Sensitivity of a subregional scale atmospheric inverse CO_2 modeling framework to boundary conditions

Mathias Göckede,[1] David P. Turner,[1] Anna M. Michalak,[2,3] Dean Vickers,[4] and Beverly E. Law[1]

Received 4 May 2010; revised 14 September 2010; accepted 27 September 2010; published 23 December 2010.

[1] We present an atmospheric inverse modeling framework to constrain terrestrial biosphere CO_2 exchange processes at subregional scales. The model is operated at very high spatial and temporal resolution, using the state of Oregon in the northwestern United States as the model domain. The modeling framework includes mesoscale atmospheric simulations coupled to Lagrangian transport, a biosphere flux model that considers, e.g., the effects of drought stress and disturbance on photosynthesis and respiration CO_2 fluxes, and a Bayesian optimization approach. This study focuses on the impact of uncertainties in advected background mixing ratios and fossil fuel emissions on simulated flux fields, both taken from external data sets. We found the simulations to be highly sensitive to systematic changes in advected background CO_2, while shifts in fossil fuel emissions played a minor role. Correcting for offsets in the background mixing ratios shifted annual CO_2 budgets by about 47% and improved the correspondence with the output produced by bottom-up modeling frameworks. Inversion results were robust against shifts in fossil fuel emissions, which is likely a consequence of relatively low emission rates in Oregon.

Citation: Göckede, M., D. P. Turner, A. M. Michalak, D. Vickers, and B. E. Law (2010), Sensitivity of a subregional scale atmospheric inverse CO_2 modeling framework to boundary conditions, *J. Geophys. Res.*, *115*, D24112, doi:10.1029/2010JD014443.

1. Introduction

[2] Understanding the controls and mechanisms that drive the terrestrial biospheric carbon cycle is important for scientists and decision makers dealing with mitigation and adaptation measures toward climate change [*IPCC*, 2007]. Based on a growing global network of observation sites, modeling frameworks have been developed that assimilate and integrate various data sources to produce spatially explicit maps of carbon exchange between biosphere and atmosphere [e.g., *Wang et al.*, 2009; *Williams et al.*, 2009]. These models make it possible to constrain wall-to-wall carbon budgets in regional to global domains and help to improve the interpretation of carbon flux variability on time scales from days to decades [*Stoy et al.*, 2009], advancing our understanding on the underlying processes and thus our ability to forecast future climate scenarios [e.g., *Friedlingstein et al.*, 2006].

[3] Top-down atmospheric inversion models [e.g., *Enting*, 2005] have been widely applied to investigate CO_2 exchange processes, with an increasing number of model implementations covering a large range of temporal and spatial scales [e.g., *Gourdji et al.*, 2008; *Gurney et al.*, 2002; *Matross et al.*, 2006; *Mueller et al.*, 2008; *Peters et al.*, 2007; *Peylin et al.*, 2005]. These models constrain surface processes from the atmospheric point of view [e.g., *Rayner et al.*, 1999; *Tans et al.*, 1990], applying sophisticated mathematical tools [e.g., *Michalak et al.*, 2004; 2005; *Peters et al.*, 2005] to determine spatially and temporally varying flux patterns that match the observed variations in atmospheric observations of CO_2 mixing ratios. The terrestrial biosphere CO_2 exchange at the surface is usually simulated through biogeochemical models of various degrees of sophistication. Atmospheric inverse modeling has been demonstrated to provide net carbon budgets for large scale domains with coarse spatial resolution, but resolving regional or smaller scale patterns in surface fluxes that are required to improve carbon cycle process understanding still remains challenging. Major obstacles when aiming at refining top-down model output are the sparseness of available observation networks [*Gurney et al.*, 2002, 2004], uncertainties associated with atmospheric transport and mixing [e.g., *Baker et al.*, 2006; *Prather et al.*, 2008], and high computational demand associated with a large number of cells in high-resolution surface grids.

[4] Atmospheric top-down modeling at subregional scales [*Lauvaux et al.*, 2008; *Matross et al.*, 2006] holds the

[1]Department of Forest Ecosystems and Society, College of Forestry, Oregon State University, Corvallis, Oregon, USA.
[2]Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan, USA.
[3]Department of Atmospheric, Oceanic and Space Science, University of Michigan, Ann Arbor, Michigan, USA.
[4]College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA.

0148-0227/10/2010JD014443

Table 1. Surface Characteristics in Oregon and Data Sources

Data Type	Number of Classes	Classes	Data Source
Ecoregion	10	Coastal Range (CR), Klamath Mountains (KM), Willamette Valley (WV), West Cascades (WC), Cascade Crest (CC),[a] East Cascades (EC), Columbia Plateau (CP), Blue Mountains (BM), Northern Basin and Range (NB), Snake River (SR)	EPA Level III Ecoregions (http://www.epa.gov/wed/pages/ecoregions.htm)
Vegetation Land Cover Types	6	Evergreen Needleleaf Forest/Mixed Forest (ENF/MF), Deciduous Broadleaf Forest (DBF), Juniper Woodland (WOOD), Shrubland (SHRUB), Grassland (GRASS), Cropland (CROP)	Forest Types, *Law et al.* [2004]; Nonforest Types, USGS NLCD[b] (http://landcover.usgs.gov/nlcd.php)[c]
Disturbance Regime	2	Wildfire (FIRE), Harvest/Clearcut (CUT)	*Law et al.* [2004], *Cohen et al.* [2002]

[a]Ecoregion added for this study as transition zone between WC and EC.
[b]USGS NLCD, U.S. Geological Survey National Land Cover Dataset.
[c]Compared to the original data, some classes herein were combined and reclassified.

potential to provide high-resolution flux products that can be used to improve understanding of mechanisms controlling terrestrial biosphere CO_2 exchange processes. It is the resolution needed for resolving sources and sinks of CO_2 within countries and states. The reduced computational demand associated with smaller domain sizes makes it possible to increase the spatial resolution of the atmospheric inversion setup. For example, the inversion in *Matross et al.* [2006] covered an area of 35° × 90° with resolution 1/6° latitude and 1/4° longitude, and *Lauvaux et al.* [2008] used a domain of 300 × 300 km at 8 km resolution. In addition, although such subregional inversions still estimate fluxes at scales that are large relative to landscape-scale variability, the underlying fluxes below the inversion resolution can be defined at the disturbance level scale (<1 km), instead of aggregating larger areas that combine different biomes and disturbance conditions.

[5] With high-resolution inversion setups, the information extracted from observations of atmospheric CO_2 mixing ratio time series can be used to train a spatially representative version of a biosphere flux model that captures the basic mechanisms and controls of the surface-atmosphere carbon exchange processes, such as the paramount importance of capturing ecosystem response to disturbance [e.g., *Law*, 2005; *Running*, 2008]. However, finer spatial resolutions increase uncertainties in atmospheric transport and mixing, which might confound information from highly resolved surface flux fields [e.g., *Dolman et al.*, 2009]. Also, restricting the model domain to subregional (<500 km) scales requires nesting the model domain into larger scale grids to solve for the variability in boundary conditions, i.e., the advected background CO_2 mixing ratios [e.g., *Rödenbeck et al.*, 2009]. The operation of a top-down modeling framework in high spatial resolution requires input of a highly resolved fossil fuel emission database and its uncertainties to consider the impact of anthropogenic CO_2 emissions on the atmospheric observations.

[6] In this study we present an atmospheric inverse modeling framework operated at a subregional scale, building on high-resolution surface maps of CO_2 emission sources with an effective spatial resolution of <1 km. The inversion optimizes individual flux base rate parameters for a total of 68 surface types in Oregon, defined as combinations of ecoregions, land cover types, and disturbance regimes within the domain. The domain of our study is the Pacific Northwest region of the United States, which is dominated by incoming "clean" air masses from the Pacific Ocean. Here we implement major changes to the modeling framework that was introduced by *Göckede et al.* [2010]: (1) inclusion of Lagrange multipliers into the optimization to enforce plausible parameter ranges; (2) addition of a variogram analysis to consider spatial correlations among prior base rate uncertainties; (3) change of the data source for interpolated surface meteorology; (4) addition of a soil water scalar to improve simulation of drought stress effects on CO_2 assimilation. Our objectives are to examine model sensitivity to the influence of advected background CO_2 acquired from the global CarbonTracker database [*Peters et al.*, 2007], and to test the model sensitivity to fossil fuel emissions estimates from the high-resolution national database VULCAN [*Gurney et al.*, 2009a]. All simulations were conducted using two optimization strategies, one assigning parameters that are constant throughout the year and the other using seasonally varying parameter sets.

2. Methods

[7] The general concept of the atmospheric inverse modeling framework employed in this study has been developed by *Gerbig et al.* [2003]. An in-depth description of the adaptation and refinement of this framework for modeling the Oregon domain in the context of this study is given by *Göckede et al.* [2010]. The following sections will briefly outline the major characteristics of the approach and the improvements in some of the components that have been implemented since the work of *Göckede et al.* [2010].

2.1. Model Domain and Observational Data

[8] This study aims at constraining the CO_2 exchange processes between the terrestrial biosphere and the atmosphere for the state of Oregon, USA. The dimensions of Oregon are 580 km east–west and 420 km north–south, with a total area of about 2.5×10^5 km^2. The distinct climate zones are the more mesic and populated western third of the

Table 2. List of Observation Sites Used in This Study

Site Name	Site Code	Measurement Height[a] (m agl)	Elevation (Tower Base) (m asl)	Latitude Position (deg)	Longitude Position (deg)
Yaquina Head	YH	12.5	115	44.67	−124.07
Mary's Peak	MPk	7.5	1248	44.50	−123.55
Mary's River Mature Fir	MF	37.9	262	44.65	−123.55
Metolius Mature Pine	MP	33.5	1253	44.45	−121.56
Burns	BU	6.0	1397	43.47	−119.69

[a]agl, above ground level.

state and the sparsely populated eastern portion with a continental, semiarid climate. They are separated by the Cascade Mountains 200 km inland and parallel to the Pacific coastline. The western ecoregions support dense coniferous forests and agriculture. The eastern ecoregions are dominated by juniper-sagebrush-grass communities adapted to the semiarid climate. The entire state is frequently affected by various degrees of drought during the summer months [e.g., *Irvine et al.*, 2002; *Law and Waring*, 1994; *Law et al.*, 2001; *Schwarz et al.*, 2004]. Please refer to *Göckede et al.* [2010] for more details and figures. We characterize our model domain by assigning 10 different ecoregions, six land cover types, and two disturbance regimes, all based on different remote sensing sources (see Table 1 for details and data sources). We will further on refer to combinations of ecoregion, land cover, and disturbance type as "surface type". The actual number of surface types present in the model domain is 68, compared to 120 possible combinations of ecoregion, land cover type, and disturbance. The horizontal resolution of the model grid is 1 km^2, with additional information on surface type distribution on the subgrid scale, thus effectively increasing the resolution beyond that threshold.

