

## Atmospheric inverse modeling to constrain regional-scale CO<sub>2</sub> budgets at high spatial and temporal resolution

Mathias Göckede,<sup>1</sup> Anna M. Michalak,<sup>2,3</sup> Dean Vickers,<sup>4</sup> David P. Turner,<sup>1</sup> and Beverly E. Law<sup>1</sup>

Received 16 April 2009; revised 27 January 2010; accepted 15 March 2010; published 14 August 2010.

[1] We present an inverse modeling framework designed to constrain CO<sub>2</sub> budgets at regional scales. The approach captures atmospheric transport processes in high spatiotemporal resolution by coupling a mesoscale model with Lagrangian Stochastic backward trajectories. Terrestrial biosphere CO<sub>2</sub> emissions are generated through a simple diagnostic flux model that splits the net ecosystem exchange into its major components of gross primary productivity and autotrophic and heterotrophic respirations. The modeling framework assimilates state-of-the-art data sets for advected background CO<sub>2</sub> and anthropogenic fossil fuel emissions as well as highly resolved remote sensing products. We introduce a Bayesian inversion setup, optimizing a posteriori flux base rates for surface types that are defined through remote sensing information. This strategy significantly reduces the number of parameters to be optimized compared with solving fluxes for each individual grid cell, thus permitting description of the surface in a very high resolution. The model is tested using CO<sub>2</sub> concentrations measured in the fall and winter of 2006 at two AmeriFlux sites in Oregon. Because this database does not cover a full seasonal cycle, we focus on conducting model sensitivity tests rather than producing quantitative CO<sub>2</sub> flux estimates. Sensitivity tests on the influence of spatial and temporal resolution indicate that optimum results can be obtained using 4 h time steps and grid sizes of 6 km or less. Further tests demonstrate the importance of dividing biome types by ecoregions to capture their different biogeochemical responses to external forcings across climatic gradients. Detailed stand age information was shown to have a positive effect on model performance.

**Citation:** Göckede, M., A. M. Michalak, D. Vickers, D. P. Turner, and B. E. Law (2010), Atmospheric inverse modeling to constrain regional-scale CO<sub>2</sub> budgets at high spatial and temporal resolution, *J. Geophys. Res.*, *115*, D15113, doi:10.1029/2009JD012257.

### 1. Introduction

[2] The capability of the terrestrial biosphere to act as a source or sink for carbon has a significant effect on the assessment of carbon budgets on scales ranging from global to local [e.g., Baldocchi, 2008; Potter *et al.*, 2003; Stephens *et al.*, 2007]. Understanding the mechanisms that drive current biosphere carbon fluxes and pools plays a vital role in predicting future ecosystem functionality and atmospheric composition in the context of global change and thus in projecting the potential anthropogenic influence on global

climate [Intergovernmental Panel on Climate Change, 2007]. Common approaches to analyzing terrestrial biosphere-atmosphere exchange processes comprise eddy-covariance flux measurement networks [e.g., Baldocchi *et al.*, 2001; Law, 2005; Valentini *et al.*, 2000] and bottom-up biogeochemistry modeling frameworks that ingest multiple data sources including atmospheric observations, ecosystem inventories, and remote sensing data [Krinner *et al.*, 2005; Potter, 1999; Thornton *et al.*, 2002]. Eddy-covariance measurements have provided detailed insight into the flux mechanisms between the surface and the atmosphere [e.g., Law *et al.*, 2002; Reichstein *et al.*, 2007], but the information is representative only at local scales [e.g., Rannik *et al.*, 2000; Schmid, 1997; Schuepp *et al.*, 1990]. Bottom-up modeling has delivered valuable insight into spatially variable flux processes on different scales, but validation of the models is usually restricted to single, well-instrumented observation sites, and evaluation of the spatial products remains challenging.

[3] Atmospheric inverse modeling [e.g., Bousquet *et al.*, 2000; Enting, 2005; Rayner *et al.*, 1999] holds the poten-

<sup>1</sup>Department of Forest Ecosystems and Society, College of Forestry, Oregon State University, Corvallis, Oregon, USA.

<sup>2</sup>Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan, USA.

<sup>3</sup>Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, Michigan, USA.

<sup>4</sup>College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA.

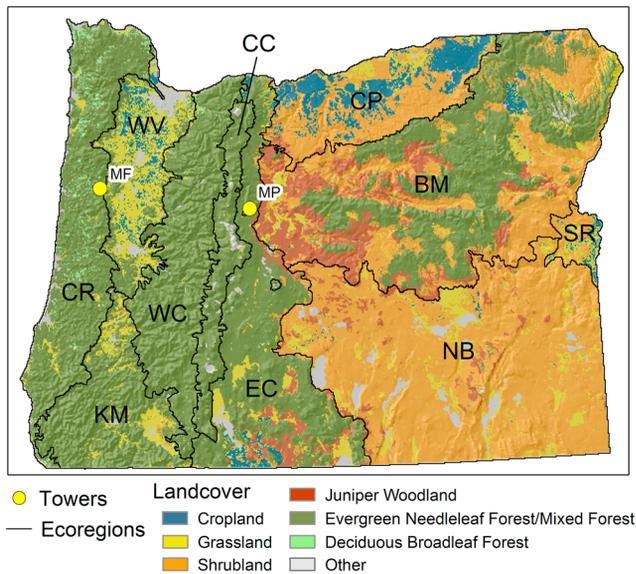
tial to overcome the shortcomings listed above. The technique allows one to extract information from atmospheric CO<sub>2</sub> observation networks to either directly constrain carbon budget estimates on various scales [e.g., *Gurney et al.*, 2002; *Michalak et al.*, 2004] or provide independent reference data to evaluate and optimize biosphere carbon flux models [e.g., *Gerbig et al.*, 2003; *Matross et al.*, 2006]. CO<sub>2</sub> concentration data may be representative of spatial scales ranging from regional to continental [*Gerbig et al.*, 2006; *Gloor et al.*, 2001], so the variability captured in the time series integrates the influence of biospheric activity on atmospheric composition over large areas. To date, most inverse atmospheric modeling studies have addressed global [e.g., *Gourdji et al.*, 2008; *Gurney et al.*, 2002; *Jacobson et al.*, 2007; *Mueller et al.*, 2008] to continental [e.g., *Peters et al.*, 2007; *Peylin et al.*, 2005] domains, whereas others used a pseudodata approach to evaluate the capabilities of their inversion setups to constrain carbon budgets [e.g., *Gourdji et al.*, 2009; *Lokupitiya et al.*, 2008; *Schuh et al.*, 2009; *Zupanski et al.*, 2007]. Concerning temporal resolution, inverse modeling approaches have been applied to solve for time steps far smaller than annual, e.g., monthly to daily varying fluxes [e.g., *Baker et al.*, 2006; *Gourdji et al.*, 2008; *Mueller et al.*, 2008]. These studies have successfully demonstrated the potential of this technique to assess overall carbon budgets and their distribution among subregions including temporal dynamics and uncertainties. However, because of the aggregation of fluxes over large regions, these approaches are usually too coarse to capture vegetation and climate variability at small subgrid scales, and thus these can provide little information on the mechanisms that drive the underlying carbon processes.

[4] Application of atmospheric inverse modeling on regional scales [e.g., *Gerbig et al.*, 2003; *Lauvaux et al.*, 2008; *Matross et al.*, 2006] permits a model setup that defines surface processes in enough detail to investigate the effect of climate anomalies (e.g., drought) or disturbance history (e.g., wildfires) on carbon fluxes and pools. This way, highly detailed prior information, e.g., from remote sensing sources, can be assimilated into the modeling framework, although the final resolution of information gained through the inverse modeling approach also depends on the setup of the atmospheric observation network. Regional-scale applications could potentially be used to monitor compliance with greenhouse gas reduction goals in the context of state-level greenhouse gas targets or international treaties. However, the high spatial resolution in the model domain setup requires an optimization strategy that limits the degrees of freedom associated with the increased number of grid cells [e.g., *Gerbig et al.*, 2006]. Also, improvements in aggregation error [*Kaminski et al.*, 2001] with high spatial resolution models may be canceled out by increased uncertainties in transport simulations [*Lin and Gerbig*, 2005] or vertical boundary layer mixing [*Gerbig et al.*, 2008]. To avoid underconstrained model setups, regional-scale inverse modeling needs to assimilate additional sources of information, such as a realistic description of the model domain from remote sensing, and an accurate definition of a prior flux model version that has been trained on eddy-covariance flux measurements.

[5] Here, we present a regional-scale inverse modeling approach developed to constrain CO<sub>2</sub> budgets in the U.S. West Coast region. This approach uses an optimization strategy that has been customized to capture the fine-scale spatial heterogeneity in surface fluxes that may have an important effect in regional-scale inverse modeling studies. One central element of this strategy is to separate the model domain surface into so-called surface types defined through remote sensing data layers for ecoregion, land cover type, and disturbance history. This strategy decouples the number of parameters to be optimized, and thus the degrees of freedom in the optimization process, from the number of grid cells or regions in the model domain and therefore makes possible a highly detailed description of the surface domain that reduces potential representation errors. Other characteristics of our inverse modeling approach include the optimization of flux base rates instead of the fluxes themselves, with individual base rates assigned to each of the major flux components of gross primary productivity (GPP), autotrophic respiration (RA), and heterotrophic respiration (RH). The modeling framework assimilates high-resolution (<1 km grid size) remote sensing data sets to characterize the surface and is operated in subdaily time steps to capture information from the observed diurnal cycle of atmospheric CO<sub>2</sub> concentration. Atmospheric transport modeling to link receptor locations to spatially distributed sources is solved in high spatiotemporal resolution by coupling the Weather Research and Forecast (WRF) mesoscale model to the Stochastic Time-Inverted Lagrangian Transport (STILT) model. The simple diagnostic carbon flux model splits total carbon flux into its main components of GPP, RA, and RH and takes into account important influence factors such as drought stress, stand age, or disturbance history. A Bayesian approach is applied to optimize the flux base rates. Because the database of well-calibrated atmospheric CO<sub>2</sub> concentrations available for modeling does not yet cover a full seasonal cycle, we refrain from computing annual CO<sub>2</sub> budgets for the study region. Instead, this study focuses on sensitivity studies of the influence of spatial and temporal resolution as well as the role of complexity in the domain surface setup.

