity to fluxes on an average monthly basis. A map of the average sensitivity of measurements to fluxes for June 2004 is shown in Fig. 2, with a contour line demarcating the 25% of grid-cells in high sensitivity areas which are used in the 2nd RMSE calculation. Second, the accuracy of the estimated a posteriori uncertainties (from Eq. 5) is evaluated by verifying the percent of $1^\circ \times 1^\circ$ "true" fluxes that fall within two standard deviations of the estimated fluxes. Ideally, 95% of fluxes should fall within this interval. Values significantly below 95% would indicate an underestimation of the true a posteriori uncertainties, whereas values substantially above 95% indicate overly conservative estimates of the a posteriori uncertainties.

3.3 Refinement and analysis of two best inversion setups

The two cases with the lowest RMSE among the "transport error" inversions are selected for further analysis. We then examine how these cases can be further improved and made more realistic by refining the setup and parameterization of their covariance matrices.

First, in a real-data inversion, there are no "true" spatial covariance parameters that can be obtained using RML-Krig, unlike in the current synthetic data setup. Therefore, using RML-Krig with a specific biospheric model in future work to estimate flux covariance parameters would bias results if that particular model under- or over-estimated the

amount of variability in the true underlying fluxes. In fact, Huntzinger et al. (2009) found that the degree of spatial variability differs strongly across common biospheric models. Therefore, in order to perform an inversion that relies solely on the atmospheric measurements to estimate both fluxes and covariance parameters, the RML-Inv results for both **R** and **Q** are applied in the two modified best inversion cases. However, given that not all cases produce reasonable spatial covariance parameters using RML-Inv (see Sect. 4.1), the case producing the optimal **Q** parameters for the given flux resolution is applied in the inversion.

Two other structural modifications are made to the covariance matrices for the two best inversion setups. The use of temporal flux covariance in **Q** is explored for the cases with a daily or 3-hourly flux resolution to help produce more accurate temporally-aggregated uncertainties. This option is not explored for the 8-day flux cases because the examined period included only four flux periods at this resolution. Finally, the use of a different variance in **R** for each of the nine towers is evaluated to investigate if further subdividing the model-data mismatch beyond average values for Tall/MBL and Short towers leads to a reduction in RMSE.

Fluxes from these two best cases with updated covariance matrices are evaluated (i) at the grid-scale, both visually and using the metrics outlined in Sect. 3.2, and (ii) by aggregating the 8-day estimates and "true" fluxes to ecoregions and the entire North American continent for comparison at these larger spatial scales.

4 Results and discussion

4.1 RML-Inv tests for spatial covariance parameters (Q)

The RML-Krig and RML-Inv estimates of the a priori spatial flux covariance parameters in **Q** are presented in Table 3. As described in Sect. 3.1, the RML-Krig parameters are derived directly from the "true" prescribed fluxes, whereas the RML-Inv parameters are inferred from the synthetic atmospheric data, in a manner analogous to what

ACPD
9, 22407–22458, 2009

would be possible in an inversion using real atmospheric data. As expected, the "true" parameters show that the overall variance of the fluxes decreases as the flux temporal resolution becomes coarser, as more of the short-term variability in the spatial flux distribution is averaged out. Also, the covariance structure of 3-hourly-average fluxes shows a strong dependence with time of day, with the highest variance during the day when photosynthetic fluxes dominate, the lowest at night, and intermediate values during the "shelf" period around dawn and dusk. The correlation range of the fluxes, on the other hand, remains relatively constant across both temporal flux resolution and time of the day for 3-hourly fluxes.

The ability of the atmospheric data to correctly recover the flux spatial covariance structure in **Q** is found to be strongly dependent on the relative temporal resolutions of the fluxes and observations. For cases where the time scales match (e.g. using 8-day average concentrations to infer the spatial covariance structure of 8-day average fluxes, or F8d/C8d), RML-Inv is able to recover the covariance parameters to within a factor of two for all cases, and much better in most cases. Conversely, when the temporal resolution of the concentrations is higher than that of the fluxes (e.g. using 3-hourly-average concentrations to infer the spatial covariance structure of 8-day average fluxes, F8d/C3h), RML-Inv is not able to satisfactorily recover covariance parameters. This indicates that consistent observation and flux resolutions must be used if the spatial covariance of the flux distribution is to be recovered directly from atmospheric observations, and that, if this guideline is followed, the atmospheric data are able to recover the spatial covariance of fluxes reasonably well for all the examined resolutions.

One caveat to the results presented here is that these RML-Inv tests for **Q** did not consider transport error. Preliminary tests showed that using RML-Inv to estimate spatial covariance parameters with added noise in the synthetic measurement vectors produced unstable results across all cases. However, the relatively few data points for one month may have contributed to this instability, and therefore, these tests should be repeated in future work using synthetic measurements created over longer time

periods.

4.2 Model-data mismatch variances (R)

For all examined cases, RML-Inv was able to satisfactorily recover the "true" temporal aggregation error, as well as the noise added to the measurements representing random transport model error (results not shown). Therefore, the atmospheric data along with the RML algorithm provide a powerful approach for correctly estimating model-data mismatch variances for atmospheric inversions.

The "true" temporal aggregation error, corresponding to the model-data mismatch caused by representing fluxes that vary at high temporal resolution (in this case, 3-hourly) as daily or 8-day average fluxes, is presented in Fig. 3 for all inversion cases except F3h/C3h (where the "true" values are zero). As expected, for all towers, the temporal aggregation error is highest when the coarsest fluxes are matched with the finest-resolution observations (F8d/C3h). The temporal aggregation error also tends to be higher for towers in highly active biospheric regions (e.g. LEF among the tall towers, HF among the short towers), where high temporal variability in nearby fluxes has a strong influence on measured concentrations.

For the F8d/C3h case, which shows the highest temporal aggregation error among the cases, this error is investigated at a 3-hourly scale for an average 24-hour period. The temporal aggregation errors for this case show a distinct diurnal cycle for the Short towers (Fig. 4), with highest values at night, and lowest values during the afternoon (18:00–24:00 UTC). Also, the errors for the Tall/MBL and Short towers are similar for afternoon observations, due to convective mixing at this time of the day which allows short towers to "see" flux variability at wider spatiotemporal scales. Conversely, at night, measurements from short towers lie within the stable nocturnal boundary layer, and are primarily influenced by near-field (in both space and time) fluxes. Therefore, given the higher aggregation errors associated with these measurements, their use in an inversion to recover 8-day average fluxes would contribute little information, and more importantly, might lead to significant errors in recovered fluxes. As previously

ACPD
9, 22407–22458, 2009

discussed, non-afternoon measurements shown in Fig. 3 for the Short towers are not used in any inversions except for two exceptions: the F3h/C3h case and a sensitivity test using the F8d/C3h case.

Inferred transport model errors, derived as the difference between the total model-data mismatch seen in actual measurements and temporal aggregation error, are presented in Table 4 for the different inversion cases and Tall/MBL vs. Short towers. As expected, the inferred transport errors are generally higher for Short relative to Tall/MBL towers and also when using higher-resolution measurement data. For example, the inferred error ranges from 2.1 to 2.5 ppm for the Tall/MBL towers when using 3-hourly concentrations vs. 1.5 to 1.8 ppm for these same towers when using daily or 8-day average data. Night-time transport model error is also inferred for the Short towers for the F3h/C3h case, and is found to be about 50% higher than the afternoon value.