[9] Our atmospheric observation network consists of five sites arranged in a west–east transect spanning Oregon between the Pacific Coast and the Great Basin (see Table 2 for details). All sites are equipped with the same custom-built basic instrument setup that monitors atmospheric CO_2 mixing ratios calibrated against standard gases provided by the National Oceanic and Atmospheric Administration's Earth System Research Laboratory (NOAA ESRL). Instruments were installed between August 2006 and May 2007 and have been providing continuous hourly averaged atmospheric CO_2 mixing ratios to present. To minimize the potential influence of boundary layer mixing biases [e.g., *Peters et al.*, 2010], we restricted the observations used for the atmospheric inversion to afternoon averaged CO_2 mixing ratios (1400–1800 local time (LT)).

[10] Details on quality control and data filtering procedures have been provided by *Göckede et al.* [2010]. Here, we improved our analysis with the addition of spatially resolved wildfires on a daily time step based on fire perimeter information (http://geomac.usgs.gov/). The data product does not include emission source strengths, so the information cannot be used to directly quantify the influence of wildfires on the CO_2 measurement data. However, we used the wildfire maps to determine for each hourly measurement the proportion of the source weight function influenced by burning areas, and we flagged time steps with a wildfire influence above a fixed threshold.

[11] We used atmospheric CO_2 mixing ratio observations from the Yaquina Head (YH) and Mary's Peak (MPk) sites to analyze potential offsets in the advected background CO_2 signal (see section 2.3). Both sites are close to the Pacific coastline and are capable of sampling incoming maritime air masses undisturbed by local sources and sinks when careful data filtering is applied, while the other three sites in the network are significantly influenced by terrestrial carbon fluxes at all times and are thus unsuitable to characterize background mixing ratios. Measurements used for the analysis of biases in advected background mixing ratios cover the period from the onset of the observations (YH, April 2007; MPk, October 2006) until December 2008, which marks the end of the currently available background CO_2 data set (CarbonTracker). Both sites provide continuous measurements of hourly averaged atmospheric CO_2 mixing ratios, with additional sensors capturing local wind speed and direction as well as atmospheric pressure, air temperature, and humidity. (For details on site characteristics and data filtering, see appendix A).

2.2. Modeling Framework

[12] This study uses a "classic" atmospheric inversion setup to constrain terrestrial biosphere CO_2 fluxes based on atmospheric observations of CO_2 mixing ratios. Receptor locations are linked to sources and sinks at the surface through an atmospheric transport modeling module (section 2.2.2). Surface fluxes within the modeling domain are taken as the sum of biospheric fluxes (section 2.2.1) and anthropogenic emissions (section 2.4), with advected boundary CO_2 conditions provided by the CarbonTracker database (section 2.3). The optimum biospheric flux fields to match the atmospheric observations are determined using Bayesian optimization of the BioFlux base rate parameters (section 2.2.3).

2.2.1. Terrestrial Carbon Flux Model (BioFlux)

[13] BioFlux is a simple diagnostic model to simulate the exchange of CO_2 between terrestrial biosphere and atmosphere at fine spatial and temporal scales. It assimilates information from different remote sensing platforms, spatially interpolated surface meteorology, and reference flux data sources such as eddy-covariance measurements. Influence factors such as disturbance regime or land cover type are considered through customized parameter sets for each surface type. At the core of BioFlux are three equations for the fluxes of gross primary production (GPP), autotrophic respiration (RA), and heterotrophic respiration (RH).

[14] GPP is calculated as the product of light use efficiency and available photosynthetically active radiation, modulated by scaling factors such as daily minimum temperature, atmospheric water vapor pressure, and cloud cover. As an upgrade to the flux equation presented by

Göckede et al. [2010], we added the soil moisture fraction as an additional constraint:

$$\mathrm{GPP} = -\mathrm{GPP_{base}} \cdot \mathrm{APAR} \cdot T_{sc} \cdot \mathrm{VPD_{sc}}\left(1 + \mathrm{CL_{wgt}} \cdot \mathrm{CL_{sc}}\right)\mathrm{Age_{GPP}} \cdot \mathrm{SW_{sc}}, \quad (1)$$

where

GPP gross primary production [g C m^{-2} time $step^{-1}$];
GPP_{base} base rate for gross primary production [g C MJ^{-1}];
APAR available photosynthetically active radiation [MJ m^{-2} time $step^{-1}$];
T_{sc} minimum temperature scaling factor;
VPD_{sc} vapor pressure deficit scaling factor;
CL_{wgt} cloudiness influence weight;
CL_{sc} cloudiness scaling factor;
Age_{GPP} age scaling factor on GPP;
SW_{sc} soil water fraction scaling factor.

[15] The new scaling factor SW_{sc} has been added to strengthen the influence of water stress on the computed CO_2 fluxes, where the atmospheric water demand alone, as represented by the water vapor pressure deficit (VPD), proved to be too weak of a constraint to capture the complex drought dynamics in the Oregon domain. SW_{sc} follows a functional form similar to that used for the minimum temperature constraint, progressively shutting off photosynthesis when the soil is drying out:

$$\mathrm{SW_{sc}} = \frac{1}{1 + \exp\left(\frac{-(\mathrm{SWC} - 0.25)}{0.1}\right)}, \quad (2)$$

where SWC is the fractional soil water content.

[16] SWC is calculated as the actual water content of the rooting zone, normalized by the water holding capacity, ranging between 1 for a completely filled soil reservoir and 0 for dried-out soils (see *Göckede et al.* [2010] for details). To avoid overestimation of moisture control on GPP through the multiplication of the two scalars VPD_{sc} and SW_{sc}, only the minimum of VPD_{sc} and SW_{sc} will be used in equation (1), while the higher value is set to one. Consequently, SW_{sc} will only have an impact on the computation of GPP if the soil water stress is higher than the VPD influence.

[17] The equations for autotrophic and heterotrophic respiration remain the same as described by *Göckede et al.* [2010]:

$$\mathrm{RA} = R_m + R_g, \quad (3)$$

$$R_m = R_{m,\mathrm{base}} Q_{10}^{[(T_{\mathrm{air}} - 20)/10]} \mathrm{fPAR}, \quad (4)$$

$$R_g = (\mathrm{GPP} - R_m) R_{g,\,frac}, \quad (5)$$

$$\mathrm{RH} = \mathrm{RH_{base}} \cdot \mathrm{TS_{sc}} \cdot \mathrm{SW_{sc}} \cdot \mathrm{Age_{RH}} \cdot \mathrm{fPAR}, \quad (6)$$

where

RA autotrophic respiration [g C m^{-2} time $step^{-1}$];
R_m maintenance respiration [g C m^{-2} time $step^{-1}$];
R_g growth respiration [g C m^{-2} time $step^{-1}$];
$R_{m,base}$ base rate for maintenance respiration [g C m^{-2} time $step^{-1}$];
Q_{10} base rate for Q_{10} temperature influence function;
T_{air} actual air temperature [°C];
fPAR fraction of photosynthetically active radiation absorbed by the canopy;
$R_{g,frac}$ fraction of assimilated carbon used in growth respiration, set to 0.25;
RH heterotrophic respiration [g C m^{-2} time $step^{-1}$];
RH_{base} base rate for heterotrophic respiration [g C m^{-2} time $step^{-1}$]
TS_{sc} soil temperature scaling factor;
SW_{sc} soil water scaling factor;
Age_{RH} age scaling factor on RH.

[18] Overall, the model makes use of eight free parameters that can be customized per surface type through model training against reference flux data sources. Only three of these eight parameters, namely the flux base rates for GPP, RA, and RH, are optimized by the inverse modeling framework, while the remaining five are treated as constant prior estimates.

[19] BioFlux model initialization follows the same concept as outlined by *Göckede et al.* [2010, section 2.3.2]. However, in this study, we used the WRF (Weather Research and Forecast, www.wrf-model.org) model rather than SOGS (Surface Observations Gridded System) [e.g., *Jolly et al.*, 2005] as a data source for daily surface meteorology. The spatial resolution of the WRF data followed the grid specifications given in section 2.2.2. All data were reprojected onto a Lambert Azimuthal Equal Area grid with 1 km horizontal resolution to mimic the format provided by DayMet. As in the previous study, we used knowledge-based high-quality PRISM (Parameter-elevation Regressions on Independent Slopes Model) [e.g., *Daly et al.*, 2008] data in monthly time steps as a reference to scale WRF precipitation fields. The use of WRF data in this context harmonizes the flux computation with the transport modeling, since the same meteorological fields that govern the simulation of turbulent flow patterns are now driving changes in photosynthesis and respiration.

2.2.2. Atmospheric Transport Modeling

[20] Atmospheric transport of CO_2 was computed on the basis of WRF model output coupled to the STILT atmospheric transport model (Stochastic Time Inverted Lagrangian Transport) [*Lin et al.*, 2003]. We used offline coupling between these two transport components; i.e., WRF generates refined three-dimensional transport fields based on initial and boundary conditions taken from National Centers for Environmental Prediction (NCEP) global data sets (final (FNL), http://dss.ucar.edu/datasets/ds083.2), which were fed into STILT to compute the high-resolution source weight functions that indicate the actual "field of view" of the measurements. STILT source weight functions directly convert spatially distributed fields of CO_2 flux source strengths into changes of atmospheric CO_2 mixing ratios at the receptor locations. The setup specifications for the WRF and STILT models were similar to those of *Göckede et al.* [2010], except for the spatial setup of the nested WRF model grids. Here, the inner grid focused on western and central Oregon (6 km resolution, 88 × 70 cells, 40 s time step), and the outer grid covered the state of Oregon (18 km resolution, 48 × 40 cells, 120 s time step). In both grids, the number of vertical levels was 27.