## 2. Atmospheric Inverse Modeling Approach

[6] The atmospheric inverse modeling framework presented in this section follows the general concept proposed by *Gerbig et al.* [2003]. Atmospheric transport modeling (section 2.4) is used to develop an influence function that links a receptor location to spatially distributed sources and sinks. This influence function is coupled to modeled terrestrial biosphere fluxes of CO<sub>2</sub> (section 2.3) to obtain the effect of photosynthesis and respiration on the atmospheric CO<sub>2</sub> concentration time series. Also considering anthropogenic fossil fuel emissions and advected background concentrations (section 2.5), this approach allows one to simulate the time series of CO<sub>2</sub> concentration for any given location and time frame within the model domain. The optimization strategy (section 2.6) builds on Bayesian inversion to optimize the correlation between modeled and observed (section 2.2) CO<sub>2</sub> concentration time series by improving flux base rates for individual surface types in the biosphere flux model.



**Figure 1.** Ecoregions and land cover types in the Oregon, USA, model domain. Topography is indicated by hillshading underneath the landcover colors. Yellow dots mark the locations of CO<sub>2</sub> concentration monitoring sites (MF, Mary’s River mature fir site; MP, Metolius mature pine site). Please refer to Table 1 for the list of ecoregion abbreviations.

## 2.1. Oregon Model Domain

[7] Our study focuses on the state of Oregon, located in the Pacific Northwest Region of the United States (Figure 1). This model domain is characterized by a significant small-scale to mesoscale variability in vegetation characteristics that poses challenges to carbon cycle modeling. The crest of the Cascade Mountains roughly splits the state into a mesic western part dominated by dense, managed Douglas-fir forest and agricultural crops, and a semiarid eastern part mainly consisting of open ponderosa pine forest and juniper-

sagebrush-grass communities. The whole state is subject to drought during the summer growing season, primarily in the central to eastern parts [e.g., Irvine *et al.*, 2002; Law *et al.*, 2001; Schwarz *et al.*, 2004]. Predominant clear-cut management practices in the western third of the state and frequent wildfires in the East created a small-scale mosaic of age classes on forested lands [Cohen *et al.*, 2003]. The Portland metropolitan area in the northwest corner of Oregon is the only major source of anthropogenic carbon emissions. Because forests in the Pacific Northwest are among the most productive globally [Luyssseart *et al.*, 2008; Myneni *et al.*, 2001], with intensive ongoing research on the role of climate change or management practices [Campbell *et al.*, 2009; Donato *et al.*, 2006; 2009; Irvine *et al.*, 2007], Oregon offers a highly relevant domain for studies on regional carbon balances.

[8] On the basis of different remote sensing sources (Table 1), we identify 10 ecoregions, 6 vegetation land cover classes, and 2 disturbance regimes for the Oregon domain. Hereafter, we use the term “surface type” to refer to any combination of ecoregion, land cover class, and disturbance regime. A stand age map was produced from Landsat data [Cohen *et al.*, 2002], where change detection was used for forests less than 30 years old, and spectral regression was used for forests older than 30 years. To increase computation efficiency, the original pixel resolution of 25 m was aggregated to a 1 km resolution, storing the five dominant combinations of surface type and stand age and their coverage percentage for each grid cell.

## 2.2. Atmospheric CO<sub>2</sub> Concentration Data

[9] Continuous well-calibrated atmospheric CO<sub>2</sub> concentration measurements from two monitoring sites in Oregon are used as input for the inverse modeling approach (Figure 1). The Metolius mature pine site [MP, 44.45°N, 121.56°W, 1310 m above sea level (asl)] is situated about 17 km north of the town of Sisters, Oregon, in the semiarid East Cascades ecoregion [Irvine *et al.*, 2007; Law *et al.*, 2004]. The inlet height of the CO<sub>2</sub> monitoring system is 33.5 m above

**Table 1.** Survey on Remote Sensing–Based Data Used to Describe Surface Characteristics in Oregon

Data Type	Number of Classes	Classes	Data Source
Ecoregion	10	CR, Coastal Range KM, Klamath Mountains WV, Willamette Valley WC, West Cascades CC, Cascade Crest <sup>a</sup> EC, East Cascades CP, Columbia Plateau BM, Blue Mountains NB, Northern Basin and Range SR, Snake River	EPA Level III Ecoregions ( <a href="http://www.epa.gov/wed/pages/ecoregions.htm">http://www.epa.gov/wed/pages/ecoregions.htm</a> )
Vegetation land cover types	6	ENF/MF, evergreen needleleaf forest/mixed forest DBF, deciduous broadleaf forest WOOD, juniper woodland SHRUB, shrubland  GRASS, grassland CROP, cropland	Forest types: Law <i>et al.</i> [2004]  Nonforest types: U.S. Geological Survey’s National Land Cover Database ( <a href="http://landcover.usgs.gov/nlcd.php">http://landcover.usgs.gov/nlcd.php</a> ). Compared with the original data, some classes herein were combined and reclassified.
Disturbance regime	2	FIRE, wildfire CUT, harvest/clear-cut	Law <i>et al.</i> [2004] Cohen <i>et al.</i> [2002]

<sup>a</sup>Ecoregion added for this study as transition zone between WC and EC.

ground level (agl). The Mary's River mature Douglas-fir site (MF, 44.65°N, 123.55°W, 310 m asl) lies about 6 km north of Blodgett, Oregon, in the Oregon Coast Range ecoregion. The initial inlet height for the CO<sub>2</sub> monitoring system of 46.6 m agl was relocated to 37.9 m agl on 26 September 2006. The two sites are equipped with custom-built CO<sub>2</sub> monitoring systems following the design by *Stephens et al.* [2006], with major modifications to stabilize the output against fluctuations in pressure, air temperatures, and water vapor. The core of the system is a LiCor LI-820 (MP site) or LI-840 (MF site) gas analyzer operating at a measurement frequency of 1 Hz. Calibration and quality control of the raw data are based on four standard gases sampled at regular intervals. Final output data are hourly averaged atmospheric CO<sub>2</sub> mixing ratios.

[10] The CO<sub>2</sub> monitoring systems were installed at both sites in August 2006. The availability of ancillary data sets required to run the atmospheric inversion (see below) restricted the present study to the latter half of 2006. Quality filters are applied to identify situations with insufficient vertical mixing as well as the transition period from a convective boundary layer in the late afternoon to the early evening stable boundary layer. We applied a disjunct tall tower concept (Appendix A) to identify situations where capping inversions disconnected the short tower observations from the upper portions of the boundary layer. Data gaps and the exclusion of the flagged data significantly reduced both data sets (MF = 21% of data remaining and MP = 24% of data remaining). Because wildfire emissions cannot be quantified yet because of their unknown source strength, measurements downstream of actively burning areas (<http://geomac.usgs.gov/>) were also excluded from the analysis. Because this reduction of the data set only removes the immediate effect wildfires have on the atmospheric mixing ratios, the capability of our modeling framework to evaluate the long-term influence of fire disturbance on subsequent terrestrial fluxes (photosynthesis and respiration) is not influenced by this filter.

### 2.3. Terrestrial Biosphere CO<sub>2</sub> Flux Model

[11] The terrestrial biosphere CO<sub>2</sub> flux model (further on referred to as BioFlux) assimilates information from remote sensing platforms, interpolated surface meteorology data sets, and eddy-covariance flux sites to simulate net ecosystem exchange (NEE) between the vegetation and the atmosphere. The model resolves CO<sub>2</sub> fluxes in hourly time steps into GPP, RA, and RH. BioFlux includes influences of forest disturbance history and stand age as well as drought stress on NEE.

#### 2.3.1. Flux Algorithms

[12] The calculation of GPP is based on the product of light use efficiency and absorbed photosynthetically active radiation (APAR), with additional scaling factors simulating the influence of temperature, vapor pressure deficit, cloud cover, and stand age on photosynthesis:

$$\text{GPP} = -\text{GPP}_{\text{base}} \cdot \text{APAR} \cdot T_{\text{sc}} \cdot \text{VPD}_{\text{sc}} \cdot (1 + \text{CL}_{\text{wgt}} \cdot \text{CL}_{\text{sc}}) \cdot \text{Age}_{\text{GPP}}, \quad (1)$$

where

- GPP gross primary productivity (g C m<sup>-2</sup> time step<sup>-1</sup>);
- GPP<sub>base</sub> base rate for gross primary productivity (g C MJ<sup>-1</sup>);
- APAR absorbed photosynthetically active radiation (MJ m<sup>-2</sup> time step<sup>-1</sup>);
- $T_{\text{sc}}$  minimum temperature scaling factor (-);
- VPD<sub>sc</sub> vapor pressure deficit scaling factor (-);
- CL<sub>wgt</sub> cloudiness influence weight (-);
- CL<sub>sc</sub> cloudiness scaling factor (-);
- Age<sub>GPP</sub> age scaling factor on GPP (-).

APAR is derived as the product of incident photosynthetically active radiation (PAR) and the fraction of available radiation in the photosynthetically active wavelengths that a canopy absorbs (fPAR), with a time step that may vary between half-hourly and daily. The functions for minimum temperature ( $T_{\text{sc}}$ ) and vapor pressure deficit (VPD<sub>sc</sub>) are derived from sigmoid functions that can be adapted to biome characteristics by fitting two parameters: inflection point and influence width (Appendix B). The sigmoid functions, ranging between 0 and 1, follow the ramp function approach described, e.g., by *Running et al.* [2000], while keeping the equations continuous and thus differentiable. The cloudiness scaling factor (CL<sub>sc</sub>) is obtained from the ratio of actual to potential incoming shortwave radiation, ranging between 0 for clear skies and 1 for complete overcast. It has been included in the equation to simulate the effect of diffuse radiation on photosynthesis. The influence of CL<sub>sc</sub> on GPP can be adapted for each biome by optimizing the cloudiness influence weight (CL<sub>wgt</sub>). The age scaling factor (Age<sub>GPP</sub>) is included to account for reduced net primary productivity in older forest stands (Appendix B).