Given that the cases for which the temporal aggregation error is highest correspond to the cases where the transport model error is lower, and vice versa, the impact of total model-data mismatch on the examined inversion cases cannot be established from these results alone. As such, this impact is explored using the six inversion cases presented in the next section.

4.3 Identification of best inversion setups

As described in Sect. 3.2, each of the inversion cases presented in Table 1 is run using synthetic measurement vectors with and without added noise to represent transport model errors. Results shown in Fig. 5 demonstrate that among the six "perfect transport" inversions, the F3h/C3h case has the lowest RMSE for all North American grid-cells, as well as for the subset of grid-cells in high sensitivity areas. Estimating daily fluxes also has a benefit over 8-day flux estimation in the high sensitivity areas, but for all North America, the difference between the daily and 8-day average flux cases is less pronounced. For reference, if the inversion were to exactly infer the mean monthly flux across the continent with no spatiotemporal variability, the RMSE would be 0.89 μmol/(m^2×s). Therefore, all "perfect transport" inversions perform better than

ACPD
9, 22407–22458, 2009

this baseline value.

As expected, the a posteriori RMSE increases for all cases when transport error is considered, although only the F8d/C8d case yielded an RMSE above the baseline for all North America. The F3h/C3h case shows the strongest impact due to transport errors, which is consistent with the large inferred errors for this case (shown in Table 4). Yet, for all North America, using 3-hourly measurements is associated with the minimal RMSE across flux resolutions, with the F8d/C3h and F3h/C3h case tying for the lowest values, and the Fd/C3h case a close second. Overall, when considering transport error, the RMSE values appear to be more similar across flux resolutions than in the "perfect transport" scenario, both within the high sensitivity areas and across all North America.

For the "perfect transport" and "transport error" inversions, the RMSE is lower in high-sensitivity areas for most cases as compared to the RMSE for the whole continent. However, despite good performance for the whole continent when considering transport error, the F8d/C3h case shows little reduction in RMSE for high-sensitivity areas. This result could be due to high temporal aggregation error associated with this case, which is particularly problematic in the near-field flux locations, but it may also point to a potential trade-off between minimizing grid-scale errors for the continent as a whole vs. in the near-field of measurements.

The fraction of true fluxes lying within two standard deviations (Eq. 5) of the estimated fluxes is presented in Table 5, with the first two columns representing the "perfect transport" and "transport error" inversions, respectively. In an ideal scenario, 95% of true fluxes should lie within this interval, with higher values representing overly conservative a posteriori error estimates, and lower values representing a posteriori uncertainty estimates that are too low. A posteriori uncertainties are found to accurately represent the true a posteriori errors for all cases for which 8-day fluxes are estimated directly. Conversely, the a posteriori uncertainties are far too low for cases in which the native resolution of the a posteriori fluxes is 3-hourly or daily. This result was found to be caused by failing to consider the temporal correlation in fluxes (or flux residuals from

the model of the trend) at these finer time scales (see Sect. 4.3).

Additional sensitivity tests are performed to investigate the value associated with including night-time data in the inversion for both Short and Tall/MBL towers. For the F8d/C3h case, including night-time data at the Short towers is found to increase the RMSE above the defined baseline value, from 0.81 to 1.05 μmol/(m^2×s) without transport error, and from 0.82 to 0.91 μmol/(m^2×s) with transport error. This substantial increase in RMSE shows that even without systematic transport errors due to the mischaracterization of night-time PBL height, using night-time measurements from shorter towers in the inversion increases overall bias in inferred fluxes.

The second set of sensitivity tests investigates the impact of removing night-time data for the Tall/MBL towers, so that only afternoon measurements are included for all nine towers. These tests were performed for the F8d/C3h, F8d/Cd and Fd/C3h cases. For the latter two cases, the RMSE is found to increase for inversions with and without transport error, both across the continent and in high sensitivity areas. However, for the F8d/C3h case, excluding night-time data with assumed transport error substantially improves performance in the near field, while the RMSE for the continent remains unchanged.

4.4 Refinement and analysis of two best inversion setups

The first set of experiments (in Sect. 4.1) concluded that, at least under perfect transport conditions, the spatial covariance parameters of the flux distribution could be successfully estimated from the atmospheric data as long as the flux and observation resolutions were consistent. The second set of experiments (in Sect. 4.2) identified the F8d/C3h and F3h/C3h cases as having the lowest RMSE for all North America among the "transport error" inversions. However the Fd/C3h case has a similarly low RMSE, and has substantially lower computational costs relative to the F3h/C3h case. Therefore, the Fd/C3h is chosen over the F3h/C3h case for further investigation. In this section, refinements to the covariance matrices are made for the F8d/C3h and Fd/C3h inversion setups (including night-time data for the Tall/MBL towers) to arrive at

ACPD
9, 22407–22458, 2009

an optimal recommended setup for regional inversions over North America.

4.4.1 Updated parameterization of covariance matrices

As previously mentioned, one of the keys for successful inversions is to correctly identify the a priori covariance structure of the flux (or flux residual) distribution in **Q**. The two "best" cases (F8d/C3h and Fd/C3h) involve the estimation of 8-day and daily average fluxes. In Sect. 4.1, it was found that the best way to use the atmospheric data with RML-Inv to obtain the flux covariance parameters for 8-day average fluxes was by using 8-day average concentrations (F8d/C8d), and for daily fluxes, it was best to use daily observations (Fd/Cd). Therefore, to simulate the best overall setups that could be implemented using real atmospheric observations without relying on external estimates of covariance parameters, the F8d/C8d and Fd/Cd RML-Inv cases are used to supply covariance parameters (shown in Table 3) for the F8d/C3h and Fd/C3h inversions, respectively. This substitution of cases is possible because the flux resolution remains consistent between the covariance estimation and inversion steps.

Temporal flux covariance is also incorporated into the **Q** matrix for the Fd/C3h case in order to help rectify the under-estimation of a posteriori uncertainties at the 8-day scale (as presented in Sect. 4.2 and Table 5). An RML-Krig analysis of daily CASA-GFEDv2 fluxes yielded an l parameter of 1.7 days (i.e. a correlation length of 5.1 days, which is consistent with the timescale of synoptic weather systems). This "true" temporal correlation parameter is incorporated into **Q** using Eq. (6).

In addition, the impact of allowing the model-data mismatch to vary among individual towers is investigated. Table 6 shows the inferred transport errors for each tower for the two examined cases. These results confirm that, even when examined individually, Tall/MBL towers generally exhibit lower transport errors relative to Short towers. The inferred transport error for LEF is slightly negative in the F8d/C3h case (but not for the Fd/C3h case), owing to the somewhat rough method for estimating transport error implemented in this study. This negative error was increased to 0.5 ppm for the following analysis.

ACPD
9, 22407–22458, 2009

4.4.2 Interpreting fluxes from the two best cases

The F8d/C3h and Fd/C3h inversions are repeated with the modifications to their covariance matrices described above, yielding further improvements. For the F8d/C3h case, the RMSE decreases from 0.82 to 0.78 μmol/(m^2×s), whereas for the Fd/C3h case, the RMSE decreases from 0.82 to 0.80 μmol/(m^2×s). These reductions in RMSE are primarily due to using specific model-data mismatch variances per measurement tower, as opposed to average variances across all Tall/MBL or Short towers. For the Fd/C3h case, the added assumption of a priori temporal flux covariance for the Fd/C3h case somewhat counteracts this reduction in RMSE. However, this case maintains an advantage relative to the F8d/C3h case in the highly-sensitive areas with an RMSE of 0.74 vs. 0.78 μmol/(m^2×s).