2.2.3. Parameter Optimization

[21] The general setup of the Bayesian approach [e.g., *Enting*, 2005; *Tarantola*, 1987] to optimize flux model parameters followed the concept outlined by *Göckede et al.* [2010]. The three flux base rates for GPP, RA, and RH were optimized for each of the 120 surface types defined for the Oregon domain, 68 of which are active. Optimized parameters were considered to be invariant with time for most of the sensitivity tests presented below, while the influence of seasonal variability for these parameters was tested in additional simulations. As an extension to the reference by *Göckede et al.* [2010], we added Lagrange multipliers [*Gill et al.*, 1986] to the Bayesian optimization to constrain the base rates to physically plausible values. The minimum plausible base rates were assumed to be 0.25. The augmented cost function L [e.g., *Enting*, 2005; *Michalak et al.*, 2004] is defined through

$$L = \frac{1}{2}(\mathbf{z}-\mathbf{Hs})^T\mathbf{R}^{-1}(\mathbf{z}-\mathbf{Hs}) + \frac{1}{2}(\mathbf{s}-\mathbf{s_p})^T\mathbf{Q}^{-1}(\mathbf{s}-\mathbf{s_p}) + \boldsymbol{\lambda}^T(\mathbf{s}-0.25), \quad (7)$$

where

- $\mathbf{z}$ atmospheric observations (vector of dimension $n \times 1$);
- $\mathbf{s}$ base rates to be optimized (vector of dimension $m \times 1$);
- $\mathbf{H}$ Jacobian transfer function linking base rates to mixing ratios ($n \times m$ matrix);
- $\mathbf{R}$ model-data mismatch covariance ($n \times n$ matrix);
- $\mathbf{s_p}$ a priori base rates (vector of dimension $m \times 1$);
- $\mathbf{Q}$ covariance matrix of errors in $\mathbf{s_p}$ ($m \times m$ matrix);
- $\boldsymbol{\lambda}$ Lagrange multipliers used to enforce the inequality constraint $\mathbf{s} \geq 0.25$ (vector of dimension $m \times 1$);
- m number of parameters to be optimized (120 surface types for three fluxes, total of 360);
- n number of observations (4 h afternoon averages for three sites 2007–2008, total of 1497).

[22] The Lagrange multiplier term is added to the objective function to enforce an inequality constraint, which, in this case, simply specifies that each base rate (i.e., each element of $\mathbf{s}$) must be $\geq$0.25. The use of Lagrange multipliers is a standard approach in constrained optimization problems. Because the Lagrange multipliers are used here for *inequality* (as opposite to equality) constraints, the system must be solved iteratively [e.g., *Gill et al.*, 1986, section 5.5.1] to meet the Karush-Kuhn-Tucker conditions [*Karush*, 1939; *Kuhn and Tucker*, 1951]. The solution method involves setting the derivative of the objective function with respect to $\mathbf{s}$ and $\boldsymbol{\lambda}$ equal to 0. For inequality constraints, the Lagrange multipliers of the points corresponding to active constraints (i.e., those where the base rates are estimated to be 0.25) must be positive. This approach has previously been applied in enforcing parameter nonnegativity in interpolation [e.g., *Barnes and You*, 1992] and inverse modeling applications in other environmental fields [e.g., *Michalak and Kitanidis*, 2004]. Note that Lagrange multipliers have been classically applied in the context of deterministic (rather than probabilistic or stochastic) optimization problems, and the uncertainties estimated here must therefore be interpreted with caution. The posterior uncertainties are still representative of the relative uncertainty of the estimated parameters, but they are not strictly speaking representative of a Gaussian model (which, by definition, would not have constraints on parameter values). This approach presents a compromise between enforcing plausible values for base rates and at the same time using an overall setup that is similar to that used in previous inversion studies.

[23] For the definition of the prior uncertainty covariance matrix $\mathbf{Q}$, we calculated prior base rates and their uncertainty using the two-stage multivariate optimization approach described by *Göckede et al.* [2010]. These a priori base rates for GPP, RA, and RH and their uncertainties were assigned to surface types aggregated to 8 km resolution. A variogram analysis [e.g., *Cressie*, 1993] was performed on this fine-resolution base rate map, with the base rates normalized by their uncertainties, to provide an estimate of the spatial scales over which base rates are correlated (i.e., the base rate correlation length). A variogram describes the expected increase in base rate variance as a function of the separation distance between two points in the model domain, and it was employed to consider correlations between uncertainties defined for, e.g., different biomes in the same ecoregion, or the base rates of a specific biome type in different ecoregions. The fitted variogram model assumed an exponential decay in covariance with separation distance [e.g., *Michalak et al.*, 2004]. The optimized correlation length was used to define a square covariance matrix that relates each 8 km grid cell to each other 8 km grid cell using the fitted exponential model. The variances in this square matrix were multiplied again by the specified standard deviations of the base rates to offset the normalization prior to the variogram analysis. To represent the expected covariance between surface types (which are defined at a scale coarser than 8 km), this covariance matrix was aggregated by averaging all the elements that correspond to a given land surface type (diagonal elements of the final $\mathbf{Q}$) and all elements that correspond to a given *combination* of land surface types (off-diagonal elements of the final $\mathbf{Q}$). Differences in disturbance regime are neglected in this analysis; i.e., the same off-diagonal covariances are used for forests disturbed by wildfires and clear-cutting, respectively. The final product is an a priori covariance matrix $\mathbf{Q}$ that includes off-diagonal elements that represent links between the prior uncertainty estimates for different surface types. Consideration of correlations between base rates reduced the effective degrees of freedom in $\mathbf{Q}$ from 360 (no off-diagonal elements) to about 194 (with off-diagonals). Please refer to appendix B for details on the definition of the model-data-mismatch covariance matrix $\mathbf{R}$.

2.3. Advected Background CO_2

[24] Regional scale top-down models focusing on constraining components of the carbon cycle need to account for variability in the advected CO_2 signal (see, e.g., Figure 1) for a correct interpretation of the observations. Inaccurate representation of this variability will force the model to attribute advection-driven changes in atmospheric CO_2 to the sources and sinks of CO_2 within the model domain, leading to systematic biases in the optimized flux estimates [*Schuh et al.*, 2010]. Four-dimensional CO_2 fields from global scale models may provide suitable data sources to solve the boundary conditions of smaller scale models, but due to the scale discrepancies a systematic "small scale" bias in global models could trigger significant shifts in regional scale results because the offsets have to be balanced by fluxes from a much smaller area.

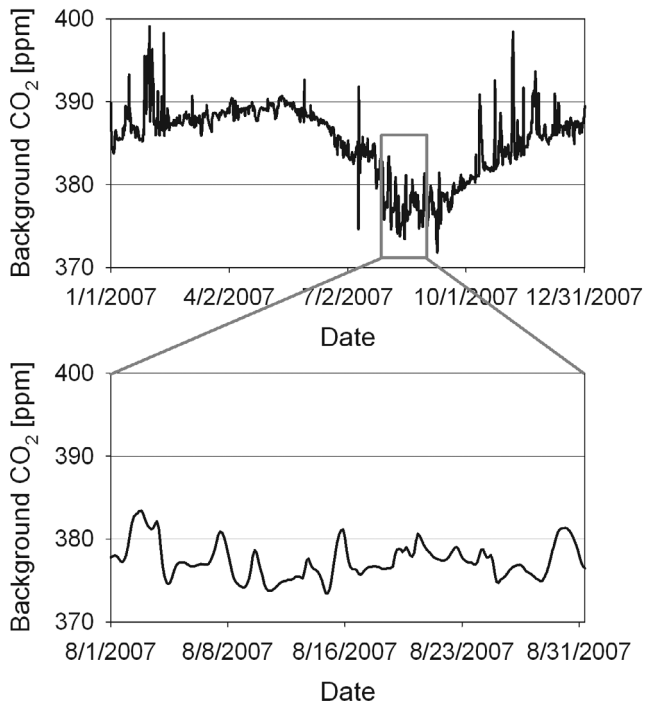

Figure 1. Temporal variability of the advected atmospheric CO_2 mixing ratios just off the Oregon coastline near the Yaquina Head site (44.5°N, 125.5°W, 100 m above sea level), taken from the CarbonTracker 2009 data set. The long-term variability shown in the upper panel represents the fluctuation in source/sink strengths over the course of the year in the Northern Hemisphere, while the shorter-term variability in the lower panel is mostly a consequence of changing transport patterns at the continental to regional scale.

[25] We extracted information from the North American grid of the 2009 CarbonTracker release (1° × 1° spatial resolution, 34 vertical levels, 3 h temporal resolution [*Peters et al.*, 2007; see also http://carbontracker.noaa.gov]) to assign CO_2 mixing ratios to the air masses at their initial entry into our model domain. CarbonTracker is a global scale data assimilation system that aims at providing improved insight into the carbon cycle feedbacks between biosphere and atmosphere in the context of climate change. Fluxes and mixing ratios simulated by CarbonTracker are trained and validated against atmospheric observations of CO_2 mixing ratios collected at 81 global monitoring sites, 17 of which are located within the contiguous United States and three in U.S. West Coast states. To obtain flux and mixing ratios fields that agree with the observations, prior fluxes taken from CASA-GFED version 2 (Carnegie-Ames-Stanford approach Global Fire Emission Database) [*Randerson et al.*, 2005; *van der Werf et al.*, 2006] in 1° × 1° resolution are multiplied by scaling factors optimized for larger ecoregions on a weekly time step using an ensemble Kalman filter technique [*Peters et al.*, 2005].

2.4. Fossil Fuel Emissions

[26] Boundary conditions for anthropogenic fossil fuel emissions are taken from the VULCAN database [*Gurney et al.*, 2009a]. This inventory of U.S. fossil fuel emissions is based on seven primary data sets summarizing, e.g., residential and commercial activities, power production, or road and aircraft emissions, available as a mixture of geocoded point sources, line sources, and area sources. Additional data are included to refine the space and time distributions, such as breaking up county level residential emissions using U.S. Census data, or downscaling onroad emissions based on GIS (geographic information system) road information [*Gurney et al.*, 2009b]. The VULCAN data sets compare favorably with annual, nationwide emission estimates published by other U.S. agencies in most sectors, and the high-resolution emission fields were shown to have a major effect on atmospheric CO_2 mixing ratio fields simulated by a transport model when compared to an earlier, coarser emission inventory [*Gurney et al.*, 2009b]. VULCAN is currently lacking a formal uncertainty assessment, so the accuracy of the fine-scale space/time distribution of emission fluxes cannot be quantified yet. However, in-depth sensitivity studies on key parameters are expected to be available soon (K. Gurney, personal communication).

[27] VULCAN provides hourly fossil fuel CO_2 emissions from multiple emission sources, aggregated into a regular 10 × 10 km grid covering the contiguous United States. For the present study we employed data from VULCAN version 1.3, which is based on inventories of the year 2002. To facilitate extrapolation of the 2002 results to the simulation years 2007 and 2008 covered in this study, we aggregated the continuous emission time series to hourly averages for weekdays, Saturdays, and Sundays for each month, which smoothes out the effect of shorter term climate anomalies that might have impacted the 2002 data set. For this extrapolation we assumed no significant trend in total statewide emission rates between the source year 2002 and the target years 2007 and 2008 [see also *Göckede et al.*, 2010].