[13] RA is calculated as the sum of a temperature-dependent maintenance respiration and a growth respiration component that reflects assimilated carbon available for growth:

$$\text{RA} = R_{\text{m}} + R_{\text{g}}, \quad (2)$$

$$R_{\text{m}} = R_{\text{m, base}} \cdot Q_{10}^{[(T_{\text{air}} - 20)/10]} \cdot \text{fPAR}, \quad \text{and} \quad (3)$$

$$R_{\text{g}} = (\text{GPP} - R_{\text{m}}) \cdot R_{\text{g, frac}}, \quad (4)$$

where

- RA autotrophic respiration (g C m<sup>-2</sup> time step<sup>-1</sup>);
- $R_{\text{m}}$  maintenance respiration (g C m<sup>-2</sup> time step<sup>-1</sup>);
- $R_{\text{g}}$  growth respiration (g C m<sup>-2</sup> time step<sup>-1</sup>);
- $R_{\text{m, base}}$  base rate for maintenance respiration (g C m<sup>-2</sup> time step<sup>-1</sup>);
- $Q_{10}$  base rate for  $Q_{10}$  temperature influence function (-);
- $T_{\text{air}}$  actual air temperature (°C);
- fPAR fraction of PAR absorbed by the canopy (-);
- $R_{\text{g, frac}}$  fraction of assimilated carbon used in growth respiration (-).

[14] Maintenance respiration is the product of a biome-specific base rate ( $R_{m, \text{base}}$ ) and the  $Q_{10}$  function that reflects the influence of temperature on respiration processes. As an additional scaling factor, fPAR is included as a proxy for leaf area index to reflect the influence of live biomass on the respiration fluxes. Growth respiration is set to a fixed ratio ( $R_{g, \text{frac}} = 0.25$ ) of assimilated carbon available after maintenance respiration has been deducted from GPP.

[15] RH is modeled using a biome-specific base rate scaled with actual soil temperature, soil moisture, stand age, and fPAR:

$$\text{RH} = \text{RH}_{\text{base}} \cdot \text{TS}_{\text{sc}} \cdot \text{SW}_{\text{sc}} \cdot \text{Age}_{\text{RH}} \cdot \text{fPAR}, \quad (5)$$

where

RH heterotrophic respiration ( $\text{g C m}^{-2} \text{ time step}^{-1}$ );  
 $\text{RH}_{\text{base}}$  base rate for heterotrophic respiration ( $\text{g C m}^{-2} \text{ time step}^{-1}$ );  
 $\text{TS}_{\text{sc}}$  soil temperature scaling factor (-);  
 $\text{SW}_{\text{sc}}$  soil water scaling factor (-);  
 $\text{Age}_{\text{RH}}$  age scaling factor on RH (-).

The scaling factors for soil temperature and soil moisture are derived using exponential functions (Appendix B), forcing higher RH fluxes with increasing temperatures and soil water content (SWC). A specific age function was used to capture increased decomposition of biomass in the years after disturbance events such as clear-cuts or fires. fPAR has been included in the algorithm as an indicator for the effect of recently assimilated carbon on RH. fPAR values smaller than 0.2 are set to 0.2 to guarantee a minimum flux rate even in sparsely vegetated areas.

### 2.3.2. Model Initialization

[16] The BioFlux model simulates carbon fluxes in regions characterized by fine-scale heterogeneity in surface types (see also section 2.1). Model algorithms include eight parameters that can be adjusted to represent the response of each surface type to environmental drivers: base rates of GPP, RA, and RH ( $\text{GPP}_{\text{base}}$ ,  $R_{m, \text{base}}$ , and  $\text{RH}_{\text{base}}$ ), inflection point and influence width for both the scaling parameters of  $T_{\text{min}}$  and  $\text{VPD}_d$  ( $T_{\text{inf}}$ ,  $T_{\text{wid}}$ ,  $\text{VPD}_{\text{inf}}$ , and  $\text{VPD}_{\text{wid}}$ ), and the cloudiness influence weight  $\text{CL}_{\text{wgt}}$ . Initial values for all eight parameters are derived by fitting BioFlux results to reference data sets (see below), whereas only the three base rates will subsequently be modified based on results of the atmospheric inversion approach (see section 2.6).

[17] The model is initialized in a two-stage process, using site-level reference flux data of GPP, RA, and RH in daily time steps. In the first stage, we use the Shuffled Complex Evolution – University of Arizona (SCE-UA) [Duan *et al.*, 1992, 1993] algorithm to identify the optimum parameter values for each reference data set. SCE-UA explores an  $n$ -dimensional solution space that is defined by upper and lower limits for each parameter and does not require prior knowledge on optima as a starting point. The Simplex downhill search algorithm used by SCE-UA effectively evolves the population of solutions toward a single best parameter set, mostly neglecting regions with lower posterior density. Because these regions of lower posterior density are important to characterize the parameter distributions,

we added the Shuffled Complex Evolution Metropolis (SCEM) algorithm [Vrugt *et al.*, 2003] as a second initialization stage to derive the parameter uncertainties required for the Bayesian inversion (see section 2.6). SCEM builds on SCE-UA but introduces a number of modifications, the most important of which is the replacement of the Simplex downhill search method with the Metropolis algorithm [Metropolis *et al.*, 1953]. The overall effect is that SCEM efficiently explores lower posterior density regions of the parameter space, calculating several tens of thousands of parameter sets and their specific posterior densities. To avoid bias introduced by the randomly chosen starting population of solutions, we used only the final 50% of the parameter sets to characterize parameter uncertainties. Only the three flux base rates per surface type were optimized in the Bayesian inversion (section 2.6); the remaining five parameters were kept constant during SCEM runs.

[18] Two AmeriFlux eddy-covariance sites were available to provide reference data for the parameter initialization. The first is the Metolius mature pine site (MP) (see also section 2.2), providing 4 years of flux data to initialize evergreen needleleaf and mixed forest (ENF/MF) biomes in the Eastern Oregon ecoregions. The second site, Wind River (45.82°N, 121.95°W, 371 m asl), is located in the humid West Cascades ecoregion just north of the Oregon border [Falk *et al.*, 2008; Shaw *et al.*, 2004]. Three years of Wind River flux data were used to initialize ENF/MF biomes in the Western Oregon ecoregions. NEE observations were partitioned into GPP and ecosystem respiration (RE), with RE further split into its two major components (RA and RH) using a fixed GPP/RA ratio. The total coverage percentage of ENF/MF biomes initialized by eddy-flux data is 43.8% of the Oregon domain.

[19] For the initialization of nonconifer biomes, GPP, RA, and RH from the Biome-BGC model [Thornton *et al.*, 2002] was used as reference data. These reference model runs were based on extensive previous measurement and modeling studies in the Oregon region [e.g., Law *et al.*, 2004, 2006; Turner *et al.*, 2003, 2004] and considered the same information to describe the model domain as in this study (see Turner *et al.* [2007] for details). We conducted Biome-BGC runs for 14 selected locations distributed throughout Oregon, each of which represented a single nonforest surface type with significant coverage in the model domain. These 14 surface types cover another 48.0% of the Oregon domain, and results were extrapolated into the remaining 8.2% for which no reference data were available.

### 2.3.3. Biosphere Flux Model Input Data

[20] The main source of spatial surface meteorology data to drive BioFlux in this study is the Surface Observations Gridded System (SOGS) [Jolly *et al.*, 2005], which interpolates site-level surface meteorology from various sources on the basis of topography. SOGS provides daily minimum and maximum temperatures, average incoming shortwave radiation, average VPD, and precipitation in continuous grids with 1 km resolution. SOGS precipitation data were improved on the basis of comparison to monthly time series data from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) [e.g., Daly *et al.*, 2008] climate mapping system, creating a hybrid product that



**Figure 2.** Spatial coverage of (black boxes) two nested Weather Research and Forecast domains used as input data for the STILT model in this study.

combines the high spatial and temporal resolution of SOGS with the knowledge-based high-quality PRISM data. Day-Met [Thornton and Running, 1999; Thornton *et al.*, 1997] was used as an additional data source to provide the daytime length between sunrise and sunset.

[21] Daily SOGS data must be interpolated into subdaily time steps to allow computation of surface CO<sub>2</sub> fluxes with a time step of 1 h. Appendix C describes our solution to approximate daily courses of incoming PAR and temperature using average daily radiation, daily maximum and minimum temperatures, geographic position, as well as time and date. Additional information on input data preparation for use in BioFlux is given in Appendix B.

[22] Gridded Moderate Resolution Imaging Spectroradiometer (MODIS) fPAR data (1 km spatial resolution and 8 day temporal resolution) were downloaded from NASA archives (<https://wist.echo.nasa.gov/api/>). A temporal interpolation routine [Zhao *et al.*, 2005] was applied to fill missing data or data flagged as low quality by the quality assurance protocol.

#### 2.4. Atmospheric Transport Modeling

[23] Atmospheric transport modeling is required to solve for the influence of a changing “field of view” or source area [e.g., Gash, 1986; Pasquill, 1972] on atmospheric measurements in heterogeneous terrain. This spatial context of a measurement, commonly defined as the footprint [e.g., Schmid, 2002; Schmid and Oke, 1990; Schuepp *et al.*, 1990], is described by a transfer function that links site-level observations to the surrounding terrain, helping to

explain fluctuations in the observed signal caused by a varying composition of sources and sinks within the source area. The atmospheric transport module couples the mesoscale atmospheric model WRF (<http://www.wrf-model.org>) with the receptor-oriented atmospheric transport model STILT [Lin *et al.*, 2003].