After accounting for temporal correlation between daily fluxes, a posteriori uncertainties at the 8-day scale rise to expected levels (see Table 5, column 3). In fact, the Fd/C3h best case captures 97% of the "true" fluxes within the 2σ confidence intervals, while the F8d/C3h best case captures 90%. This difference between cases is most likely due to the under- and over-estimation of the variances in **Q** inferred by RML-Inv for 8-day vs. daily fluxes, which in turn impacts the uncertainty estimation (see Eq. 5).

Grid-scale 8-day a posteriori fluxes from the two inversions as compared to the "true" 8-day CASA-GFEDv2 fluxes are shown in Fig. 6. Overall, both inversions detect large-scale patterns of sources and sinks seen in the true fluxes at the 8-day scale, although the F8d/C3h inversion is better able to detect neutral fluxes and net sources across the continent than the Fd/C3h case. As expected, the inferred fluxes show significantly less grid-scale spatial variability relative to the true fluxes, due to the sparse atmospheric network along with well-mixed air sampled at the towers, the spatial correlation structure in **Q**, and the lack of auxiliary environmental variables within the trend that would be included in a real data inversion (e.g. Gourdji et al., 2008).

Fluxes are recovered more precisely in the vicinity of observation towers, as reflected in the lower RMSE in the highly sensitive areas (as seen in Fig. 5). However, grid-

ACPD
9, 22407–22458, 2009

scale a posteriori uncertainties also grow as the atmospheric constraint decreases for far-field fluxes (Mueller et al., 2008), so that the 2σ confidence intervals around the inferred flux estimates should contain the "true" fluxes about 95% of the time regardless of distance from the measurement towers. Figure 7 presents maps of this relative error for both inversion cases, locating grid-cells where the "true" flux falls within one, two, or more than two standard deviations of the estimated fluxes. For both cases, the confidence intervals appear to be accurate for most of the under-constrained areas shown in Fig. 2, although areas with high relative errors are still evident in the boreal north, Central America and surprisingly in the eastern US in the near vicinity of multiple towers. The underestimation of uncertainties in well-constrained areas may be due to sharp gradients in the underlying fluxes which are smoothed out due to the spatial correlation length scales in **Q**. Consistent with the percent of "true" fluxes captured within the 2σ confidence intervals, the F8d/C3h case shows a higher proportion of fluxes outside both the 1σ and 2σ confidence intervals relative to the Fd/C3h case.

Figure 9 presents estimated fluxes and their uncertainties aggregated to ecoregions for all June 2004 (Fig. 8), showing that the 95% confidence intervals from both inversions successfully contain the true flux for about half of the regions. At this aggregated scale, the inversions perform well when the ecoregion-scale flux remains close to the mean North American flux, also shown in Fig. 9. This occurs because the inversion tends to revert to the model of the trend ($\mathbf{X}\boldsymbol{\beta}$) for this highly under-constrained problem, which in the simple setup presented here is represented as a simple monthly continental mean. However, among the two cases, the F8d/C3h case shows more variability across ecoregions, yielding estimates significantly closer to the truth (at 1σ) in the Tropics and Subtropics and the Desert and Xeric Shrubland, two highly under-constrained regions. In other regions, there is little difference in the aggregated fluxes between the two inversions, although an area-weighted RMSE at the ecoregion scale shows that the F8d/C3h case is slightly more accurate than the Fd/C3h case (10.4 vs. 11.9 $\mathrm{g\,C/(m^2 \times month)}$).

The poor performance of both cases relative to the "true" fluxes in ecoregions such

ACPD
9, 22407–22458, 2009

as the Tropics and Subtropics, the Temperate Broadleaf and Mixed Forests, and the Tundra would be alleviated with improvements in the atmospheric monitoring network in later years, and also with the inclusion of auxiliary data in real-data inversions (e.g. Gourdji et al., 2008) which can further help to refine the model of the trend ($\mathbf{X}\boldsymbol{\beta}$) in time and space. Yet with the simple inversion setup presented here and using only nine measurement towers unevenly spaced across the continent, both the Fd/C3h and F8d/C3h best cases estimate the mean North American flux to within 2σ of the "truth", with this latter case yielding a slight overestimate of the total continental sink for this month by only about 10%.

5 Summary and conclusions

The synthetic data inversion experiments presented in this work are designed to identify an optimal geostatistical inverse modeling setup for estimating high resolution North American carbon fluxes using a realistic observational network. The lessons learned, however, inform future regional grid-scale inversions using both the geostatistical and the synthesis Bayesian inversion approaches.

5.1 Covariance matrices

The RML-Inv approach, which could also be applied with some modifications (Michalak et al., 2005) to a Bayesian inversion setup, was shown to be an objective and accurate method for estimating model-data mismatch errors. For the a priori spatial flux covariance parameters, RML-Inv produced reasonable results as long as the temporal resolution of the observations was consistent with the temporal resolution of the fluxes for which the covariance parameters were being inferred. However, this result awaits further confirmation with simulated transport error over longer time periods. To appropriately select spatial covariance parameters in future real-data geostatistical inversions, it may be helpful to compare the results from RML-Inv using atmospheric

ACPD
9, 22407–22458, 2009

data with a range of results inferred by RML-Krig using fluxes from various biospheric models.

When estimating 3-hourly or daily-average fluxes, it was found to be important to include temporal correlation in the a priori flux covariance matrix, which allowed the inversions to yield more accurate a posteriori uncertainty estimates at aggregated temporal scales. Although not explored in this study, it may also be necessary to include temporal covariance for inversions estimating 8-day fluxes for periods longer than a single month.

Finally, estimating a separate model-data mismatch variance per measurement tower helped to improve performance relative to a setup using average values across Tall/MBL vs. Short towers. Preliminary tests also show that when including night-time data for Tall/MBL towers, using separate model-data mismatch values by time of day can help to improve results. As the measurement network expands in future years, estimating different model-data mismatch variances for each tower and time of day may not be feasible using RML-Inv, although an approach similar to that applied in Mueller et al. (2008) may prove beneficial, where a scaling parameter is instead estimated on initial estimates of the measurement variances.

5.2 Inversion setup

The current study demonstrates a benefit across flux resolutions, when considering random transport model error, to using atmospheric observations at sub-daily temporal resolutions to infer fluxes across the continent at both grid and aggregated scales. Overall, an inversion using 3-hourly average observations to estimate 8-day fluxes was found to be optimal for minimizing bias in inferred fluxes given the current measurement network. This case also has the advantage, relative to the cases estimating 3-hourly or daily fluxes, of lower computational cost, which would be advantageous for inversions spanning multiple months or years.

The benefit shown here associated with using sub-daily measurements in regional inversions is mostly consistent with the Law et al. (2002, 2003) studies. Law et al.

ACPD
9, 22407–22458, 2009

(2002) showed in a series of inversions with no model-data mismatch that RMSE was minimized when using sub-daily measurements to infer monthly mean fluxes. However, when considering transport error associated with incorrect phasing, averaging the synthetic measurements to synoptic scale (≈2–5 days) gave more accurate inversion results than using either sub-daily or daily average measurements (Law et al., 2003). Law et al. (2003) also considered transport error associated with changes in the magnitude of non-baseline events, which is more analogous to the type of random error considered here. For this type of transport error, as well as measurement calibration errors, they found that it was still best to use higher-resolution measurements (i.e. 4-hourly) relative to averaged synoptic-scale measurements. The appropriate concentration averaging interval is most likely relative to the spatial and temporal scale of flux estimation, as well as the specific transport model used. Regardless, transport errors due to phasing as well as other systematic biases in WRF/STILT should be further investigated in future work.