3. Results and Discussion

[28] A sensitivity analysis of the impact of the changes in the modeling framework compared to that of *Göckede et al.* [2010] on performance is presented in appendix C. Results indicate that the inclusion of the Lagrange multipliers significantly improved model performance, while the other modifications did not cause statistically significant changes in the results when evaluated against associated uncertainties. Still, the revised modeling framework stabilizes the results by assimilating additional prior information.

3.1. Sensitivity of the Statewide CO_2 Budget on Advected Background CO_2 and Fossil Fuel Emissions

[29] To investigate the overall effect of changes in advected background CO_2 mixing ratios on the net statewide CO_2 budget, we analyzed how the correction of hypothetical systematic biases in CarbonTracker would affect the simulated CO_2 fluxes. CO_2 budget results were obtained for the state of Oregon and the year 2007 without considering seasonality in optimized parameters. Background mixing ratios taken from CarbonTracker were modified using shifts between −2 and 1 ppm, where negative values indicate a hypothetical high bias in CarbonTracker, so all advected background values were reduced for the flux optimization. Offsets were assumed constant throughout the year. We found a quasi-linear relationship between offsets in the advected background mixing ratios and shifts in the optimized CO_2 budget, with an average

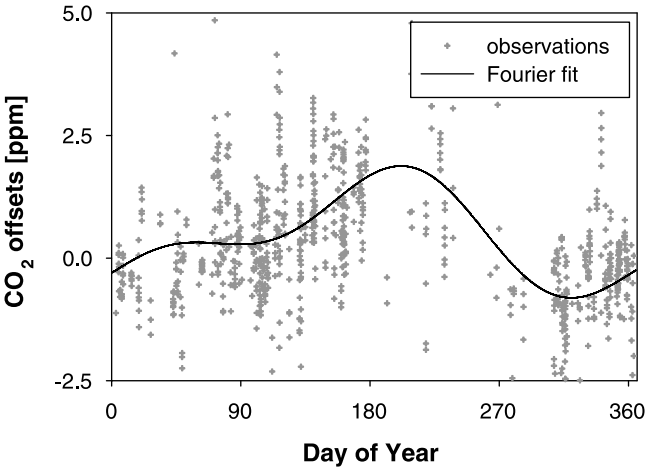

Figure 2. Offsets between advected CarbonTracker data and observed atmospheric CO_2 mixing ratios plotted against Julian day, with positive values indicating a high bias in CarbonTracker. Observations are taken from both the Yaquina Head (YH) and the Mary's Peak (MPk) sites, while the solid line indicates a truncated Fourier function fitted to the data, which is used to correct for background biases in the inversion.

slope of 34.9 [Tg C ppm^{-1}] for the state of Oregon in 2007. Because the slope reaches about the same order of magnitude as the annual statewide CO_2 sink (see also section 3.3), it follows that the computation of statewide CO_2 budgets is highly sensitive to potential biases in advected background mixing ratios.

[30] We conducted similar tests for the fossil fuel emissions taken from the VULCAN database. In a first test stage, original fossil fuel fluxes were multiplied by scaling factors ranging between 0.25 and 3, linearly altering the emissions in time and space. The simulated statewide biosphere CO_2 exchange proved to be less sensitive to these modifications than what was seen for the advected background, with a doubling of the VULCAN emissions (10.2 Tg C yr^{-1}) requiring an annual increase of about 7.5 Tg in vegetation uptake to fit the simulated to the observed atmospheric CO_2 mixing ratios. Since the VULCAN state-level annual emission estimates agree well with other emission data sources, a significant net offset of more than 20% in the provided statewide emissions is highly unlikely, and thus the impact on the simulated CO_2 budget is assumed to be small. To test the more realistic uncertainty of temporal and/or spatial shifts in the fossil fuel flux patterns while maintaining the same statewide annual emissions budget, in a second stage we created seven VULCAN scenarios as input for the inverse modeling framework. Scenario setup included smoothing of emissions to monthly or annually constant values, shifting emission fractions from one season to the remaining months, or spatially smoothing emissions to a grid of 50 × 50 km. None of these emission scenarios caused the optimized annual biosphere CO_2 exchange to shift by more than 2%. We therefore conclude that the inverse modeling results for the state of Oregon are robust toward potential systematic biases in the VULCAN data set, so fossil fuel fluxes are not regarded as a major source of uncertainty for statewide results.

3.2. Advected Background CO_2 Correction

[31] We assume potential offsets in advected CO_2 mixing ratios from CarbonTracker have no interannual variability, so results are analyzed and plotted against Julian days. For both sites that represent clean incoming air from the Pacific (YH and MPk), we found a pronounced seasonal trend with positive offsets in summer and negative offsets in winter. Offset characteristics are similar at both sites even though the site characteristics and the setup of the data filters differ significantly between the two sites. Subtle temporal shifts in seasonality are mostly attributed to the distribution of data gaps over the course of the year, which is influenced by availability of atmospheric observations and occurrence of periods dominated by westerly winds. The good agreement between offsets observed at both sites rules out a systematic effect of measurement conditions on the overall findings and confirms the setup of the data filtering methods applied to exclude systematic biases.

[32] Figure 2 displays the seasonal trends of offsets between advected CarbonTracker data and atmospheric observations from both YH and MPk sites after application of the data filtering described in appendix A. We fit a truncated Fourier function to the offsets to produce a continuous representation of the seasonality of advected background CO_2 mixing ratio biases from CarbonTracker in the Oregon model domain. The continuous function is subsequently used to correct CarbonTracker data sets in the inverse modeling framework. For the curve fitting, a full year of data was copied to the beginning and end of the time series, respectively, to minimize boundary effects. The resulting function has a seasonal amplitude of about 2.7 ppm and produces an annually averaged offset of 0.45 ppm. Application of the continuous Fourier function reduces the root-mean-square error (RMSE) from 0.531 ppm (offsets only) to 0.026 ppm (offsets minus Fourier function). Application of these background offsets to correct CO_2 boundary conditions provided by CarbonTracker significantly shifted the simulated net regional scale carbon fluxes, and produced better comparisons of simulated CO_2 mixing ratios with the reference CO_2 observations (see detailed results in section 3.3).

[33] The findings presented in the last two sections emphasize the high importance of accurate CO_2 mixing ratio boundary conditions for inverse modeling frameworks operated on regional to continental scales. Very small potential offsets in background CO_2 mixing ratios can trigger large relative shifts in simulated biospheric flux fields in regions sensitive to these offsets, and results presented here demonstrate that current high-resolution global scale products such as CarbonTracker can be subject to seasonally varying offsets that significantly affect regional scale CO_2 budgets. This suggests that inverse modeling frameworks depending on external data to estimate their CO_2 boundary conditions should use part of the observations to check for potential offsets in these background mixing ratios to avoid biases in the simulated fluxes.

[34] The work presented here cannot be used to directly determine the extent to which the shifts in fluxes due to background offset correction found for the Oregon domain would also be present for other model domains. Detailed

Table 3. Results for Root-Mean-Square Error Between Observed and Simulated Atmospheric CO_2 Mixing Ratios, and the Associated Coefficients of Determination[a]

		Root-Mean-Square Error			Coefficient of Determination (R^2)		
Site	Year	Prior	Posterior	% Reduction	Prior	Posterior	Difference
MF	2007	3.633	2.707	25.5	0.735	0.788	0.053
	2008	4.498	3.517	21.8	0.709	0.742	0.033
MP	2007	2.678	2.398	10.5	0.690	0.709	0.019
	2008	2.907	2.640	9.2	0.675	0.721	0.046
BU	2007	1.678	1.662	1.0	0.776	0.777	0.001
	2008	1.626	1.645	−1.2	0.811	0.808	−0.003

[a]All results are based on afternoon averaged (1400–1800 LT) data in daily time step.

information on CarbonTracker biases has been noted earlier on their Web page (http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/profile_stats.html), where a more in-depth discussion on spatial and temporal variability of the offsets between model output and observations can be found. The seasonal amplitudes of the offset between our observations and CarbonTracker data shown in Figure 2 correspond well with results for other sites presented on the Web page. Assuming these offsets are systematic and apply to larger scales, inverse modeling CO_2 flux simulations for, e.g., the North American continent could be biased at the margins of the model domain, where the influence of the boundary conditions is strongest, and offset correction could be an important step toward reconciling inversion-derived flux estimates with the findings by bottom-up biogeochemical models. A systematic high bias in CO_2 mixing ratios of air masses entering North America might also hint at an underestimation of the sink strength of the North Pacific Ocean basin and/or terrestrial vegetation upstream, which would be balanced in a North American inversion by increasing the estimate of the North American terrestrial sink. However, the offsets observed in this study could potentially also represent a local anomaly, which might bias regional flux pattern in continental scale modeling setups, but may not alter the total annual flux estimates significantly on continental scales. Given the high sensitivity of inverse modeling toward advected background CO_2 mixing ratios, and the potential impacts toward improving flux estimates on regional to continental scales, this subject should be further investigated using different model domains and additional observation sites.

3.3. Simulated CO_2 Mixing Ratios and Statewide Budget

[35] The seasonal amplitude between summertime low and wintertime high mixing ratios observed at the three sites decreases from ~20 ppm at the westernmost (Mary's River mature fir, MF) site over ~15 ppm at the center of the state (Mary's Peak (MP) site) to only about 10 ppm in the east (Burns, BU). This gradient reflects the change in vegetation activity as a response to the transition from humid climate in the West to semiarid conditions in the East. CO_2 mixing ratio signal variability on time scales between 1 and 10 days also decreased along that gradient, thus reducing the assigned model-data-mismatch uncertainties (see also appendix B) for the MP and BU sites, since residuals between observed and simulated CO_2 mixing ratios are significantly lower than those at the MF site (Table 3).