[24] WRF is a mesoscale atmospheric model that can be used for both operational forecasting and atmospheric research. It has been set up here to generate refined three-dimensional transport fields as offline input for the STILT model, stored in 20 min intervals, with a spatial resolution corresponding to the specific WRF grid resolution (see below). Adding WRF as a component to the atmospheric transport module allows computation of customized high-resolution meteorological fields that conserve mass, momentum, entropy, and scalars and include parameterized convective mass fluxes. We use the Advanced Research WRF Version 2.0 (described, e.g., in Skamarock *et al.* [2005] and Wang *et al.* [2005]) in a software package customized for coupling with STILT. Details on the physics schemes used in our WRF setup are given in Table 2. Initial and boundary conditions were taken from the National Centers for Environmental Prediction (NCEP) final global analyses (<http://dss.ucar.edu/datasets/ds083.2>), which is available in 1° × 1° resolution on 26 pressure levels 4 times daily. We selected a nested design with two domains (Figure 2): an outer grid covering all of Oregon in 12 km resolution (70 × 60 cells, 60 s time step) and an inner grid focusing on the central Western part of the state with a resolution of 4 km (73 × 55 cells, 20 s time step). The number of vertical levels is 27 for both domains, and feedback between the grids is activated. A continuous model run covered the entire observation period (August to December 2006), with Four-Dimensional Data Assimilation [e.g., Stauffer and Seaman, 1994; Stauffer *et al.*, 1991] grid nudging activated above the planetary boundary layer for the outer grid simulations to align the model to the NCEP reference data 4 times daily. Nudging coefficients were set to 1 × 10<sup>-4</sup> s<sup>-1</sup> for wind and temperature and 1 × 10<sup>-5</sup> s<sup>-1</sup> for moisture, as recommended by Deng and Stauffer [2006].

[25] STILT computes the field of view of atmospheric CO<sub>2</sub> concentration measurements by releasing an ensemble of particles at the receptor point and following their trajectories backward in time. Integrating the positions of these particles specifies the relative influence of each surface pixel on the concentration measurements. These source weight functions can then be combined with spatially and temporally explicit estimates of terrestrial carbon fluxes to simu-

**Table 2.** Physics Schemes Used in WRF Model Setup

Category	Selected Options
Microphysics	WRF single-moment three-class scheme
Long-wave radiation	Rapid radiative transfer model
Short-wave radiation	Dudhia scheme
Surface layer	Monin-Obukhov (Janjic Eta) scheme
Land surface	NOAH land-surface model (unified NCEP/NCAR/AFWA scheme)
Boundary layer	Mellow-Yamada-Janjic (Eta) TKE scheme
Cumulus parameterization	Grell-Devenyi ensemble scheme

late atmospheric CO<sub>2</sub> concentration changes. In our study, these surface flux fields consisted of biosphere CO<sub>2</sub> fluxes computed by the BioFlux model and anthropogenic CO<sub>2</sub> emissions provided by fossil fuel inventories (see section 2.5). The magnitude of the simulated CO<sub>2</sub> concentration change is dependent on local source/sink strength, the size of the target volume, and the particle residence time within this target volume. The spatial integral over the source weight function gives the total concentration change the ensemble experiences as it passes through the model domain. Provided the initial CO<sub>2</sub> concentration of each air parcel is known (see also section 2.5), this approach allows simulating atmospheric CO<sub>2</sub> concentrations for any location and time step within the model domain.

[26] STILT requires three-dimensional transport fields as input, generated in our study using the WRF model. A total of 250 particles were released at the receptor height in hourly intervals and followed backward in time for 72 h or until they reached the boundary of the outer WRF model domain (Figure 2). For footprint computation, the particle positions were projected on a map with a resolution of  $0.01^\circ \times 0.01^\circ$ ; however, the horizontal size of the grid cells in the footprint area was dynamically adjusted with increasing distance from the receptor location (see *Gerbig et al.* [2003] for details), so that the highest resolution was only applied in close vicinity to the tower, whereas in the outer areas of the footprint, grid cells could reach a maximum size of  $0.25^\circ \times 0.25^\circ$ . The column height of the target volume was set to be 0.5 times the boundary layer height. Whenever the vertical position of a particle was smaller than this column height at its specific location, the particle was assumed to be affected by surface fluxes. The use of convective mass fluxes provided by WRF was activated such that STILT sampled the vertical profiles of mass fluxes within the updrafts and downdrafts computed by WRF, and the STILT particles followed these updrafts or downdrafts with probabilities proportional to the WRF mass fluxes.

## 2.5. Fossil Fuel Emission Data and CO<sub>2</sub> Boundary Conditions

[27] The anthropogenic CO<sub>2</sub> emissions are taken from a gridded data set provided in a spatial resolution of  $10 \times 10$  km by the Vulcan project [*Gurney et al.*, 2008]. Vulcan analyses are based on data from the Clearinghouse for Inventories and Emission Factors (<http://www.epa.gov/ttn/chief/index.html>) provided by the U.S. Environmental Protection Agency (EPA) and include additional information on mobile sources, power plants, and U.S. census data [*Gurney*, 2008]. All data are interpolated to a 10 km grid resolution with emissions in hourly time steps for the year 2002, which is the year of the Vulcan spatial analysis. To allow extrapolation into different years in the context of this study, we aggregated the data into hourly averages for weekdays, Saturdays, and Sundays of each month. State-level emission inventories, which can be used to compute scaling factors for this temporal extrapolation, are currently only available through 2005; however, because United States-wide inventories, which are available through 2006 [*Environmental Protection Agency*, 2008], indicate rather stable emission rates during the years 2000–2006 (total CO<sub>2</sub> emissions

increased by 1.3% between 2002 and 2006), we assumed no significant changes in emissions between 2002 and 2006.

[28] To simulate absolute CO<sub>2</sub> concentrations at a given receptor location, initial concentrations need to be added to the concentration changes that the ensemble of particles experiences within the model domain (section 2.4). These initial concentrations are assigned in the same way for particle trajectory starting locations inside or outside the model domain, respectively. CO<sub>2</sub> concentration boundary conditions are taken from the high-resolution North American grid of the 2007B release of CarbonTracker [*Peters et al.*, 2005, 2007]. Built by the National Oceanic and Atmospheric Administration's Earth System Research Laboratory, CarbonTracker generates continuous CO<sub>2</sub> mole fraction grids by coupling surface CO<sub>2</sub> exchange models to atmospheric transport modeling and optimizing results against atmospheric CO<sub>2</sub> observations using an ensemble Kalman filter. The CarbonTracker output used for this study provides four-dimensional grids of CO<sub>2</sub> mole fractions in  $1^\circ \times 1^\circ$  horizontal resolution, 34 vertical levels, and six-hourly temporal resolution.

## 2.6. Parameter Optimization

### 2.6.1. Classical Bayesian Approach

[29] We apply a classical Bayesian approach [e.g., *Enting*, 2005; *Tarantola*, 1987] to find an optimum base rate set for the BioFlux model using information extracted from the observed atmospheric CO<sub>2</sub> concentrations. The optimization strategy has the following basic characteristics: First, we optimize flux base rates, which can be interpreted as sensitivities to external drivers such as radiation and temperature, so we train a flux model instead of adjusting the fluxes themselves. Second, NEE is divided into its major components GPP, RA, and RH in the optimization process, assigning individual scaling factors to the base rate of each of these. A similar concept was applied previously by, e.g., *Bousquet et al.* [1999] and *Zupanski et al.* [2007], who decomposed NEE into photosynthesis and RE in their optimization strategy. Third, all optimized flux base rates are assumed to be constant over time. This strategy is based on the hypothesis that seasonal trends and phenological states are well captured through the meteorological drivers and the MODIS fPAR product, leaving a flux base rate that is uniform in time. Fourth, the model domain is classified into surface types, each with a uniform set of base rates. This feature of our approach decouples the number of parameters to be optimized from the horizontal resolution of the surface grid (see also section 3.3), thus enabling fine-scale variability in the setup of the model domain. Structuring the model domain into surface types is thus an important piece of prior information we feed into the optimization routine to maximize the use of information provided by the atmospheric observations. The setup of surface types, however, needs to be chosen carefully to avoid aggregation errors caused by inhomogeneous flux mechanisms within assigned surface types. To fully exploit the benefits provided by the highly detailed surface description in the optimization; that is, to be able to optimize the surface types in remote regions of the domain, a higher density of sites would be required than used in this study.

[30] Flux base rates (one each for the fluxes of GPP, RA, and RH for every surface type; see also section 2.3.2) were optimized on the basis of information extracted from atmospheric observations and prior flux estimates, by finding a minimum of the Bayesian cost function  $L_s$  [e.g., Enting, 2005; Michalak *et al.*, 2004]:

$$L_s = \frac{1}{2}(\mathbf{z} - \mathbf{H}\mathbf{s})^T \mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}) + \frac{1}{2}(\mathbf{s} - \mathbf{s}_p)^T \mathbf{Q}^{-1}(\mathbf{s} - \mathbf{s}_p), \quad (6)$$

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 concentrations ( $n \times m$  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);
- $m$  number of parameters to be optimized;
- $n$  number of observations.

In this study, the number of parameters to be optimized,  $m$ , equals the number of surface types in the domain setup (e.g., 120 in the base case scenario) multiplied by 3 (one base rate each for GPP, RA, and RH). The Jacobian transfer function,  $\mathbf{H}$ , reflects the influence of the spatially distributed biospheric flux base rates on atmospheric CO<sub>2</sub> concentrations as computed by the coupled STILT-BioFlux models, aggregated by surface type ( $m$  columns) and time step ( $n$  rows). To create the atmospheric observation vector,  $\mathbf{z}$ , modeled results for advected background concentration and the influence of fossil fuel emissions on atmospheric CO<sub>2</sub> are presubtracted from the measurements of absolute CO<sub>2</sub> concentration. So both  $\mathbf{z}$  and the product  $\mathbf{H}\mathbf{s}$  represent the change in atmospheric CO<sub>2</sub> concentration induced by biospheric CO<sub>2</sub> fluxes.

[31] The cost function in equation (6) can be minimized using the following solution [Enting, 2005; Tarantola, 1987]:

$$\hat{\mathbf{s}} = \mathbf{s}_p + \mathbf{Q}\mathbf{H}^T(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}_p) \quad \text{and} \quad (7)$$

$$\mathbf{V}_{\hat{\mathbf{s}}} = \mathbf{Q} - \mathbf{Q}\mathbf{H}^T(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\mathbf{Q}, \quad (8)$$

where

- $\hat{\mathbf{s}}$  posterior optimized set of base rates;
- $\mathbf{V}_{\hat{\mathbf{s}}}$  covariance of the posterior uncertainties for the estimated base rates.