Night-time data from Short towers lying within the nocturnal boundary layer were found to be subject to high temporal aggregation and transport model errors. Consistently, using these measurements in an inversion to estimate 8-day average fluxes was shown to increase grid-scale bias. Whether to include night-time data in future inversions for the Tall/MBL towers is less clear. For most examined cases, this data was found to improve inversion performance. However, for the F8d/C3h case, a case with high temporal aggregation error, night-time data was found to bias near-field fluxes, although it proved beneficial in under-constrained areas, particularly the Tropics and Subtropics and Desert and Xeric Shrubland ecoregions. The choice of whether to include night-time data for Tall/MBL towers in future regional inversions will need to be re-evaluated in the context of an expanding measurement network (e.g. ≈40 operational calibrated CO_2 measurement towers in 2008; A. Andrews, personal communication) which will make more grid-cells part of the "near-field", as well as persistent night-time transport model biases, which could potentially further reduce the value of including these measurements in an inversion.

5.3 Future work

This study employed a very simple model of the trend that did not include any auxiliary environmental variables. This choice was made in order to focus solely on the atmospheric constraint provided by continuous measurements from the 9 towers operational in 2004. However, as previously shown in Gourdji et al. (2008), the model of the trend used in the geostatistical inverse modeling approach can include variables such as satellite-derived vegetation indices, analyzed environmental data from numerical weather prediction models, and fossil fuel and agricultural inventories, among others. Auxiliary variables help to inform possible spatial patterns in carbon fluxes, but more importantly they provide information about processes that may be controlling flux variability, particularly important for adjusting fluxes in under-constrained regions.

The use of such auxiliary variables in future real data geostatistical inversions will help to address some of the limitations associated with the synthetic data inversions presented here. First, these variables help to capture sharper features in the flux field (see Gourdji et al., 2008) than what is visible through the atmospheric data alone (and shown in Fig. 6). Second, as the model of the trend is able to explain more of the variability in the recovered fluxes, the flux residuals decrease in magnitude, thereby reducing the influence of the spatiotemporal correlation structure in $\mathbf{Q}$ on the final flux estimates. This will be especially important when point source fossil fuel emissions are estimated in addition to more smoothly-varying biospheric fluxes.

In summary, the results from the synthetic data experiments presented in this work suggest that even a fairly sparse network of continuous CO_2 measurements, used with no auxiliary information or prior estimates of flux variability in time or space, can be used to infer relatively unbiased estimates of the net continental CO_2 surface flux over North America at an 8-day temporal resolution. The incorporation of additional atmospheric data and auxiliary variables in future real data geostatistical inversions can only help to further improve the recovery of fluxes at finer spatial resolutions.

ACPD
9, 22407–22458, 2009

Acknowledgements. This work was supported by the National Aeronautics and Space Administration under Grant No. NNX06AE84G "Constraining North American Fluxes of Carbon Dioxide and Inferring Their Spatiotemporal Covariances through Assimilation of Remote Sensing and Atmospheric Data in a Geostatistical Framework" issued through the ROSES A.6 North American Carbon Program. In addition, S. Gourdji received support from the NASA Earth System Science Fellowship.

We would especially like to thank J. Eluskiewicz, J. Henderson and T. Nehrkorn at Atmospheric and Environmental Research (AER) Corporation for their work in customzing WRF/STILT. The WRF v2.2 model was initialized with and used analysis nudging from the 32-km North American Regional Reanalysis from the National Centers for Environmental Prediction for the outermost 40-km grid. All nested grids were run with 30 vertical levels going from the surface to approximately 20.5 km altitude, with approximately 8 levels below 1000 m.

The authors would also like to thank the following scientists for sharing their atmospheric measurements: M. Torn and M. Fischer (LBNL), D. Worthy (Environment Canada), and S. Wofsy (Harvard University). We also thank M. Rodell for making the GLDAS dataset available to the community, as well as W. Peters for helpful input in the early stages of this study. Finally, the PUORG research group gave helpful feedback throughout the course of this work.

References

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y. H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, Global Biogeochem. Cy., 20, GB1002, doi:10.1029/2004GB002439, 2006.

Bakwin, P. S., Tans, P. P., Hurst, D., and Zhao, C.: Measurements of carbon dioxide on very tall towers: Results of the NOAA/CMDL program, Tellus, 50B, 401–415, 1998.

Carouge, C., Bousquet, P., Peylin, P., Rayner, P. J., and Ciais, P.: What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes – Part 1: Potential of the network, Atmos. Chem. Phys. Discuss., 8, 18591–18620, 2008a, http://www.atmos-chem-phys-discuss.net/8/18591/2008/.

Carouge, C., Peylin, P., Rayner, P. J., Bousquet, P., Chevallier, F., and Ciais, P.: What can we learn from European continuous atmospheric CO2 measurements to quantify regional fluxes – Part 2: Sensitivity of flux accuracy to inverse setup, Atmos. Chem. Phys. Discuss., 8, 18621–18649, 2008b, http://www.atmos-chem-phys-discuss.net/8/18621/2008/.

Cressie, N. A. C.: Statistics for Spatial Data, Wiley, J., New York, 900 pp., 1991.

Denning, A. S., Collatz, G. J., 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: 1. Surface carbon fluxes, Tellus, 48B, 521–542, 1996.

Geels, C., Doney, S. C., Dargaville, R., Brandt, J., and Christensen, J. H.: Investigating the sources of synoptic variability in atmospheric CO_2 measurements over the Northern Hemisphere continents: a regional model study, Tellus, 56B, 35–50, 2004.

Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A. T., Dargaville, R., Aalto, T., Brandt, J., Christensen, J. H., Frohn, L. M., Haszpra, L., Karstens, U., Rödenbeck, C., Ramonet, M., Carboni, G., and Santaguida, R.: Comparing atmospheric transport models for future regional inversions over Europe – Part 1: mapping the atmospheric CO_2 signals, Atmos. Chem. Phys., 7, 3461–3479, 2007, http://www.atmos-chem-phys.net/7/3461/2007/.

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: 2. Analysis of COBRA data using a receptor-oriented framework, J. Geophys. Res., 108(D24), 4757, doi:10.1029/2003JD003770, 2003.

Gerbig, C., Lin, J. C., Munger, J. W., and Wofsy, S. C.: What can tracer observations in the continental boundary layer tell us about surface-atmosphere fluxes?, Atmos. Chem. Phys., 6, 539–554, 2006, http://www.atmos-chem-phys.net/6/539/2006/.

Gerbig, C., Körner, S., and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, Atmos. Chem. Phys., 8, 591–602, 2008, http://www.atmos-chem-phys.net/8/591/2008/.

Gourdji, S. M., Mueller, K. L., Schaefer, K., and Michalak, A. M., Global monthly-averaged CO_2 fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, J. Geophys. Res., 113, D21115, doi:10.1029/2007JD009733, 2008.

Gurney, K., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y. H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J.,

Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C. W.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415(6872), 626–630, 2002.

Haszpra, L.: On the representativeness of carbon dioxide measurements, J. Geophys. Res., 104(D21), 26953–26960, 1999.

Huntzinger, D. N., Michalak, A. M., Gourdji, S. M., and Mueller, K. L.: An expanded approach for assessing and comparing variability of biospheric carbon flux estimates with an application to North America, Geophys. J. Res.-Biogeosci., in review, 2009.

Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three-dimensional global inverse model of the atmospheric transport, 2. Inversion of the transport of CO_2 in the 1980's, J. Geophys. Res., 104(D15), 18555–18581, 1999.

Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G., On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106, 4703–4715, 2001.

Kitanidis, P. K.: Introduction to Geostatistics: Applications in Hydrogeology, Cambridge University Press, New York, 249 pp., 1997.

Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K. J., Ciais, P., Denning, A. S., and Rayner, P. J.: Mesoscale inversion: first results from the CERES campaign with synthetic data, Atmos. Chem. Phys., 8, 3459–3471, 2008, http://www.atmos-chem-phys.net/8/3459/2008/.

Law, R. M., Rayner, P. J., Steele, L. P., and Enting, I. G.: Using high temporal frequency data for CO_2 inversions, Global Biogeochem. Cy., 4, 1053, doi:10.1029/2001GB001593, 2002.

Law, R. M., Rayner, P. J., Steele, L. P., and Enting, I. G.: Data and modelling requirements for CO_2 inversions using high frequency data, Tellus, 55B, 512–521, doi:10.1034/j.1600-0560.2003.0029.x, 2003.

Law, R. M., Rayner, P. J., and Wang, Y. P.: Inversion of diurnally-varying synthetic CO_2: network optimisation for an Australian test case, Global Biogeochem. Cy., 18, GB1044, doi:10.1029/2003GB002136, 2004.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: the Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res., 108(D16), 4493, doi:10.1029/2002JD003161, 2003.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Grainger, C. A., Stephens, B. B., Bakwin, P. S., and Hollinger, D. Y.: Measuring fluxes of trace gases at

regional scales by Lagrangian observations: application to the CO_2 Budget and Rectification Airborne (COBRA) study, J. Geophys. Res., 109, D15304, doi:10.1029/2004JD004754, 2004.

Matross, D. M., Andrews, A., Pathmathevan, M., Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Gottlieb, E. W., Lee, J. T., Zhao, C., Bakwin, P. S., Munger, J. W., and Hollinger, D. Y.: Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based, and satellite data, Tellus, 57B, 344–358, 2006.

Michalak, A. M., Bruhwiler, L., and Tans, P. P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, J. Geophys. Res., 109, D14109, doi:10.1029/2003JD004422, 2004.

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, J. Geophys. Res., 110, D24107, doi:10.1029/2005JD005970, 2005.

Mueller, K. L., Gourdji, S. M., and Michalak, A. M.: Global monthly-averaged CO_2 fluxes recovered using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements, J. Geophys. Res., 113, D21114, doi:10.1029/2007JD009734, 2008.

Nehrkorn, T., Eluszkiewicz, J., Wofsy, S. C., Lin, J. C., Gerbig, C., Longo, M., and Freitas, S., Coupled Weather Research and Forecasting/Stochastic Time-Inverted Lagrangian Transport (WRF/STILT) model, Meteorol. Atmos. Phys., submitted, 2009.

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

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., and Kassem, K. R.: Terrestrial ecoregions of the world: a new map of life on earth, Bioscience, 51(11), 933–938, 2001.

Patra, P. K., Law, R. M., Peters, W., Rödenbeck, C., Takigawa, M., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C., S.-J. Lin, Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Vautard, R., Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of

hourly atmospheric CO_2: Analysis of synoptic-scale variations for the period 2002–2003, Global Biogeochem. Cy., 22, GB4013, doi:10.1029/2007GB003081, 2008.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, 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, P. Natl. Acad. Sci., 104(48), 18925–18930, 2007.

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., Dalu, G. A., Snook, J. S., Lee, T. J., and Kittel, T. G. F.: Nonlinear influence of mesoscale land-use on weather and climate, J. Climate, 4(11), 1053–1069, 1991.

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

Rodell, M., Houser, P. R., Jambor, U., et al.: The Global Land Data Assimilation System, B. Am. Meteorol. Soc., 85, 381–394, 2004.

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/.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 2. Technical Note 468+STR, MMM Division, NCAR, Boulder, CO, 88 pp., online available at: http://www.mmm.ucar.edu/wrf/users/docs/arw_v2.pdf, 2005.

Stull: Transilient turbulence theory: 1. the concept of eddy-mixing across finite distances, J. Atmos. Sci., 41(23), 3351–3367, 1984.

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

Table 1. Inversion cases considered in this study.

Case	Flux resolution	Observation resolution
F3h/C3h*	3-hourly	3-hourly
Fd/C3h	Daily	3-hourly
Fd/Cd	Daily	Daily
F8d/C3h	8-day	3-hourly
F8d/Cd	8-day	Daily
F8d/C8d	8-day	8-day

* This case uses 24 hours of atmospheric observations for all towers, while all other cases use only afternoon (18:00–24:00 UTC) data for the five Short towers (see Table 2).

ACPD
9, 22407–22458, 2009

Table 2. Measurement locations.

Tower	Location	Coordinates	Height	Maintained by	Type
LEF	Park Falls, Wisconsin	45.93 N, 90.27 W	396 m	NOAA/GMD	Tall
WKT	Moody, Texas	31.32 N, 97.33 W	457 m	NOAA/GMD	Tall
BRW	Barrow, Alaska	71.32 N, 156.60 W	10 m	NOAA/GMD	MBL
SBL	Sable Island, Nova Scotia	43.93 N, 60.02 W	25 m	Met Service Canada	MBL
AMT	Argyle, Maine	45.03 N, 68.68 W	107 m	NOAA/GMD	Short
ARM	Norman, Oklahoma	36.62 N, 97.50 W	60 m	US Dept. of Energy	Short
CDL	Candle Lake, Saskatchewan	53.99 N, 105.12 W	30 m	Met Service Canada	Short
FRD	Fraserdale, Ontario	49.84 N, 81.52 W	40 m	Met Service Canada	Short
HF	Petersham, Massachusetts	42.54 N, 72.17 W	30 m	Harvard University	Short

ACPD
9, 22407–22458, 2009

Table 3. A priori flux covariance parameters estimated using RML-Krig, which are considered the true values, and RML-Inv, which are obtained from the synthetic atmospheric data. RML-Krig was run only once for the daily and 8-day flux resolutions (i.e. Fd, F8d).

Case	Variance σ^2 [$(\mu\text{mol}\,(\text{m}^{-2}\,\text{s}^{-1}))^2$]		Range parameter l_x [km]	
	RML-Krig	*RML-Inv*	*RML-Krig*	*RML-Inv*
F3h/C3h "day"	29.2	53.7	590	510
F3h/C3h "night"	4.7	4.2	740	580
F3h/C3h "shelf"	10.7	16.8	390	710
Fd/C3h	1.9	7.0	560	260
Fd/Cd		3.5		470
F8d/C3h	0.8	10.2	620	80
F8d/Cd		2.2		240
F8d/C8d		0.6		600

ACPD
9, 22407–22458, 2009

Table 4. Inferred Tall/MBL and Short tower transport error standard deviations for inversion cases.

Case	Transport error [ppm]	
	Tall/MBL towers	*Short towers*
F3h/C3h	2.5	4.4
Fd/C3h	2.2	3.9
Fd/Cd	1.8	3.8
F8d/C3h	2.1	3.7
F8d/Cd	1.5	3.6
F8d/C8d	1.8	2.3

Table 5. Percentage of true fluxes within two standard deviations of all North American a posteriori grid-scale fluxes. The theoretical value is 95% for all cases.