3.3.1. CO_2 Mixing Ratio Results

[36] Net changes between prior and posterior CO_2 mixing ratio time series vary considerably among the three sites (Table 3). Comparing residuals and correlation between observed and simulated data, significant improvement in RMSE could be obtained for the MF and MP sites, with reductions ranging between 9% and 25% for individual years. For both sites, the optimization also considerably improved the coefficient of determination (R^2). For BU, where prior RMSE was already lower than the optimized results obtained at the other two sites, changes between prior and posterior results were less pronounced, with RMSE reductions between −1% and 1% and R^2 differences close to zero. Compared to prior results, mean posterior simulated CO_2 mixing ratios were increased at the MF site (+0.7 ppm), slightly negative at the MP site (−0.16 ppm), and about neutral at the BU site (−0.03 ppm). Differences between prior and posterior simulated CO_2 mixing ratios were most pronounced in spring and summer (Figure 3), with shifts in spring always positive (i.e., posterior higher than prior) and shifts in summer positive for MF, neutral for BU, and with different signs in both observation years for MP. Absolute offsets between simulated and observed CO_2 mixing ratios were reduced for all sites, with the highest improvement observed for summer (−0.8 ppm) followed by spring and fall (−0.4 ppm), and the smallest shifts found in winter (−0.2 ppm).

[37] The RMSE and R^2 values appear to be dominated by temporal and spatial variability in the local vegetation signal. Observed patterns can be attributed to three major influence factors: (1) structural shortcomings in the BioFlux model, with associated uncertainties scaling with the magnitude of the flux; (2) spatial and temporal bias in the climate drivers, which will have higher impact on active vegetation; and (3) transport model uncertainties, where the net effect on the inversion depends on the spatial variability of the surface flux signal. At the MF site, which is dominated by the very active forest of Western Oregon, all three factors contribute to elevated uncertainties, while at the BU site in Eastern Oregon, the first two of these factors are significantly reduced with overall vegetation activity being low, and also the potential bias due to transport errors is low because of the length scale of flux variability is small. The MP site is influenced by both major climate zones, so it is important for data interpretation to capture the transition from humid western forests to semiarid eastern shrublands through highly resolved surface flux maps.

3.3.2. Statewide CO_2 Flux Budget

[38] To test for potential model biases (e.g., BioFlux algorithms simplifying seasonal trends in photosynthetic rates with seasonal phenology), simulated statewide annual budgets for terrestrial biospheric CO_2 exchange were calculated for two different parameter optimization strategies. The first (constant parameters) follows the setup outlined above, with flux base rates held constant throughout the modeling time frame. We compared these results to simulations using seasonally varying parameters for winter (December–February), spring (March–May), summer (June–August), and fall (September–November), where no temporal covariance was assumed between seasons a priori. Both optimization strategies included the correction of advected background CO_2 mixing ratios described in section 3.2 and combined observations from 2007 and 2008. We found no significant difference between both optimiza-

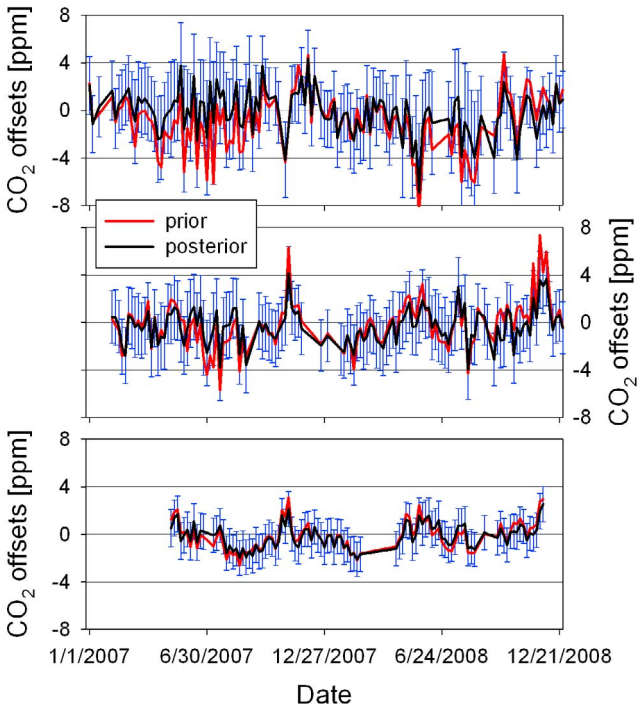

Figure 3. Afternoon averages (1400–1800 LT) of offsets (modeled minus measured) in CO_2 mixing ratios for the (top) mixed forest (MF), (middle) Metrolius mature pine (MP), and (bottom) Burns (BU) sites. Results are given before (prior mixing ratios, red) and after (posterior mixing ratios, black) optimization. For all time steps that contain measurements that passed the quality filter, vertical blue bars indicate the total model-data-mismatch uncertainty for posterior results (one standard deviation). Results were binned into 5 day averages to facilitate visualization of trends and differences between the prior and posterior results.

tion strategies for net uptakes averaged for 2007 and 2008, and only small deviations of up to 2% between results based on seasonal and constant parameter sets for individual years (Table 4). The posterior annually averaged net CO_2 uptake for both the seasonally varying parameter optimization (35.4 ± 10.6 Tg C) and the constant parameter setting (35.8 ± 5.2 Tg C) was reduced compared to the prior flux estimates (40.0 ± 23.8 Tg C). Comparison of observed versus simulated CO_2 mixing ratios indicated that the use of seasonally varying parameters did not change the RMSE at the MF site significantly (Table 3: average increase in RMSE reduction of 0.15%). For the other two sites, however, a considerable improvement in RMSE was observed upon switching from constant to seasonal parameter sets. There was an increase in RMSE reduction, averaging 9.8% (constant) to 18.4% (seasonal) for MP and from −0.1% to 7.0% for BU. Compared to constant parameter sets, the use of seasonally varying parameters increased the R^2 difference between prior and posterior results by 0.03 for the MP site and 0.01 for MF and BU.

Table 4. Simulated Terrestrial Biospheric CO_2 Budget for the State of Oregon[a]

Year	Prior	Posterior, Seasonal	Posterior, Constant
2007	45.1 ± 24.4	40.7 ± 11.4	39.2 ± 6.4
2008	34.8 ± 23.1	30.1 ± 10.9	32.3 ± 6.1
2007/2008 average	40.0 ± 23.8	35.4 ± 11.1	35.8 ± 6.2

[a]All values given in Tg CO_2-C yr^{-1}. Positive values indicate that the vegetation acts as a net sink for CO_2. Indicated uncertainty ranges represent a single standard deviation.

[39] The spatial distribution of annually averaged net ecosystem exchange fluxes optimized using constant parameter sets (Figure 4) indicates no significant shift in the overall distribution of sources and sinks before and after optimization. Both the prior and posterior flux fields show the highest uptake of CO_2 by the vegetation in the humid western half of the state and significantly reduced vegetation activity in the semiarid eastern region. The posteriori fluxes are relatively smooth in their spatial distribution and display only minor gradients at the transitions between ecoregions. Shifts in absolute source/sink strengths between prior and posterior fluxes are significant, though, and vary considerably between ecoregions. For most of the state, vegetation uptake was reduced after optimization, except for the Coast Range ecoregion and some areas in eastern Oregon dominated by agriculture and grassland. The highest net reductions in overall uptake were found for the forested areas in the northeast of the state (Blue Mountains (BM) ecoregion) and around the Cas-

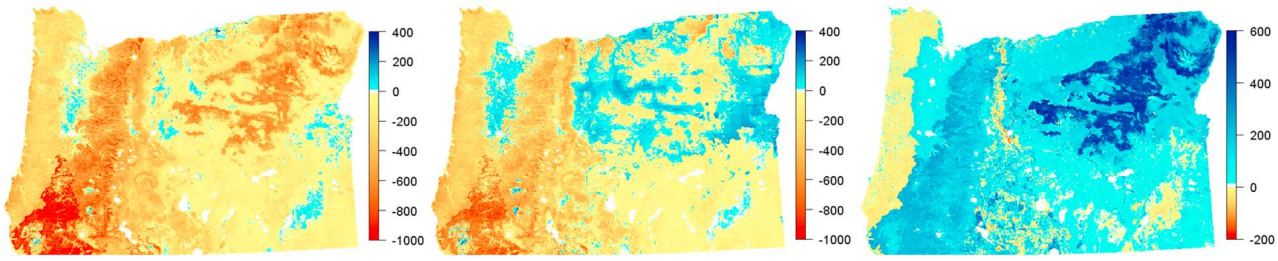

Figure 4. Results for (left) prior net ecosystem exchange (NEE), (middle) posterior NEE based on temporally invariant parameters, and (right) difference between both (posterior–prior). Please note the different scale in the right panel. All fluxes represent annual budgets given in g CO_2-C m^{-2}, averaged for the two simulation years 2007 and 2008. Negative values (warm colors) in annual flux budgets (Figure 4, left, and Figure 4, middle) indicate that the biosphere is acting as a net sink of CO_2, while negative values in the difference map (Figure 4, right) indicate that a higher sink strength was assigned to a specific area through the optimization process. Maximum values were truncated to optimize the display of spatial flux patterns.

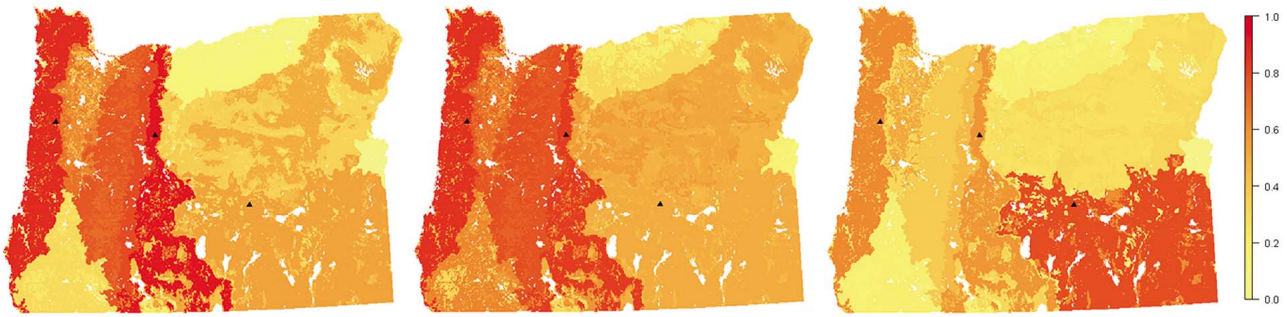

Figure 5. Uncertainty reduction for the fluxes of (left) gross primary production (GPP), (middle) autotrophic respiration (RA), and (right) heterotrophic respiration (RH). All results are calculated as 1 minus the ratio of posterior to prior uncertainty, so high values indicate a high information gain. Black triangles indicate site positions.

cade mountains (WC, West Cascades; CC, Cascade Crest; EC, East Cascades). The high relative changes in the eastern Oregon forests reflect the lack of prior information for these ecosystems, where no direct flux measurements were provided to train the initial flux model parameters.