## 2.6.2. Definition of the Error Covariance Matrices

[32] The diagonal elements of the prior uncertainty covariance matrix  $\mathbf{Q}$  are filled with values derived from the SCEM optimization runs (see section 2.3.2). In this study, we neglected cross correlations between uncertainties for different surface types and flux components (i.e., GPP, RA, and RH), so only the diagonal elements of  $\mathbf{Q}$  were filled. Variances ranges for the base rates of GPP, RA, and RH were 0.65–1.42, 8.48–9.75, and 1.16–1.58, respectively (see also the normalized spatial distributions in Figure 6).

[33] The diagonal elements of the model-data mismatch matrix,  $\mathbf{R}_i$ , are calculated as the sum of six individual error sources in the form of variances, following the concept suggested by Gerbig *et al.* [2003]. No potential temporal correlations were considered for any of the errors comprised in  $S_{\varepsilon}$ , so all off-diagonal elements in the model-data mismatch matrix  $\mathbf{R}$  remained 0.

$$\mathbf{R}_i = S_{\text{veg}} + S_{\text{part}} + S_{\text{eddy}} + S_{\text{transp}} + S_{\text{aggr}} + S_{\text{ocean}}. \quad (9)$$

$S_{\text{veg}}$  is the vegetation signal uncertainty, combining measurement uncertainties of the observations and uncertainties in advected background concentrations and fossil fuel fluxes (see also section 2.5). Measurement root mean square error (RMSE) was determined on the basis of measurements of a reference gas with known CO<sub>2</sub> concentration that was sampled at regular intervals and found to be 0.11 ppm (MF site) and 0.12 ppm (MP site), respectively. The background concentration uncertainty is set to 2.35 ppm, which is the standard deviation given on the CarbonTracker website (<http://www.esrl.noaa.gov/gmd/ccgg/carbontracker>) for residuals between modeled and measured CO<sub>2</sub> concentrations for shipboard measurements on the Pacific Ocean. We adopted 30% of the total simulated atmospheric CO<sub>2</sub> concentration change attributed to anthropogenic emissions as a conservative estimate of the uncertainty for fossil fuel emissions (range = 0.00–1.24 ppm). No influence of wildfires is considered here because time windows with wildfire emission influences on atmospheric observations were filtered out (see section 2.2). All individual standard deviations listed previously were squared to get the variances before summing up to  $S_{\text{veg}}$ , as was done for those standard deviations listed in the next paragraphs.

[34]  $S_{\text{part}}$  represents the uncertainty introduced by simulating atmospheric transport processes at a given location with a limited number of trajectories based on a stochastic model. We repeated 200 model runs with the same STILT settings as described in section 2.4 for three selected transport situations at each of the two monitoring sites to derive this statistical uncertainty. The average RMSE over all cases was 0.47% of the total modeled CO<sub>2</sub> concentration (sum of biosphere and fossil fuel fluxes and advected background) and was set to 0.5% for the calculation of the model-data mismatch covariance (range = 1.79–2.13 ppm).

[35]  $S_{\text{eddy}}$  describes the error due to unresolved eddies; that is, the variance of CO<sub>2</sub> concentration within the mixing layer profile that is created by turbulence scales not captured by the modeling approach. To accurately describe this uncertainty, upper air information is required, e.g., provided by aircraft campaigns. Because no such information is available for the spatial domain and time frame of this study, we adopted the results of Gerbig *et al.* [2003] and set  $S_{\text{eddy}}$  to a fixed value of 2 ppm. This value is the upper limit of their indicated error range and should be regarded as a conservative estimate for our study because the quality filtering of the CO<sub>2</sub> observation already excludes weak mixing conditions.

[36] The uncertainty introduced into the inverse modeling approach by uncertainties in the transport simulations,  $S_{\text{transp}}$ , is computed as the sum of two components: The first represents the uncertainties in the transport fields and can be calculated by an approach described by Lin and Gerbig

**Table 3.** Model Setup for the Sensitivity Tests Presented in Sections 3.1 to 3.4.

Section	Test	Time Step	Spatial Resolution	Surface Types
3.1	Base case scenario	1 h	<1 km	120
3.2	Temporal averaging	Variable, four classes	<1 km	120
3.3	Spatial averaging	4 h	Variable, 14 classes	120
3.4	Model domain setup	4 h	<1 km	Variable, five scenarios

[2005]. We adopted some of their settings (correlation time scale = 240 min, horizontal correlation scale = 120 km, and vertical correlation scale = 900 m) and derived a wind speed standard deviation of  $2.5 \text{ m s}^{-1}$  by comparing WRF wind fields with data from two Oregonian radiosonde stations (Salem and Medford). This approach yielded an average transport field uncertainty of 2.4 ppm. The second component of  $S_{\text{transp}}$  represents the influence of boundary layer height uncertainties on the computed fluxes. Because the available radiosonde data have too coarse a vertical resolution to serve as a reference data source and no further data source for upper air information are available in the model domain, we multiply the vegetation signal by a fixed factor to obtain the influence of the boundary layer height uncertainty on the modeled CO<sub>2</sub> concentration signal. Following *Gerbig et al.* [2003, 2008], we set that factor to 0.3, which is a rough guideline translated into an average error of 3.5 ppm. Overall, the uncertainty in our high-resolution WRF fields should be reduced compared with the 35 km resolution ECMWF fields that *Gerbig et al.* [2008] based their findings on. So although parts of this accuracy gain might be offset because of the complex terrain in the Oregon model domain, the chosen factor of 0.3 is assumed to provide conservative estimates of boundary layer height uncertainty.

[37] Using the “classic” definition of the aggregation error,  $S_{\text{aggr}}$ , it would be negligible for our modeling setup because the biosphere fluxes are calculated on a 1 km resolution grid and even include subgrid-scale variability (see also section 2.1). However, for the presented inverse modeling setup, it seems more appropriate to interpret the aggregation error as the uncertainty introduced by flux base rate variability within a given surface type. The flux base rates to be optimized reflect the sensitivity of the biosphere CO<sub>2</sub> fluxes to external drivers, with each surface type consisting of only one plant functional type, affected by a single disturbance regime, and situated in an ecoregion with relatively uniform ecological characteristics. Because this definition of surface types creates largely homogeneous units, and the major part of the remaining internal heterogeneity is captured by spatially variable definitions of stand age, fPAR, and climate, we assume the aggregation error to be of minor importance here. The error introduced by neglecting ocean fluxes,  $S_{\text{ocean}}$ , is almost negligible as well because particle trajectories can only cover a short fetch over the ocean until they reach the western boundary of the outer WRF meteorology domain (Figure 2), where the air parcels are initialized with starting concentrations from Carbon-Tracker that include ocean fluxes. Both  $S_{\text{aggr}}$  and  $S_{\text{ocean}}$  are therefore set to a fixed value of 0.1 ppm.

### 3. Results and Discussion

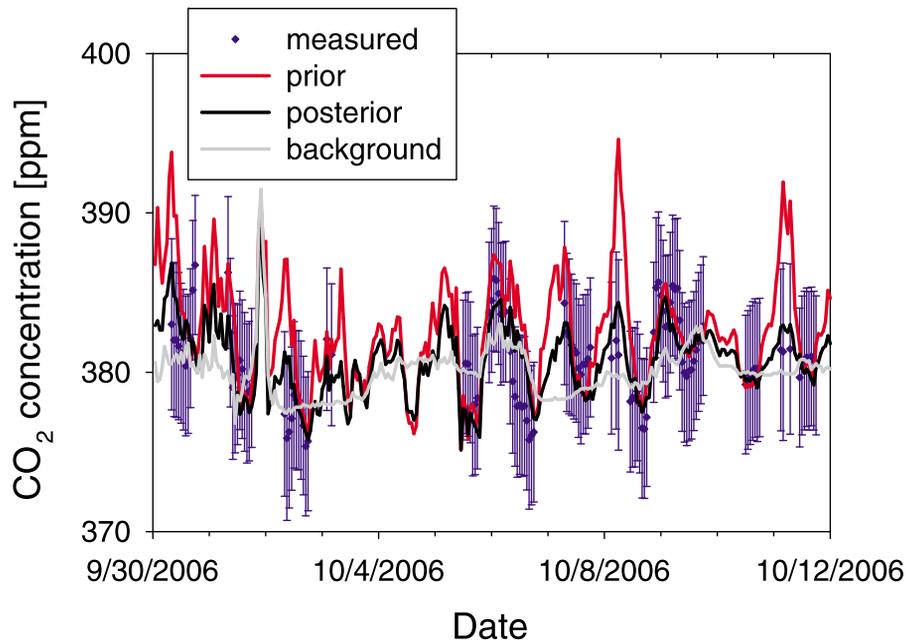
[38] The inverse modeling approach described in detail in section 2 was first applied on a “base case” scenario (time

step = 1 h, spatial resolution <1 km, number of surface types = 120) to demonstrate the overall performance of the modeling framework in the Oregon domain. Results include surface CO<sub>2</sub> fluxes, but because the currently available database does not cover a full annual cycle, these findings are not assumed to be representative beyond the observation period. We therefore restrict the discussion on qualitative aspects and overall plausibility. Using these results as a reference, we further analyzed the influence of temporal averaging (section 3.2) and spatial resolution of the regular surface grid (section 3.3) on the model performance. Finally, in section 3.4, we explored the influence of the definition of ecoregions, land cover classes, disturbance regime, and stand age on the quality of the model output. The model setup for each sensitivity test is summarized in Table 3.

#### 3.1. Measured Versus Modeled CO<sub>2</sub> Concentrations for a Base Case Scenario

[39] The base case scenario presented in this section uses the highest spatial (1 km grid size, including up to five subgrid classes per pixel) and temporal (1 h time steps) resolution settings available for our model setup and assimilates all available classes for ecoregion, land cover type, and disturbance regime to form the highest number of potential surface types (120; see also section 2.1 and Table 1). Figure 3 gives an example of hourly measured versus modeled CO<sub>2</sub> concentrations time series for a 12 day period at the MP site. The given time window was chosen because it includes relatively high vegetation flux activity, a high number of observations passing the quality filter, and significant shifts between prior and posterior model results. The correlation between measured and modeled results is of comparable quality for the times and sites not shown. Vertical blue lines give the total data uncertainty, the square root of the model-data mismatch  $S_{\epsilon}$  (see also section 2.6.2), for each time step, which summarized elements such as transport error, background uncertainty, and so on. Both prior and posterior modeled time series closely follow the measured daily and synoptic trends in most of the periods. Differences between prior and posterior model versions are most obvious during the night, where differences between simulated CO<sub>2</sub> concentrations can exceed 10 ppm, whereas afternoon minima are relatively close in most cases, with absolute differences mostly less than 1 ppm. For the optimized fluxes, nighttime maxima were considerably lower compared with the prior version, with a significantly improved correspondence to those nighttime measurements that passed the quality screening.