Case	No transport error [%]	With transport error [%]	Refined inversions with transport error [%]
F3h/C3h	72	64	
Fd/C3h	76	77	97
Fd/Cd	77	74	
F8d/C3h	92	93	90
F8d/Cd	94	95	
F8d/C8d	94	95	

ACPD
9, 22407–22458, 2009

Table 6. Inferred transport error standard deviations for each tower for the Fd/C3h and F8d/C3h inversion cases.

Tower	Fd/C3h [ppm]	F8d/C3h [ppm]
LEF	1.5	0.0
WKT	3.8	3.7
BRW	1.1	1.1
SBL	2.1	2.2
AMT	4.2	4.8
ARM	4.0	3.9
CDL	1.0	1.7
FRD	3.1	3.1
HF	5.3	5.3

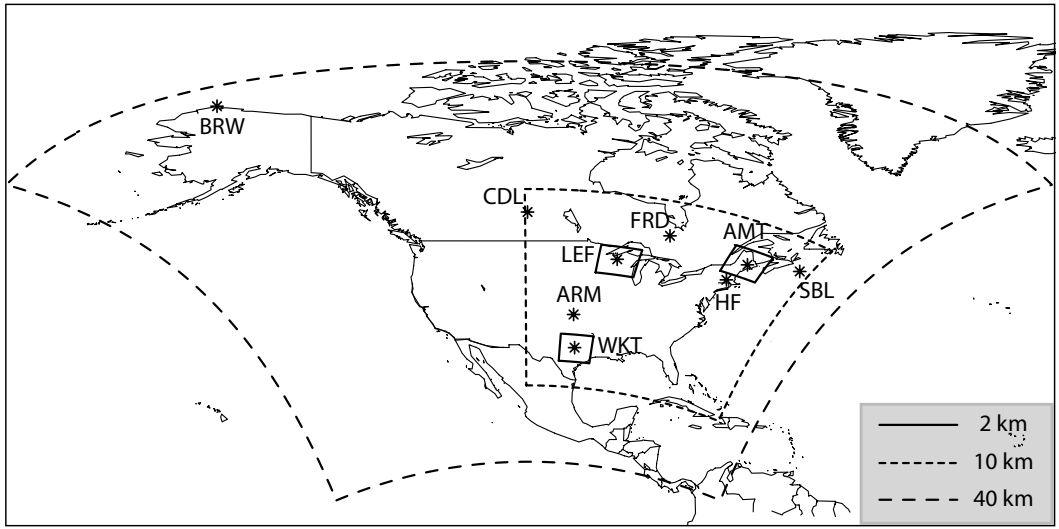

Fig. 1. Measurement tower locations and nested 2, 10 and 40 km WRF grid domains.

ACPD
9, 22407–22458, 2009

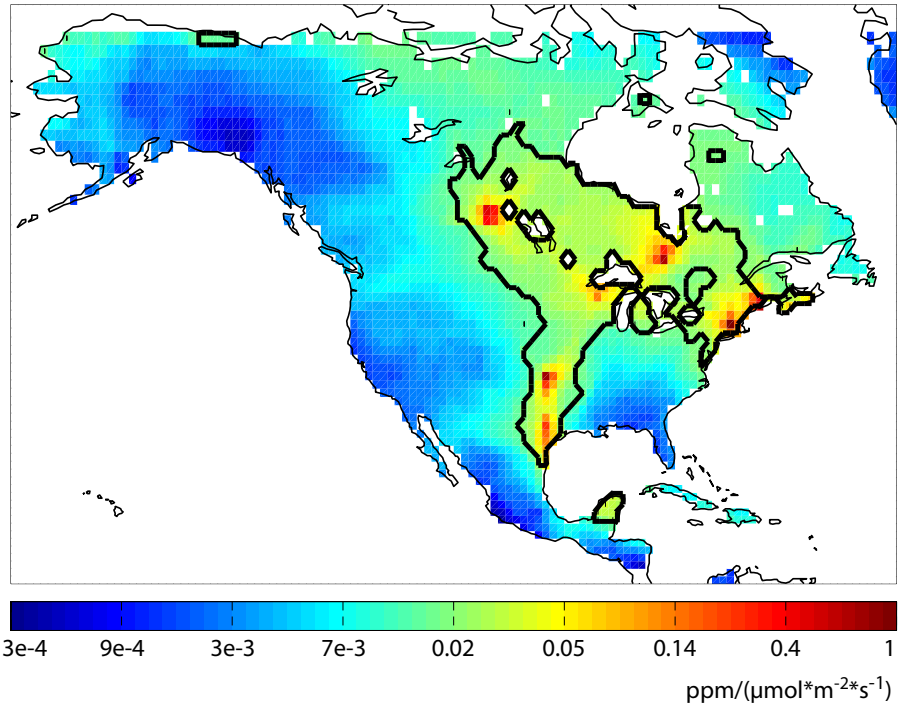

Fig. 2. Average concentration footprint for the 9 towers for the month of June, 2004, shown on a logarithmic scale. Black contour line identifies those grid-cell locations in the top 25% percentile of monthly average sensitivity to measurements.

ACPD
9, 22407–22458, 2009

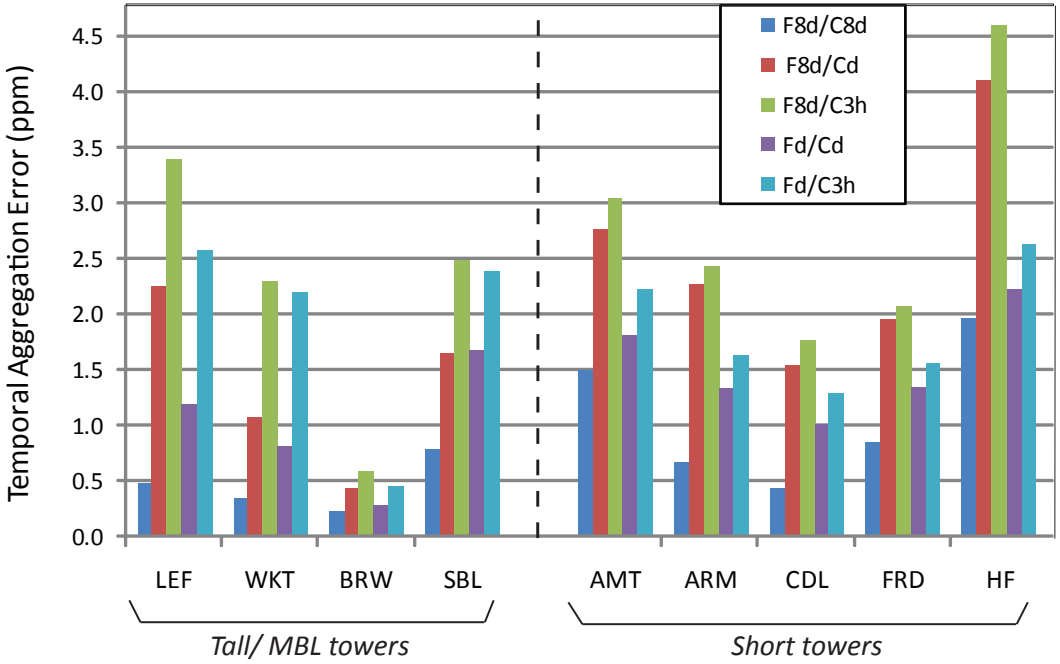

Fig. 3. Temporal aggregation error for 6 cases by tower.