3.3.3. Influence of Background Correction and Evaluation of Uncertainties

[40] To test the impact of the CO_2 background correction on simulated fluxes and CO_2 mixing ratios, we ran the same simulation setup as outlined above without correcting for CarbonTracker biases through the Fourier function (Figure 2). As already outlined in sections 3.1 and 3.2, correcting for background offsets significantly impacts the simulated annual CO_2 budget for the state of Oregon: Without the correction, terrestrial biosphere CO_2 uptake is estimated at 67.7 Tg CO_2-C yr^{-1}, so the sink strength of 35.8 Tg CO_2-C yr^{-1} presented in Table 4 represents a 47% reduction compared to the uncorrected reference. Correcting for background offsets also generally improved the correlation between observed and simulated posterior CO_2 mixing ratios. Except for two site years (MF 2007 and BU 2007), posterior RMSE between observations and simulations was lower when background correction was applied (difference in posterior RMSE, corrected minus uncorrected, for MF, −0.17 ppm; MP, −0.001 ppm; and BU, 0.02 ppm). In all cases, a higher posterior R^2 was obtained when we applied the background correction (difference in posterior R^2, corrected minus uncorrected, for MF, 0.03; MP, 0.02; and BU, 0.03).

[41] Uncertainty estimates for the statewide flux budgets (Table 4) are based on prior and posterior parameter covariance matrices, multiplied by scaling factors that reflect the flux computation in BioFlux based on the surface meteorology. Compared to prior results, a posteriori uncertainties were reduced significantly for both optimization strategies, with larger reductions obtained for the constant parameter setting since all information available from the atmospheric observations is used to improve a single set of parameters, while in the seasonal case this information is spread among four full sets of parameters. These uncertainty estimates represent a comprehensive assessment of the influence of potential biases in, e.g., surface meteorology, BioFlux algorithms, or fossil fuel emissions, as these sources of uncertainty were incorporated into the covariance matrix of prior errors (**Q**) and the model-data mismatch matrix **R**. Uncertainty reductions relative to the prior represent our confidence in the amount of information that can be extracted from observations of atmospheric CO_2 mixing ratios, considering all associated uncertainties such as transport or mixing biases (see appendix B). In addition, uncertainty numbers in Table 4 include biases associated with offsets in advected CO_2 mixing ratios and the implemented correction procedure. To acknowledge potential biases introduced through the uneven seasonal distribution of reference data and the choice of the interpolation polynomial (Figure 2), as a conservative estimate we attributed an uncertainty to this process that equals the bias introduced by a linear shift of 0.1 ppm in all corrected background mixing ratios (3.5 Tg CO_2-C; see section 3.1). This background uncertainty was added to the uncertainties derived from the posterior parameter covariance matrices to yield the total uncertainties displayed in Table 4.

[42] To evaluate the spatial distribution of information content extracted from the atmospheric observation network, uncertainty reduction was computed for each surface type as 1 minus the ratio of posterior to prior uncertainty (Figure 5). Spatial patterns in these graphs are aligned with the shapes of the ecoregions, since uncertainty reduction was not computed for each matrix grid cell individually, but averaged for each surface type (see *Göckede et al.* [2010] for details). Results vary between ecoregions, depending on their relative position within the field of view of each observation site, and also between the three flux components optimized (GPP, RA, and RH). Highest reductions can generally be obtained for those ecoregions containing the observation sites (CR, Coastal Range; EC, East Cascades; NB, Northern Basin), while for the more remote regions outside the tower footprints (SR, Snake River; CP, Columbia Plateau; KM, Klamath Mountains) flux parameter uncertainties remain high.

[43] The comparison between optimization runs using seasonally varying versus constant parameters indicates that simulated surface fluxes are barely affected by that difference in the setup, while considerable changes were observed for the correlation between observed and simulated CO_2 mixing ratios. Optimized RMSE results are mostly improved at the two sites located in semiarid regions of Oregon, which might indicate that the BioFlux model is oversimplifying carbon cycle processes that are specific to that climate zone (e.g., the handling of extreme drought stress). However, without additional ancillary information, e.g., eddy-covariance

fluxes and/or water balance measurements in the Great Basin, the causes of the observed shifts in simulated CO_2 mixing ratios cannot be isolated. Because the use of seasonally varying parameters significantly increases the number of degrees of freedom in the optimization process, this procedure always carries the risk of overfitting the model to the observations. We therefore prefer the use of temporally constant parameter sets for this modeling framework. The comparison of simulations with constant and seasonally varying parameter sets can serve as a useful diagnostic of potential shortcomings in the model framework, which subsequently can be analyzed in more detail using additional data sets and/or sensitivity tests.

[44] Additional biases might be introduced into the results by transport and mixing uncertainties, which are particularly important in high-resolution inverse modeling setups and in the presence of considerable spatial gradients in flux source strengths in the model domain. Earlier comparisons of simulated boundary-layer heights with radiosonde observations [see *Göckede et al.*, 2010] indicated reasonable agreement, but were compromised by the low vertical resolution in the reference data. More recent comparisons of WRF output with profile measurements from a 2003 aircraft campaign (data not shown) indicates good agreement between simulated and observed temperature profiles. We anticipate the availability of high-resolution SODAR profile data in our model domain that can provide vertically resolved boundary-layer wind fields to validate our transport and mixing simulations. Until these tests have confirmed the accuracy of the simulated meteorological fields, we will reduce the potential impact of transport and mixing biases by restricting the simulation starting times to afternoon situations with a well-developed boundary layer.

4. Conclusion

[45] We presented a comprehensive suite of enhancements for a previously published atmospheric inverse modeling framework [*Göckede et al.*, 2010] to constrain biosphere CO_2 fluxes at subregional scales. A special focus within the context of this study was placed on the role of potential biases in advected background CO_2 mixing ratios and anthropogenic fossil fuel emission fluxes. In addition, we tested the impact of allowing seasonally varying parameters in the optimization on simulated statewide CO_2 budgets.

[46] Systematic biases in advected background CO_2 mixing ratios were shown to have a significant impact on the simulated statewide CO_2 flux budgets, with annually averaged offsets of only 0.1 ppm shifting the flux budgets by 3.5 Tg C yr^{-1} (~10% of the net annual uptake) for the state of Oregon. Using observation sites close to the Pacific Coast that are capable of monitoring "clean" incoming maritime air masses, we found a systematic, seasonally varying offset in background CO_2 mixing ratios provided by the CarbonTracker database with an annual amplitude of 2.7 ppm. Correcting for these offsets reduced the simulated annual statewide CO_2 budget by 31.9 Tg C yr^{-1}, or 47% of the total CO_2 uptake. The inverse modeling framework proved to be robust against potential biases in the fossil fuel fluxes provided by the VULCAN data set, with a 0.7 Tg C yr^{-1} change in biospheric CO_2 uptake resulting from a potential 10% offset in anthropogenic emissions. We conclude that a precise definition of advected CO_2 background mixing ratios is paramount for inverse modeling frameworks operating on regional to continental scales, where small but systematic mixing ratio offsets can cause significant shifts in simulated flux fields. Correcting for biases has the potential to reduce gaps in simulation results between inverse modeling and bottom-up biogeochemical models, which may contribute an important step toward reconciling these two modeling strategies. The low sensitivity to relative changes in fossil fuel emissions are likely to be associated with the low overall emission level for the state of Oregon.

Appendix A: Data Filtering for Evaluation of Advected Background CO_2

[47] The Yaquina Head observation system is situated on a cape near Newport, Oregon. The inlet elevation of 127 m above sea level facilitates an unobstructed incoming flow field for all westerly wind directions. However, the location is subject to complex local to regional scale wind patterns that include katabatic flows and diurnal cycles of land-sea and sea-land breezes, respectively, which hold the potential to recirculate local air masses influenced by terrestrial vegetation and anthropogenic emissions [*Riley et al.*, 2005]. To identify these recirculated air masses, we used data filters to flag the above-mentioned transport situations. To ensure synoptic flow conditions with advected maritime air masses directly from the west, we established a diagnostic box with a latitude extension of 1° centered at the site location (see also Table 2), and an east-west extension between the site longitude position and a western boundary at −124.5°. All particle trajectories as simulated by the coupled WRF-STILT model had to enter this box through its western boundary. Second, we required hourly averaged wind directions measured at the YH site to fall within the westerly sector (225°–315°) within the 6 h preceding the measurement and the average wind speed within the same time frame had to be higher than 2 m s^{-1}. This data filtering procedure left 479 hourly measurements assumed to represent "clean" incoming maritime air masses, thus suitable for direct comparison with simulated background mixing ratios.

[48] The observation site on Mary's Peak sits on top of the highest mountain in the Oregon Coast Range, about 42 km from Yaquina Head, with a tower base elevation of 1248 m above sea level (asl). The mountaintop is dominated by open grassland, while the surrounding Coast Range mainly consists of heavily managed conifer forest. The elevation and proximity to the coast allow the sampling of incoming maritime air masses undisturbed by terrestrial sources and sinks under certain conditions. Most of the Coast Range hills in the westerly wind sector peak below 500 m asl, so that under stable atmospheric stratification at night the instruments on top of Mary's Peak will be mostly above the nighttime boundary layer and, thus, decoupled from the local surface fluxes. To identify suitable time frames, we first required the integrated footprint function as computed by WRF-STILT to indicate virtually no surface influence on the actual measurement, with a threshold set to <0.01 for the ratio of mixing ratio change [Δppm] and surface flux unit [micromole of CO_2 m^{-2} s^{-1}]. Second, we set up a similar control box as described above for the YH site, requiring all accepted particle trajectories to be advected through the narrow window at its western boundary. Third, we only

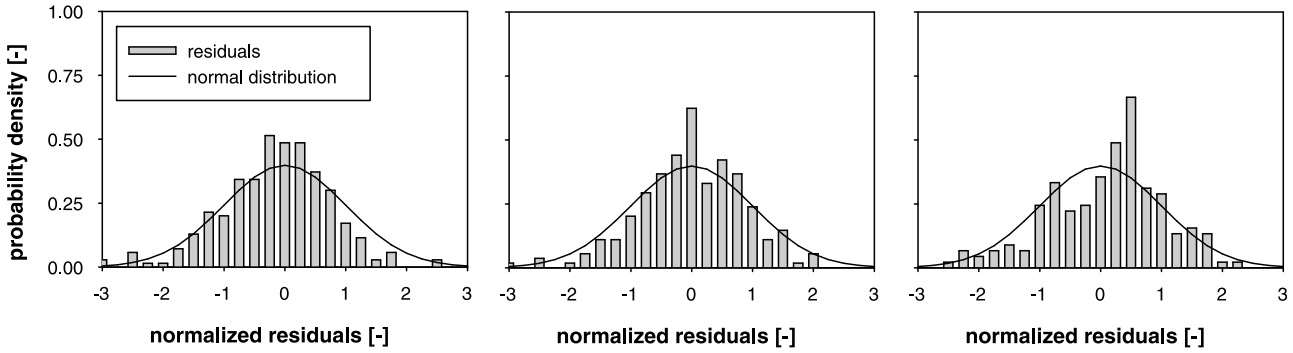

Figure B1. Frequency distribution of residuals between observed and simulated CO_2 mixing ratios for 2007 data, normalized with the model-data mismatch uncertainties (left, MF; middle, MP; right, BU). All residuals shown are based on afternoon (1400–1800 LT) averaged mixing ratios of atmospheric CO_2 (ppm). Average values for the normalized residuals are 0.90, 0.94, and 0.99 for the MF, MP, and BU sites. Residuals were binned into classes of 0.25. The black line shows a standard normal distribution (i.e., mean of 0 and variance of 1).

considered nighttime measurements between 2000 and 0400 LT for this study to exclude convective boundary layer mixing. This set of filters provided 419 hourly averaged CO_2 observations for this comparison study.