[40] To demonstrate the performance of the modeling framework at longer time scales, mean afternoon (2–5 P.M.) CO<sub>2</sub> concentrations are plotted in daily time steps in Figure 4. The overall positive trend in the measurement data due to the transition in atmospheric CO<sub>2</sub> from summertime minimum to wintertime maximum is evident in the simulated



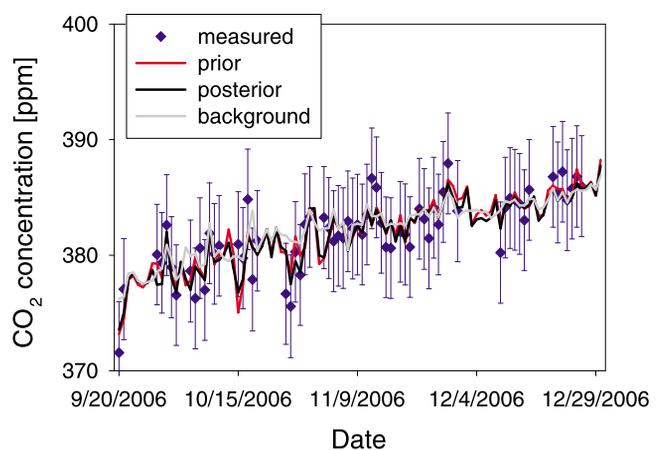
**Figure 3.** Hourly time series of measured versus modeled CO<sub>2</sub> concentrations for a 12 day period at the Metolius mature pine site. Blue diamonds give the time steps that passed the quality filtering and were used as input for the optimization. Vertical blue lines give the model-data mismatch uncertainty, the square root of  $S_{\epsilon}$  (see also section 2.6.2). Solid lines give modeled CO<sub>2</sub> concentration (red) before and (black) after optimization and (gray) background CO<sub>2</sub> concentration.

background concentration signal. Patterns in the deviations of actual CO<sub>2</sub> concentration from that background signal, which may be caused by either regional-scale transport processes or spatiotemporal variations in surface fluxes, are generally followed closely by the model. As already shown in Figure 3, prior and posterior model versions follow similar trends around the peak of daytime concentration drawdowns, with the differences in afternoon averaged CO<sub>2</sub> concentrations rarely exceeding 1 ppm (Figure 4). However, we emphasize that such small differences in the atmospheric concentrations may be associated with changes in the surface flux fields that can accumulate to significantly different flux budgets when integrated during a year.

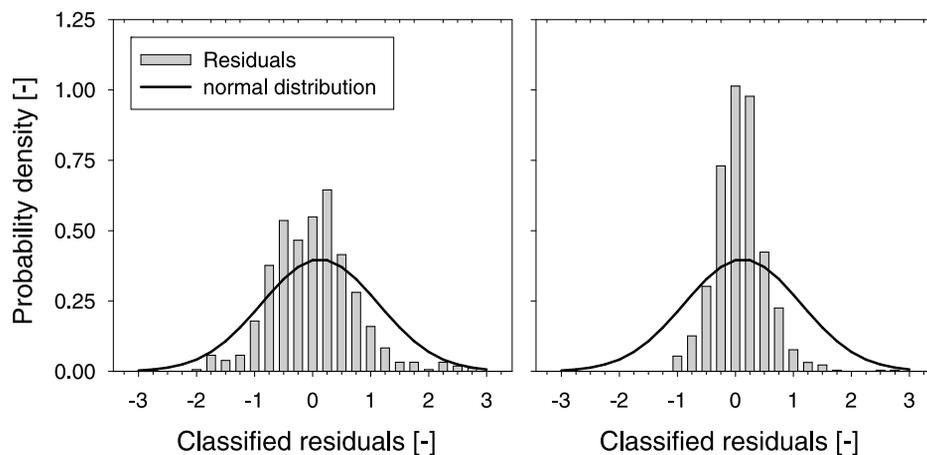
[41] Figure 5 shows the normalized distribution of the residuals between observed CO<sub>2</sub> concentrations and model results on the basis of a posteriori parameters. All values were normalized with the model-data mismatch uncertainties as described in section 2.6.2. For both sites, the distributions are centered at 0, indicating no overall bias in how the inverse modeling framework reproduces the observational data set. Compared with a standard normal distribution (i.e., 0 mean and variance of 1) the spread in the normalized residuals is lower at both sites. The mean squared normalized residual values are 0.64 for the MF site and 0.21 for the MP site, indicating that the model-data mismatch used in the inversion is conservative as intended. Particularly for the MP site, results show that the model-data mismatch variance prescribed in the inversion might be reduced to further improve the use of information contained in the atmospheric observations. For future studies, we are anticipating the availability of new Sonic Detection and Ranging data sets on boundary layer structure to better constrain the boundary layer height uncertainty and of detailed information on

uncertainty in advected boundary conditions from Carbon-Tracker that will be customized for our model domain.

[42] For the base case scenario discussed, Figure 6 displays the spatial distribution of prior and posterior base rates, the relative change of base rates after optimization, the spatial distribution of normalized prior uncertainties, and the uncertainty reduction for the parameters. In all panels, only the dominating surface type for each  $1 \times 1$  km pixel was considered. Note that base rates are sensitive to the spatial



**Figure 4.** Afternoon averages (2–5 P.M.) of measured versus modeled CO<sub>2</sub> concentrations in daily time step for the MP site. Total model-data-mismatch uncertainty for the observations is indicated by vertical blue bars. Modeled results are given (red) before (prior fluxes) and (black) after (posterior fluxes) optimization. The advected background is included as a gray line.



**Figure 5.** Empirical distribution of the residuals [(left) Mature Fir site and (right) Mature Pine site] between observed CO<sub>2</sub> concentrations and predicted observations on the basis of the a posteriori estimates, normalized by the model-data mismatch uncertainty. 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).

patterns of fPAR, hence the relatively high RA base rates in Eastern Oregon, which compensate for the low fPAR there. Another diagnostic modeling approach to simulate biosphere CO<sub>2</sub> fluxes [Turner *et al.*, 2006] used a Beer's law transformation of fPAR to biomass. This does not guarantee more accurate representation of biosphere processes, but it resulted in a more stable RA base rate across vegetation types and ecoregions. However, field observations indicate that base rates change markedly across ecoregions and vegetation types [Campbell and Law, 2005]. At this time, the linear transformation from fPAR to biomass adopted in the BioFlux model has been chosen to stabilize the model against uncertainty in fPAR data.

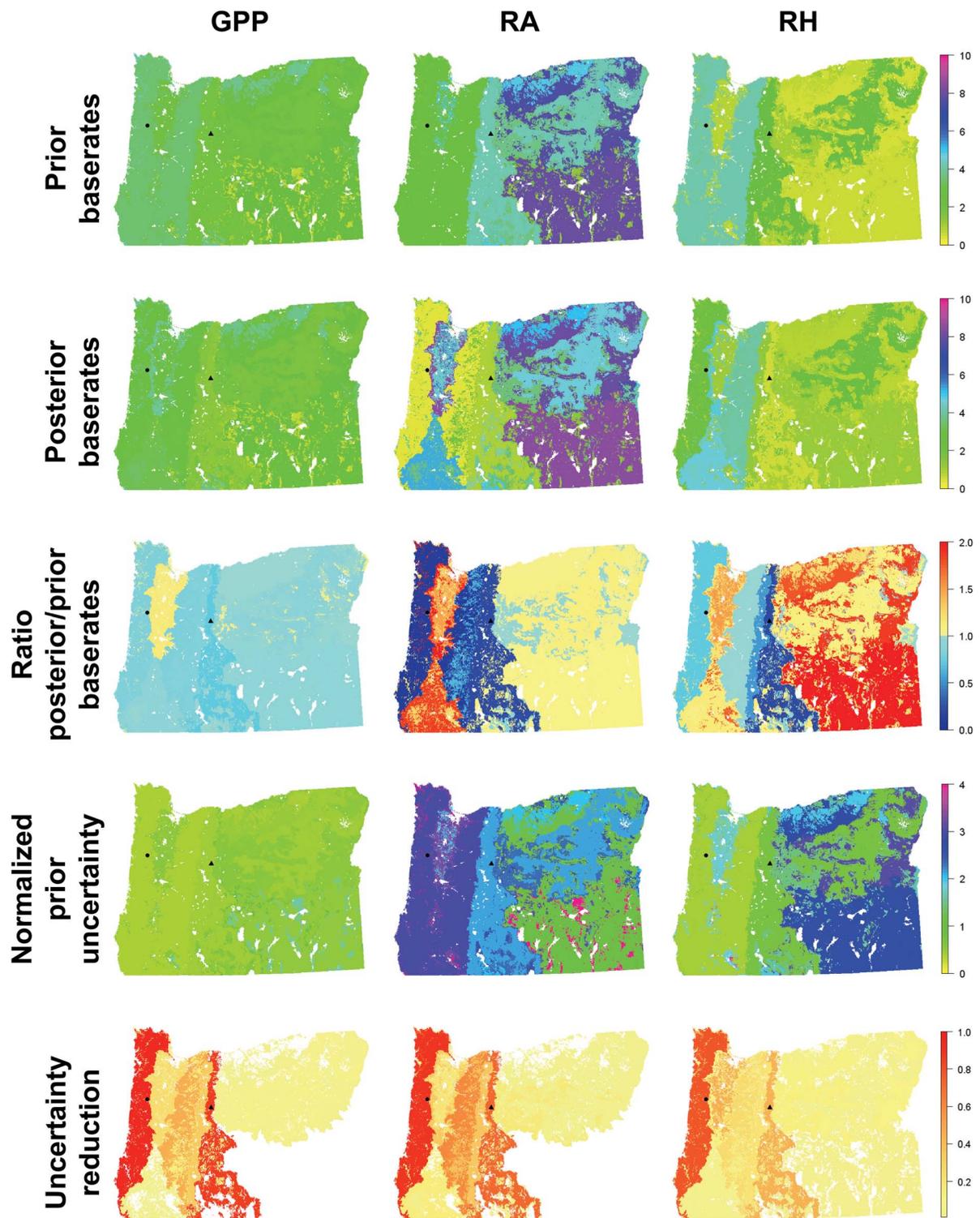
[43] The shifts in base rate values after optimization are expressed by the ratio of posterior to prior base rates, with warm colors (yellow to red) indicating that base rates have been increased by the optimization process and cold colors (blue to cyan) indicating reduced base rates. GPP base rates only experienced minor changes through the optimization, with slightly reduced base rates for most of the state. Changes in respiration base rates were more pronounced, particularly for the RH, indicating biases in prior knowledge on the spatial distribution of respiration. However, the atmospheric data available for inverse modeling do not cover a full seasonal cycle yet, and the shifts observed in base rates in this study may be influenced by the focus on fall and wintertime. In particular, the significant increase in RA in the KM and WV ecoregions seems implausible and needs to be reevaluated with more data in future studies.