ACPD
9, 22407–22458, 2009

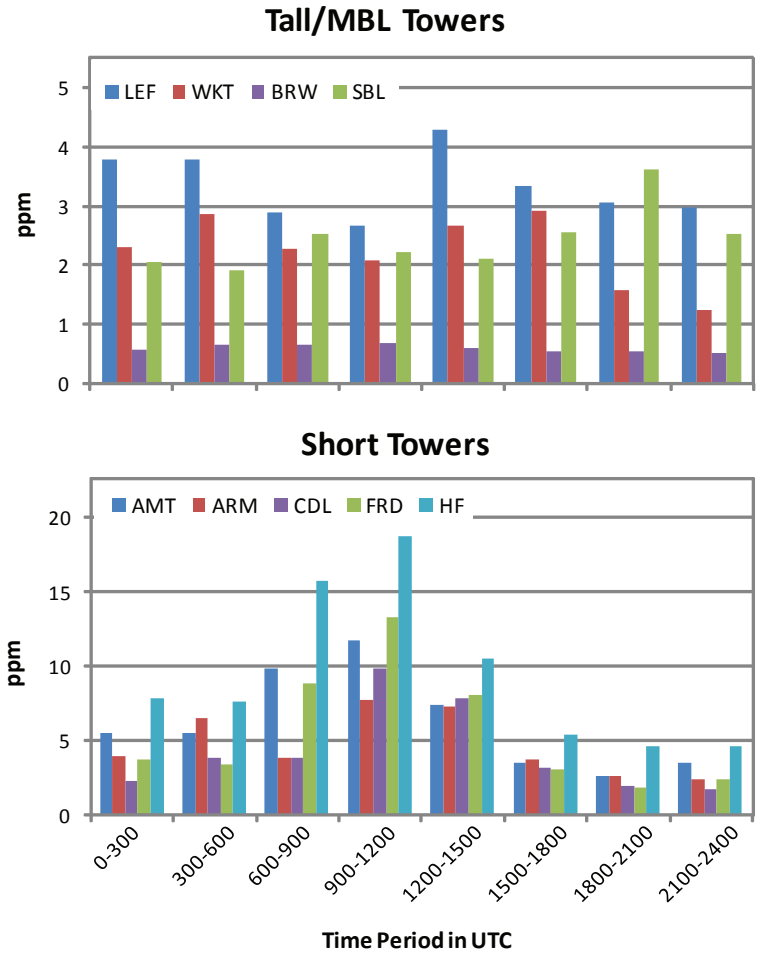

Fig. 4. Diurnal cycle of temporal aggregation error for the F8d/C3h case for Tall/MBL and Short towers. Please note different scales for the different types of towers.

ACPD
9, 22407–22458, 2009

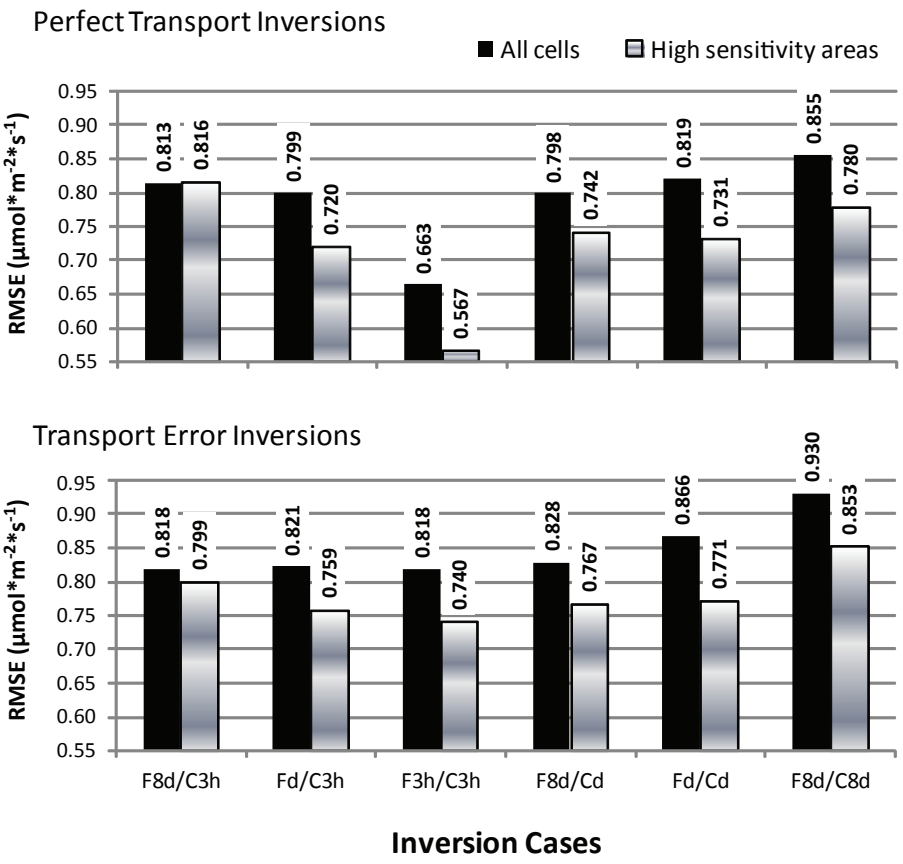

Fig. 5. A posteriori RMSE at the 8-day grid-scale comparing fluxes from the "perfect transport" and "transport error" inversions to the true CASA-GFEDv2 fluxes. RMSE is shown for all North American grid-cells, as well as for the subset of grid-cells in the high sensitivity areas identified in Fig. 2.

ACPD
9, 22407–22458, 2009

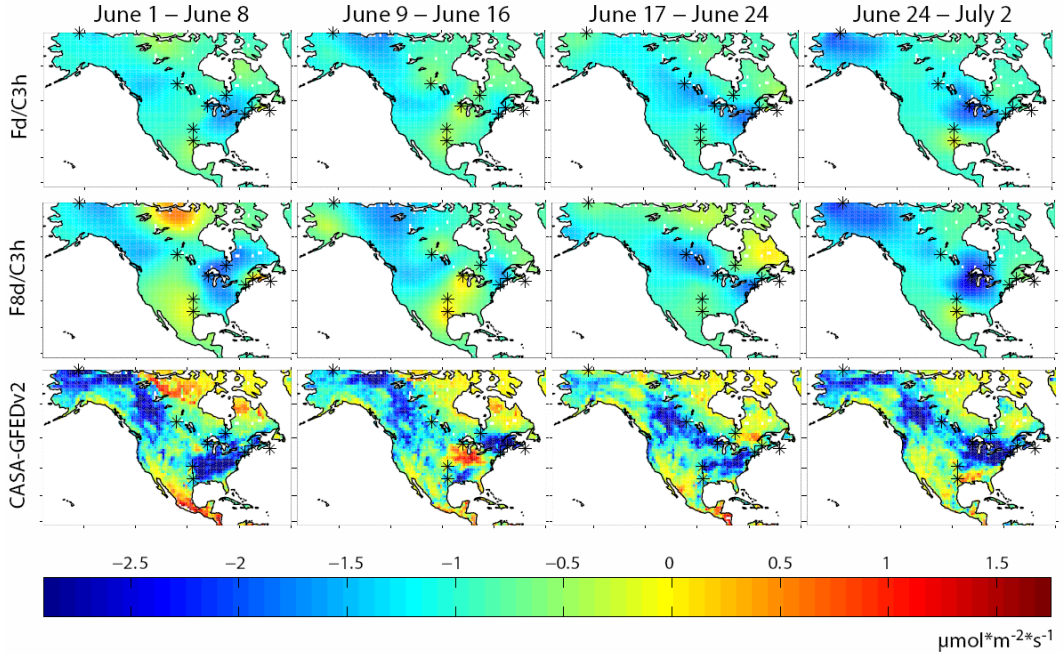

Fig. 6. Grid-scale fluxes from the two "best" inversion cases (Fd/C3h and F8d/C3h) for the four 8-day periods from 1 June to 2 July, 2004 as compared to the "true" CASA-GFEDv2 fluxes.