Appendix B: Definition of Prior Uncertainties and Model-Data Mismatch

[49] The definition of the diagonal elements of the model-data mismatch matrix **R** was refined compared to the procedure given by *Göckede et al.* [2010] to maximize the information gain from the atmospheric observations. The overall concept still follows that of *Gerbig et al.* [2003], calculating each diagonal matrix element R_i as the sum of six individual components defined below (equation (B1)), and neglecting potential temporal correlations that could be defined in the off-diagonal elements. To evaluate the plausibility of the model-data-mismatch definitions, we analyzed the distributions of the residuals between observed and simulated atmospheric CO_2 mixing ratios after optimization, normalized with the model-data-mismatch uncertainties. We adapted some of the model-data-mismatch components (see details below) to obtain a well-balanced setup of **R** that produces normalized residuals with an averaged value slightly lower than 1 and a distribution closely following a standard normal distribution (i.e., 0 mean and variance of 1).

$$R_i = S_{\mathrm{veg}} + S_{\mathrm{part}} + S_{\mathrm{eddy}} + S_{\mathrm{transp}} + S_{\mathrm{aggr}} + S_{\mathrm{ocean}}. \qquad \text{(B1)}$$

[50] The vegetation signal is derived as the observed CO_2 mixing ratios minus simulated background mixing ratios and fossil fuel emissions. Its uncertainty, S_{veg}, combines three components: (1) the scatter in the atmospheric observations, derived through regular sampling of a sample gas with known mixing ratio, and found to average 0.15 ppm over all three sites; (2) the uncertainty in fossil fuel fluxes taken from the VULCAN database, conservatively estimated to be 30% of the actual emissions; and (3) the uncertainty in the advected background CO_2 mixing ratios, which was obtained by comparing the simulated advected signal to observations at two monitoring sites sampling "clean" incoming air masses and set to a constant value of 1.235 ppm.

[51] The trajectory uncertainty, S_{part}, represents the stochastic uncertainty introduced by sampling a limited number of trajectories to represent the air masses influencing a receptor location, and it is defined as a percentage of the absolute simulated CO_2 mixing ratios. Values were reevaluated based on new model runs compared to those of *Göckede et al.* [2010], and they were found to be [0.19; 0.12; 0.07]% for the MF, MP, and BU sites, respectively. We will use similar square bracket notation further below to indicate uncertainty settings with individual values for those three sites. Due to the lack of vertical profile observations in the model domain, it was not possible to accurately define the uncertainty introduced by unresolved eddies, S_{eddy}. This parameter was therefore chosen to be one of two uncertainty components to be optimized for obtaining a well-balanced distribution of normalized residuals and was set to [1.0; 1.0; 0.5] ppm.

[52] The second uncertainty parameter that was adapted to optimize the distribution of normalized residuals was the transport field uncertainty, which is the first of two components of the overall transport uncertainty, S_{trans}. As for S_{eddy}, also for this component the reference observations to customize the settings for the Oregon domain are scarce, and did not allow for a reliable and representative data-based definition. The transport field uncertainty was set to final values of [1.6; 1.4; 0.4] ppm. The second component of S_{trans}, which represents the influence of boundary layer height uncertainty, was taken over from *Göckede et al.* [2010] as 0.3 times the modeled biospheric CO_2 flux signal. The remaining two model-data-mismatch components, S_{aggr} and S_{ocean}, were set to constant values of 0.1 ppm as suggested by *Göckede et al.* [2010].

[53] The distributions of the residuals between observed and simulated CO_2 mixing ratios for 2007, normalized with the model-data-mismatch uncertainties described above, are shown in Figure B1. For all three sites, these distributions closely followed the normal distributions and are centered at zero, indicating no systematic bias in the simulated mixing ratios as produced by the inverse modeling framework. The average values of the normalized residuals were found to be

Table C1. Results for Reduction in Root-Mean-Square Error Between Observed and Simulated Atmospheric CO_2 Mixing Ratios, and the Associated Changes in Coefficients of Determination[a]

	Reduction in Root-Mean-Square Error (%)						Change in Coefficient of Determination (R^2)					
Site	Ref[b]	LM	Q_{diag}	H_2O	WRF 2006	SOGS 2006	Ref	LM	Q_{diag}	H_2O	WRF 2006	SOGS 2006
MF	**23.6**	19.2	23.1	23.0	**21.9**	21.0	**0.043**	0.033	0.041	0.043	**0.017**	0.018
MP	**9.8**	6.1	9.0	8.9	**2.2**	2.5	**0.033**	0.027	0.027	0.031	**0.000**	0.001
BU	**−0.1**	−2.9	0.6	0.3			**−0.001**	−0.005	0.000	0.000		

[a]All numbers give absolute values obtained from comparing the respective prior and posterior results. Accordingly, results of the sensitivity tests need to be compared to their specific references (bold) to interpret differences between simulations. All results are based on afternoon averaged (1400–1800 LT) data in daily time step.

[b]Ref, upgraded optimization setup as described in section 2.2; LM, optimization without Lagrange multipliers; Q_{diag}, optimization without off-diagonal elements in **Q** matrix; H_2O, surface fluxes without soil water influence on GPP; WRF 2006, reference simulation for September–Dec 2006, using WRF data; SOGS 2006, Simulation for September–Dec 2006 using SOGS data.

0.90, 0.94, and 0.99 for the MF, MP, and BU sites, respectively, indicating a successful adaptation of the uncertainty components for maximizing the use of information provided by the atmospheric observations.

Appendix C: Model Framework Sensitivity to Component Upgrades

[54] Four major components of the inverse modeling framework presented in this manuscript have been modified relative to *Göckede et al.* [2010], who first described the approach in detail: (1) inclusion of Lagrange multipliers into the optimization to enforce plausible parameter ranges; (2) addition of a variogram analysis to consider spatial correlations among prior base rate uncertainties, introduced as off-diagonal elements in the **Q** covariance matrix; (3) change of the data source for interpolated surface meteorology; (4) addition of a soil water scalar to improve simulation of drought stress effects on CO_2 assimilation. This appendix presents diagnostics to evaluate the impact of each of these changes on the model performance, analyzing the agreement of simulated atmospheric CO_2 mixing ratios with observations as well as the associated changes in statewide terrestrial CO_2 fluxes relative to posterior uncertainties.

[55] To test the sensitivity of the model performance with respect to each of the four major aspects, we used the full model version as described in section 2.2 as the reference scenario. Starting with this reference, in four separate case studies we reset each of these aspects to the setting described by *Göckede et al.* [2010], while keeping the other three aspects unchanged. For example, to isolate the influence of the Lagrange multipliers on the model performance, we ran an optimization without using Lagrange multipliers, but still including variogram results, the new data source for surface meteorology, and the soil water influence on GPP fluxes as described for the reference scenario. In three of these cases (Lagrange multipliers, variogram results, soil water influence on GPP), this simply meant turning off a new component in the upgraded modeling framework, while for the surface meteorology the test switched the new data source (WRF) to the previously used SOGS (Surface Observations Gridded System) [e.g., *Jolly et al.*, 2005]. Simulations are based on model drivers and observations from data years 2007 and 2008, except for the analysis on surface meteorology data sets, where simulations were run for the last four months of 2006 since SOGS is currently only available to us through the end of 2006.

[56] Exclusion of Lagrange multipliers had the most significant impact on the simulated annual statewide CO_2 flux budget, increasing the net uptake by an average of 39.1% (35.4% for 2007, 43.6% for 2008) when this component of the optimization algorithm was switched off. Flux changes introduced by the exclusion of variogram results as off-diagonal elements in the prior uncertainty matrix **Q** were only minor, with an average reduction of net uptake by −1.7% (−1.4, −1.8%). A slightly higher change in surface fluxes was observed when turning off the soil water influence on photosynthesis, where average statewide CO_2 uptake increased by 8.1% (7.5%, 8.9%). Switching surface meteorology from WRF to SOGS increased the terrestrial biosphere CO_2 source for the period September–December 2006 from −6.2 Tg C to −11.1 of Tg C, a difference of 4.9 Tg C (79.7%). However, we do not consider this difference as representative for the impact of the surface meteorology data source on the presented inverse modeling framework because (1) fluxes were only obtained for a partial year, (2) the total number of observations (i.e., afternoon averaged CO_2 mixing ratios from the period September–December 2006) was too small (n = 121) to provide solid results, and (3) relative changes are very high because of the small net budget for this short time period in fall and winter 2006. Accordingly, more temporal overlap between those two data sources is required to produce meaningful results. Evaluated against the assigned posterior uncertainties in simulated statewide CO_2 budgets (±17.3%; see also Table 4), only the exclusion of Lagrange multipliers produces significant changes (~3 standard deviations), while the other tests do not result in statistically significant results. For the simulation run excluding Lagrange multipliers, simulated flux fields also appeared to be less plausible compared to the reference case, with steeper gradients between different surface types.

[57] Table C1 summarizes the key statistics for the differences in observed minus prior CO_2 mixing ratios and observed minus posterior CO_2 mixing ratios, respectively. Compared to the reference simulation, these sensitivity studies produced very similar results in terms of RMSE reduction and changes in the coefficient of determination, except for the optimization excluding Lagrange multipliers where the reduction in RMSE between prior and posterior results was 3%–4% lower compared to the reference simulation, and the improvement in R^2 between prior and posterior was lower by 0.004–0.01. These findings emphasize that comparing statistics on simulated CO_2 time series can highlight major deficiencies in the inverse

model setup, as here in the case of the excluded Lagrange multipliers. At the same time, the optimization approach appears to be capable of finding plausible flux fields for various setups of the modeling framework, so analyzing only the simulated CO_2 time series will not be sufficient to correctly interpret the model performance.