[44] The uncertainty reduction is computed as 1 minus the ratio of posterior uncertainty (see section 2.6.1) to prior uncertainty (section 2.3.2). Because uniform base rates are assigned for each surface type, the spatial patterns in the uncertainty reduction shown in Figure 6 are decoupled from the shape of the concentration footprints. If each grid cell would be optimized separately, the uncertainty reduction would peak at the receptor locations of the two observations sites used for this study and gradually decline with increasing distance. In this study, the high amount of information that is available for areas close to the tower

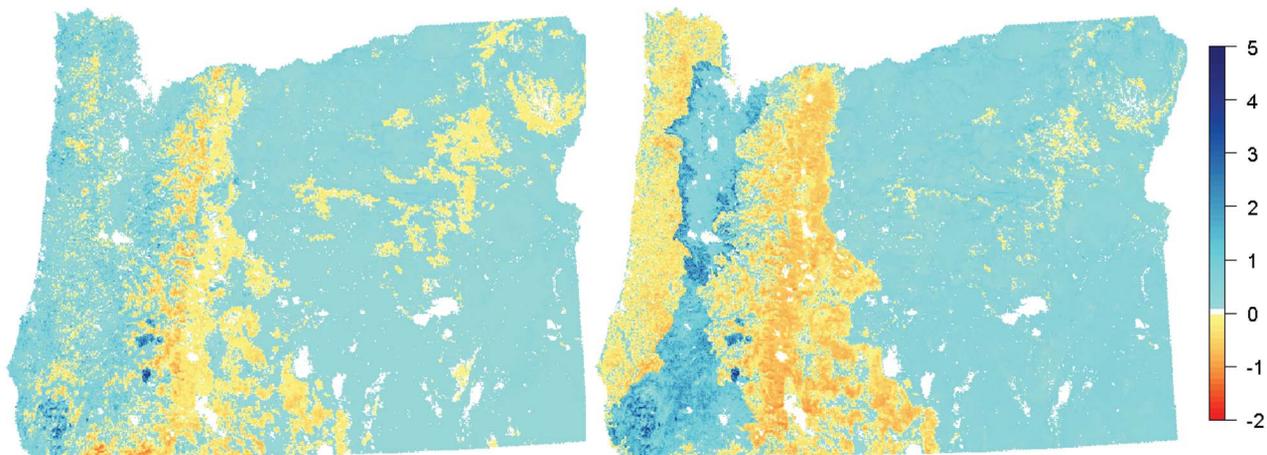
positions is projected on the entire surface type. Accordingly, the highest uncertainty reduction was found for the surface types “Coast Range/evergreen needleleaf forest/cut” that contains the MF site and “East Cascades/evergreen needleleaf forest/fire” that surrounds the MP site. Both surface types appear as dark red bands in the uncertainty reduction maps. Besides these two surface types, significant uncertainty reductions for all three flux components were only observed in the western ecoregions (WV and WC), both of which are covered by a significant portion of the observation footprint. Overall, the spatial maps of uncertainty reduction indicate that the available information is focused on the surface types surrounding the towers, indicating that more sites and longer observation periods are needed to obtain significant uncertainty reductions in representative parameter estimates for the entire state.

[45] Figure 7 presents the combined effect of the base rate changes shown in Figure 6 on the modeled NEE for the state of Oregon for the examined period. Note that the given daily averaged values are based on results from only 4 months of simulations (September–December 2006) and thus are not assumed to be representative beyond the observation period. The flux maps indicate considerable regional shifts between prior (Figure 7, left) and posterior (Figure 7, right) fluxes, with higher CO<sub>2</sub> uptake in the forested regions (particularly CR, WC, and EC) after optimization and stronger sources in agricultural regions (WV). The posterior flux gradient between the KM ecoregion and the adjacent forested ecoregions seems to be unrealistic, so results for this part of the domain have to be considered preliminary. We expect such gradients to smooth out with the addition of new observation sites, and more site years from the two sites used for this study. The range of averaged values is comparable for both prior and posterior fluxes, with high absolute values (>4 g C m<sup>-2</sup> d<sup>-1</sup>) found only in small areas in the WC and KM ecoregions that have recently been burned and thus are dominated by increased decomposition fluxes from dead biomass.

[46] Because many of the nighttime measurement data are excluded through the quality filtering because of inadequate



**Figure 6.** Maps of prior base rates, posterior base rates, the change in posterior base rates as expressed through the ratio of prior to posterior values, the prior uncertainty estimates normalized by the prior flux base rates, and uncertainty reduction after optimization for the three fluxes of (left) GPP, (middle) RA, and (right) RH. Both the ratio of posterior to prior base rates and the normalized prior uncertainties were truncated, with maximum values set to 2 and 4, respectively, to improve visualization of spatial patterns. Tower positions are given by the black dot (MF site) and black triangle (MP site).



**Figure 7.** Daily averaged net ecosystem exchange (NEE) for the period September to December 2006. (left) Prior fluxes. (right) Posterior fluxes after optimization. All values are given in grams of carbon per squared meter per day ( $\text{g C m}^{-2} \text{d}^{-1}$ ), with positive values indicating a net source to the atmosphere and negative values a net sink.

vertical mixing, the capability of the modeling framework to simulate complete daily cycles cannot be thoroughly evaluated on the basis of the currently available data. Vertical mixing processes, which also determine the height of the mixing layer, are a significant source of error in atmospheric inverse modeling [Gerbig *et al.*, 2008], particularly during weak turbulence situations at night. However, Figure 3 demonstrates the strength of the quality filtering to identify error-prone time frames and exclude them from the optimization process. Moreover, because of the output of convective mass fluxes from WRF, the influence of vertical mixing errors during daytime could be reduced compared with the use of less sophisticated meteorological drivers such as Eta Data Assimilation System.

### 3.2. Sensitivity Test on Temporal Averaging

#### 3.2.1. Test Setup and Rationale

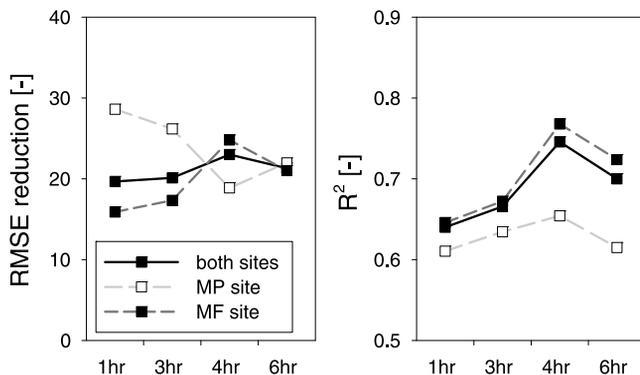
[47] The choice of the temporal averaging scale for CO<sub>2</sub> concentrations affects the performance of the inverse modeling framework in several respects. On one hand, the diurnal cycle of atmospheric CO<sub>2</sub> concentrations can be described most accurately with the highest temporal resolution, and generally speaking, more input time steps mean more available information for the Bayesian optimization approach to improve the flux base rates. On the other hand, high temporal resolution may include more scatter in the data that cannot be resolved by both the terrestrial CO<sub>2</sub> flux model and the transport model, reducing accuracy or even causing artifacts in the output. The ideal temporal averaging interval will smooth out the scatter that is not resolvable by the model while keeping the time steps as short as possible for a minimum loss of information content.

[48] To find the optimum temporal averaging scale, we aggregated hourly measurements and simulation output into bins of 3, 4, and 6 h before running the Bayesian optimization. Aggregated time steps for which less than half of the original 1 h data were available after quality filtering were discarded. For each aggregation version, the measurement error used in the model-data mismatch matrix  $\mathbf{R}$  was adjusted by the factor  $n^{0.5}$ , with  $n$  being the averaging

interval (in h). The other components of  $\mathbf{R}$  remained constant, adding to the conservative nature of our model-data mismatch definitions. For all simulations, we compute the RMSE between measured and modeled CO<sub>2</sub> concentrations before and after optimization and calculate the relative reduction in RMSE obtained through the optimization (Figure 8). Besides this relative measure of information gain through the inverse optimization, we used the coefficient of determination ( $R^2$ ) between measurements and optimized model results as a second statistical measure to evaluate model performance.

#### 3.2.2. Temporal Averaging Test Results

[49] Figure 8 indicates that the effects of temporal averaging on the quality of the optimized results are small as long as the time step is 6 h or less. For the RMSE reduction, trends at both sites are anticorrelated, with the best result for the MF site and the worst for the MP site both obtained with 4 h of averaged data. The net effect averaged for both sites



**Figure 8.** (left) Reduction (%) in the root mean square error (RMSE) after optimizing the results and (right) the absolute coefficient of determination ( $R^2$ ) for the optimized results, both obtained for five different temporal averaging intervals. Results are given for each of the sites separately [(black dashed line) MF and (gray dashed line) MP] and (black solid line) as a composite of both.

still peaks for 4 h of averaging, however. The coefficient of determination has a maximum for 4 h of averaging for both sites. We conclude that for averaging intervals of 4 h, the accuracy gain due to the removal of scatter that is not resolvable by the model outweighs the information loss due to the reduction of the available data set and oversimplification of the diurnal cycle.