ACPD
9, 22407–22458, 2009

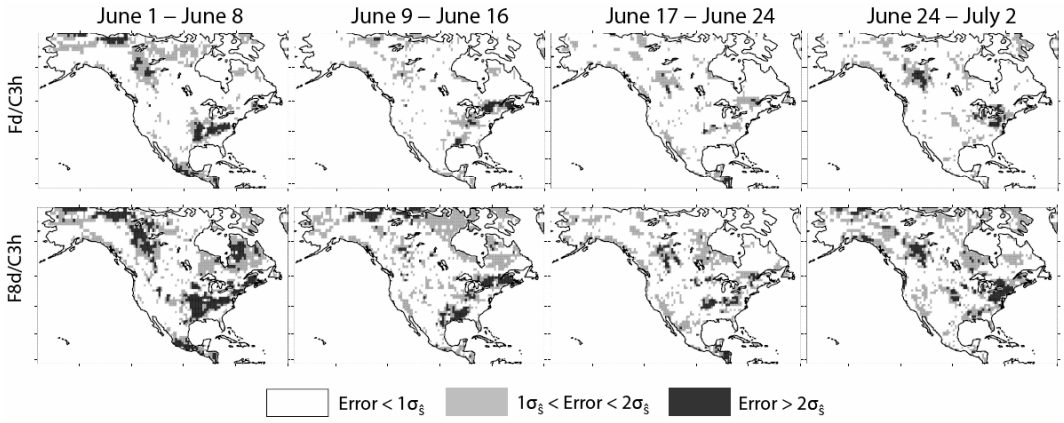

Fig. 7. Map showing the normalized error for the grid-scale a posteriori fluxes from the two "best" inversion cases (Fd/C3h and F8d/C3h), i.e. which fluxes are within 1σ or 2σ intervals or more than 2σ away from the true fluxes for the four 8-day periods.

ACPD
9, 22407–22458, 2009

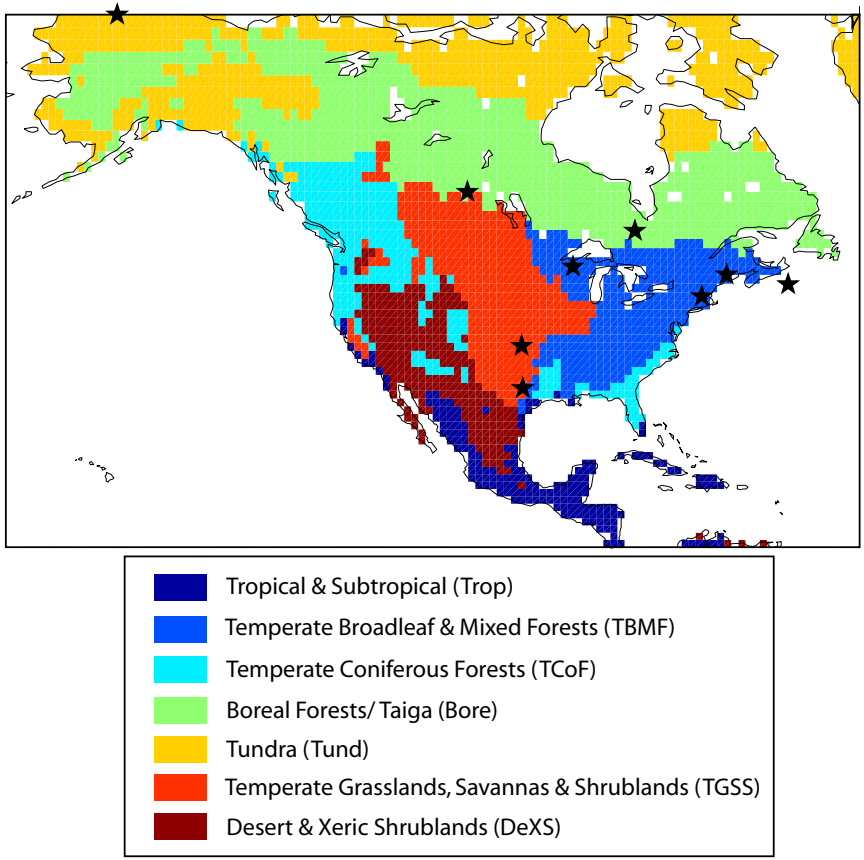

Fig. 8. Ecoregion map, modified from Olson et al. (2001), used for spatial aggregation of a posteriori fluxes and uncertainties.

ACPD
9, 22407–22458, 2009

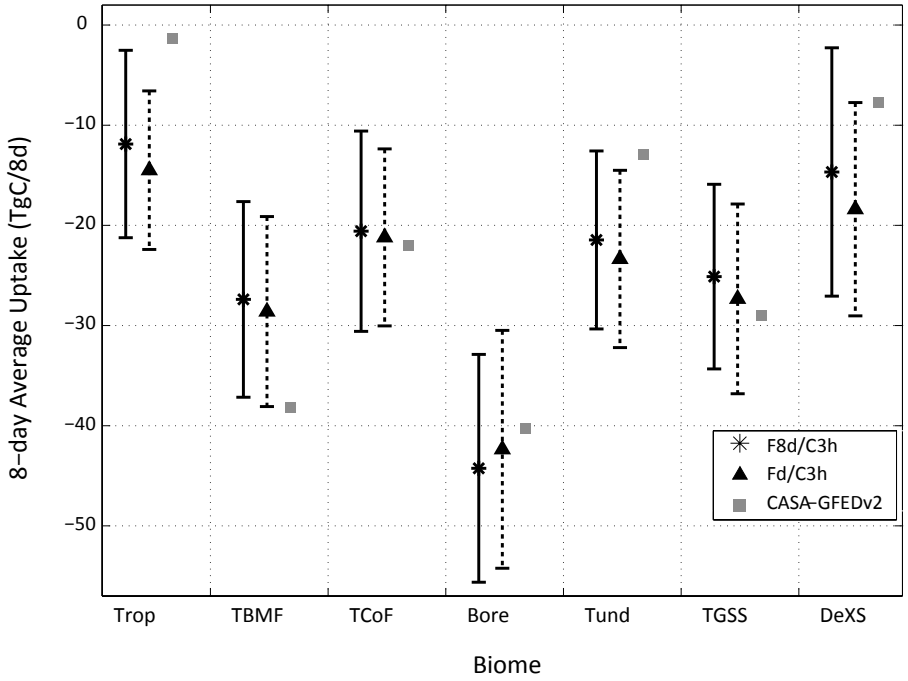

Fig. 9. Comparison of total 8-day average fluxes and uncertainties ($\pm 2\sigma$) for different ecoregions as well as the entire North American continent, aggregated a posteriori for the two best inversion cases and "true" CASA-GFEDv2 fluxes.

ACPD
9, 22407–22458, 2009