[58] **Acknowledgments.** This research was supported by the Office of Science (BER), U.S. Department of Energy (DOE, grant DE-FG02-07ER64361), for the North American Carbon Program study, "Integrating Remote Sensing, Field Observations, and Models to Understand Disturbance and Climate Effects on the Carbon Balance of the West Coast U.S". We further thank Steven Wofsy (Harvard Univeristy), John Lin (University of Waterloo, Canada) and Christoph Gerbig (MPI for Biogeochemistry in Jena, Germany) for developing and providing the STILT software package, Janusz Eluszkiewicz and Thomas Nehrkorn (both AER Inc.) for mesoscale modeling support, James Kathilankal and Manuela Huso (both Oregon State University) for statistical support, Wouter Peters and Andrew Jacobson (both NOAA ESRL) for support with the CarbonTracker product, Kevin Gurney (Purdue University) for provision of the VULCAN fossil fuel data sets, Matthias Falk (UC Davis) for the provision of the Wind River eddy-covariance fluxes, Dave Ritts (Oregon State University) for remote sensing product support, Christoph Thomas (Oregon State University) for the provision of the Metolius eddy-covariance data set, Kent Davis, Jon Boro, and Matt Trappe (all Oregon State University) for setting up and maintaining the CO_2 mixing ratio measurement program, Anthony Svejcar and Raymond Angell (both USDA Agricultural Research Station, Burns, OR) for maintaining the observation site at Burns, Britton Stephens (NCAR) and Steven Wofsy for advice in the development of our continuous CO_2 mixing ratio monitoring systems, and Steven Wofsy for the provision of the COBRA-NA 2003 measurements.

References

Baker, D. F., et al. (2006), TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, *Global Biogeochem. Cycles*, *20*, GB1002, doi:10.1029/2004GB002439.

Barnes, R. J., and K. H. You (1992), Adding bounds to kriging, *Math. Geol.*, *24*(2), 171–176.

Cohen, W. B., et al. (2002), Characterizing 23 years (1972–1995) of stand replacement disturbance in western Oregon forests with Landsat imagery, *Ecosystems*, *5*(2), 122–137.

Cressie, N. A. C. (1993), *Statistics for Spatial Data*, 928 pp., Wiley, New York.

Daly, C., et al. (2008), Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, *Int. J. Climatol.*, *28*(15), 2031–2064.

Dolman, A. J., C. Gerbig, J. Noilhan, C. Sarrat, and F. Miglietta (2009), Detecting regional variability in sources and sinks of carbon dioxide: a synthesis, *Biogeosciences*, *6*(6), 1015–1026.

Enting, I. G. (2005), *Inverse Problems in Atmospheric Constituent Transport*, 412 pp., Cambridge Univ. Press, New York.

Friedlingstein, P., et al. (2006), Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison, *J. Clim.*, *19*(14), 3337–3353.

Gerbig, C., et al. (2003), 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.

Gill, P. E., W. Murray, and M. H. Wright (1986), *Practical Optimization*, 401 pp., Academic, London.

Göckede, M., A. M. Michalak, D. Vickers, D. Turner, and B. Law (2010), Atmospheric inverse modeling to constrain regional scale CO_2 budgets at high spatial and temporal resolution, *J. Geophys. Res.*, *115*, D15113, doi:10.1029/2009JD012257.

Gourdji, S. M., K. L. Mueller, K. Schaefer, and A. M. Michalak (2008), 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.

Gurney, K. R., et al. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, *415*(6872), 626–630.

Gurney, K. R., et al. (2004), Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks, *Global Biogeochem. Cycles*, *18*, GB1010, doi:10.1029/2003GB002111.

Gurney, K. R., et al. (2009a), *VULCAN Science Methods Documention, Version 1.1*, 38 pp, available online at http://www.purdue.edu/eas/carbon/vulcan/Vulcan.documentation.v1.1.pdf.

Gurney, K. R., et al. (2009b), High resolution fossil fuel combustion CO_2 emission fluxes for the United States, *Environ. Sci. Technol.*, *43*(14), 5535–5541.

Intergovernmental Panel on Climate Change (IPCC) (2007), *Climate Change 2007: Synthesis Report*, 52 pp., available online at http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf.

Irvine, J., B. E. Law, P. M. Anthoni, and F. C. Meinzer (2002), Water limitations to carbon exchange in old-growth and young ponderosa pine stands, *Tree Physiol.*, *22*(2–3), 189–196.

Jolly, W. M., J. M. Graham, A. Michaelis, R. Nemani, and S. W. Running (2005), A flexible, integrated system for generating meteorological surfaces derived from point sources across multiple geographic scales, *Environ. Modell. Software*, *20*(7), 873–882.

Karush, W. (1939), *Minima of Function of Several Variables with Inequalities and Side Constraints*, Univ. of Chicago, Chicago, Ill.

Kuhn, H. W., and A. W. Tucker (1951), Nonlinear Programming, in *Second Berkeley Symposium*, pp. 481–492, Univ. of California Press, Berkeley, Calif.

Lauvaux, T., et al. (2008), Mesoscale inversion: First results from the CERES campaign with synthetic data, *Atmos. Chem. Phys.*, *8*(13), 3459–3471.

Law, B. E. (2005), Carbon dynamics in response to climate and disturbance: Recent progress from multiscale measurements and modeling in AmeriFlux, in *Plant Responses to Air Pollution and Global Change*, edited by K. Omasa et al., pp. 205–213, Springer, Tokyo.

Law, B. E., and R. H. Waring (1994), Combining remote sensing and climatic data to estimate net primary production across Oregon, *Ecol. Appl.*, *4*(4), 717–728.

Law, B. E., et al. (2001), Spatial and temporal variation in respiration in a young ponderosa pine forest during a summer drought, *Agr. Forest Meteorol.*, *110*(1), 27–43.

Law, B. E., et al. (2004), Disturbance and climate effects on carbon stocks and fluxes across Western Oregon USA, *Global Change Biol.*, *10*(9), 1429–1444.

Lin, J. C., et al. (2003), 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.

Matross, D. M., et al. (2006), Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, *Tellus, Ser. B*, *58*(5), 344–358.

Michalak, A. M., and P. K. Kitanidis (2004), Application of geostatistical inverse modeling to contaminant source identification at Dover AFB, Delaware, *J. Hydraul. Res.*, *42*, 9–18.

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.

Michalak, A. M., et al. (2005), Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, *J. Geophys. Res.*, *110*, D24107, doi:10.1029/2005JD005970.

Mueller, K. L., S. M. Gourdji, and A. M. Michalak (2008), 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.

Peters, W., et al. (2005), An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, *J. Geophys. Res.*, *110*, D24304, doi:10.1029/2005JD006157.

Peters, W., et al. (2007), An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *Proc. Natl. Acad. Sci. U. S. A.*, *104*(48), 18,925–18,930.

Peters, W., et al. (2010), Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, *Global Change Biol.*, *16*(4), 1317–1337.

Peylin, P., et al. (2005), Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1. Inverse methodology, *Atmos. Chem. Phys.*, *5*, 3173–3186.

Prather, M. J., X. Zhu, S. E. Strahan, S. D. Steenrod, and J. M. Rodriguez (2008), Quantifying errors in trace species transport modeling, *Proc. Natl. Acad. Sci. U. S. A.*, *105*(50), 19,617–19,621.

Randerson, J. T., et al. (2005), Fire emissions from C_3 and C_4 vegetation and their influence on interannual variability of atmospheric CO_2 and $\delta^{13}CO_2$, *Global Biogeochem. Cycles*, *19*, GB2019, doi:10.1029/2004GB002366.

Rayner, P. J., I. G. Enting, R. J. Francey, and R. Langenfelds (1999), Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, *Tellus, Ser. B*, *51*(2), 213–232.

Riley, W. J., J. T. Randerson, P. N. Foster, and T. J. Lueker (2005), Influence of terrestrial ecosystems and topography on coastal CO_2 measure-

ments: A case study at Trinidad Head, California, *J. Geophys. Res.*, *110*, G01005, doi:10.1029/2004JG000007.
Rödenbeck, C., C. Gerbig, K. Trusilova, and M. Heimann (2009), A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, *Atmos. Chem. Phys.*, *9*(14), 5331–5342.
Running, S. W. (2008), Ecosystem Disturbance, Carbon, and Climate, *Science*, *321*(5889), 652–653.
Schuh, A. E., et al. (2010), A regional high-resolution carbon flux inversion of North America for 2004, *Biogeosciences*, *7*, 1625–1644.
Schwarz, P. A., B. E. Law, M. Williams, J. Irvine, M. Kurpius, and D. Moore (2004), Climatic versus biotic constraints on carbon and water fluxes in seasonally drought-affected ponderosa pine ecosystems, *Global Biogeochem. Cycles*, *18*, GB4007, doi:10.1029/2004GB002234.
Stoy, P. C., et al. (2009), Biosphere-atmosphere exchange of CO_2 in relation to climate: A cross-biome analysis across multiple time scales, *Biogeosciences*, *6*, 2297–2312.
Tans, P. P., I. Y. Fung, and T. Takahashi (1990), Observational constraints on the global atmospheric CO_2 budget, *Science*, *247*(4949), 1431–1438.
Tarantola, A. (1987), *Inverse Problem Theory. Methods for Data Fitting and Model Parameter Estimation*, 613 pp., Elsevier Sci., New York.
van der Werf, G. R., et al. (2006), Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmos. Chem. Phys.*, *6*(11), 3423–3441.
Wang, Y.-P., C. M. Trudinger, and I. G. Enting (2009), A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales, *Agric. For. Meteorol.*, *149*(11), 1829–1842.
Williams, M., et al. (2009), Improving land surface models with FLUXNET data, *Biogeosciences*, *6*(7), 1341–1359.

M. Göckede, B. E. Law, and D. P. Turner, Department of Forest Ecosystems and Society, College of Forestry, Oregon State University, 321 Richardson Hall, Corvallis, OR 97331, USA. (mathias.goeckede@oregonstate.edu)
A. M. Michalak, Department of Civil and Environmental Engineering, University of Michigan, 183 EWRE Bldg., Ann Arbor, MI 48109, USA.
D. Vickers, College of Oceanic and Atmospheric Sciences, Oregon State University, 104 COAS Administration Bldg., Corvallis, OR 97331, USA.