[50] The sensitivity of the inverse modeling approach to the temporal averaging is influenced by the quality of the observation data, i.e., the scatter introduced by measurement uncertainty, and by the effectiveness of the data quality filtering. Also, the capability of the modeling framework to simulate the natural short-term dynamics in CO<sub>2</sub> concentration time series plays a major role. A previous study with a comparable model setup [Matross *et al.*, 2006] demonstrated that coupling STILT transport fields to highly resolved surface flux fields can reproduce observed CO<sub>2</sub> concentration changes in hourly time steps when using tall tower or aircraft data. Our finding that the best optimization results can be obtained for time series smoothed into 4 h bins may be associated with our lower measurement heights that increase the potential effect of transport errors in the near field. Also, for short towers, the measurement uncertainty because of the heterogeneity in source/sink strength in the footprint is higher compared with observations in the higher boundary layer. Results may vary with different setups in the modeling framework, e.g., the spatial resolution of the model domain, and the temporal resolution of surface flux fields or advected background signal.

### 3.3. Sensitivity Tests on Spatial Averaging

#### 3.3.1. Test Setup and Rationale

[51] A common strategy in inverse modeling approaches is to split the domain into independent grid cells and optimize the fluxes or the flux base rates (i.e., sensitivities to external drivers such as radiation or temperature) for each cell separately. For this setup, the choice of the size of the grid cells is critical because, with a high-grid resolution, more parameters need to be optimized, and the problem may become underconstrained, whereas with a low-grid resolution, additional uncertainty is introduced when aggregating nonlinear processes or averaging out subgrid-scale variability. Because in the modeling framework described herein flux base rates are optimized for specific surface types, the number of parameters to be optimized does not depend on the spatial resolution of the model domain but depend only on the number of surface types defined (see also section 3.4). Therefore, in our study, the spatial resolution setup influences only the level of detail in the surface domain description through the representation of small-scale variability in surface types. The sensitivity tests described in the next paragraphs were conducted to investigate how a highly detailed surface domain description influences the performance of the modeling framework and whether the correlation between the observed and modeled CO<sub>2</sub> concentrations can be improved.

[52] Using the base case surface map (see section 2.1) as a reference, we aggregated the surface domain into 14 different map versions, with grid resolutions ranging from  $1 \times 1$  km to  $40 \times 40$  km. For these 14 maps, each grid cell contained only a single majority surface type, and all subgrid-scale information was discarded. Aggregating into

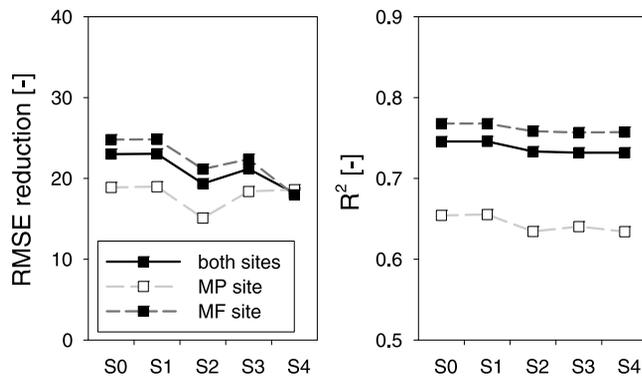
coarser grids did not significantly change proportions between the 10 ecoregions in the Oregon domain. However, for the coarser maps (resolution  $>20$  km), the spatial variability between map versions increased, particularly around the Cascades where the position of ecoregions shifted significantly (see also next paragraphs). For land cover, aggregation generally shifted the proportions from the minor classes to the dominating ones, increasing evergreen needleleaf forest and shrubland while reducing all other land cover classes. The same holds true for the disturbance types; where in each ecoregion, the dominating disturbance type was strengthened with coarser grid size, whereas the minor class was reduced in coverage area.

[53] The choice of the surface map resolution did not have an effect on the computation of the STILT footprints (see section 2.4), which were projected on the same  $0.01^\circ \times 0.01^\circ$  grid for all cases, including the dynamic adaptation of grid cell size with increasing distance from the receptor location. For the larger grid cells in the far field of the footprint, surface fluxes were averaged over larger areas for all setups (a grid cell of  $0.25^\circ \times 0.25^\circ$  covers an area of approximately  $20 \times 30$  km). Consequently, the resolution of the surface maps in the far field is not expected to have a significant effect on the inverse modeling output, as long as the proportions of surface types are not significantly biased, which is the case in our model domain. However, smoothing out the fine-scale variability in surface maps in the near field close to the receptor may have a large effect on the model's performance, particularly if shifts occur between land cover types with significant differences in flux signals (e.g., grassland to evergreen needleleaf forest).

#### 3.3.2. Spatial Averaging Test Results

[54] Concerning the quality of the modeled CO<sub>2</sub> time series, using the same statistical measures as described in section 3.2, we found no significant trend in the influence of spatial averaging on the model output. For both the RMSE reduction and the  $R^2$  results, trends were only subtle and of opposing sign between the two sites for model runs on the basis of the higher resolution maps ( $<16$  km). For the coarser maps, the scatter in the results increased with the grid resolution. Also, good fits between observed and modeled CO<sub>2</sub> concentration time series often could only be obtained because of the assignment of increasingly unrealistic flux base rates; that is, net fluxes tended toward very high or very low values, with extreme and implausible gradients between surface types on local to subregional scales.

[55] The limited effect of horizontal resolution in the model setup can be explained in part by the rather sparse observational network available for this study; that is, the high amount of information provided by a detailed surface map would be more effective with a larger number of sites. In addition, the near field of the footprint area is dominated by forests at both sites, so aggregation did not cause significant shifts in the surface types. In the far field, STILT eliminates fine-scale variability using larger grid sizes in the footprint with increasing distance from the receptor. Because the potential information from spatial patterns in subgrid-scale heterogeneity is supposed to be negligible over such transport distances, while larger grid cells reduce the effect of systematic bias in the transport patterns, this dynamic adjustment helps to strengthen the model. With



**Figure 9.** As in Figure 8 but for different levels of complexity in surface flux model setup. Classes S0 through S4 stand for the five scenarios tested (0, base Case; 1, OneDis; 2, medAge; 3, 2zones; 4, noEco). See text for definition of the scenarios.

changes in surface-type composition negligible (<2%) for aggregated grids with horizontal resolutions <8 km, aggregation to grid sizes of 4–6 km would provide a good compromise between obtaining an accurate description of the surface characteristics and computational efficiency. These shifts in coverage proportions in the maps of ecoregion/land cover/disturbance are dependent on the average length scale of variability in the model domain and thus vary from region to region.

### 3.4. Influence of Flux Model Domain Setup

#### 3.4.1. Test Setup and Rationale

[56] The computation of biosphere CO<sub>2</sub> fluxes uses highly resolved spatial information on ecoregions, land cover types, and disturbance regimes to differentiate the model domain into surface types. In addition, stand age information is provided for forested ecosystems. The base case scenario described in section 2.1 (results shown in section 3.1) used all available remote sensing information, enabling definition of 120 potential surface types (number of parameters to be optimized,  $m = 360$ ) in the Oregon domain (see also Table 1). Here, we conduct a sensitivity study to investigate how the setup of the surface domain influences the model output, and what level of complexity is required to arrive at plausible results for flux base rates.

[57] For the sensitivity test, we gradually simplified the definition of surface types by discarding or simplifying parts of the remote sensing information, regarding the base case scenario (section 3.1) as scenario 0. In scenario 1 (OneDis), we neglected the differentiation of disturbance types into “clear-cut” and “fire,” reducing the number of potential surface types to 60 ( $m = 180$ ). Disturbance is still considered here through the definition of forest stand age, but the differing recovery characteristics of burned and cut areas after the disturbance event are lost. For scenario 2 (medAge), all forests are assigned a uniform stand age of 70 years, which is the median age for Oregon forests derived from federal inventory data [Hudiburg *et al.*, 2009]. Using these maps, the disturbance effect is completely neglected ( $m = 180$ ). Scenario 3 (2zones) makes use of the detailed age information again but simplifies the 10 original ecoregions to only 2, reducing the number of potential surface types to

12 ( $m = 36$ ). The domain is separated into a humid western part (formerly ecoregions CR, WV, WC, and KM) and a semiarid eastern part (formerly EC, BM, CP, NB, SR, and CC) in the simplest attempt to capture the climatic gradient from the Pacific coast to the high desert in Oregon. For scenario 4 (noEco), the ecoregion information is dropped completely, leaving only the land cover classes to form six potential surface types ( $m = 18$ ). In all cases, no adjustments were made to the model-data mismatch matrix  $\mathbf{R}$  to facilitate scenario comparisons. This neglects a potential increase in aggregation error with decreasing numbers of surface types, but because we defined very conservative values for  $\mathbf{R}$  in the first place, we do not expect this to have affected the model results.

#### 3.4.2. Model Domain Setup Results

[58] We again used the RMSE reduction and the  $R^2$  between the measured and optimized CO<sub>2</sub> concentrations as measures of model performance (Figure 9). Results indicate that the consideration of disturbance regimes has negligible influence on the quality of the optimized CO<sub>2</sub> concentration time series because both statistical measures were stable when simplifying the base case to only one disturbance regime (oneDis). Further simplification to scenario 2 with a uniform stand age (medAge), however, has a significant effect on model performance, particularly on RMSE reduction at the MP site. This finding may reflect the highly variable age structure in the ecoregions dominating the MP observations (EC and WC), which are characterized by a fine-scale mosaic of disturbance with ages ranging from 0 to 400+ years. Additional tests using median ages of 30 and 120 years, respectively (results not shown), produced similar output as the medAge scenario, showing that the approach is even insensitive to the absolute value of the assigned stand age. However, for the latter comparison, a closer look at the optimized base rates revealed that changes in the age functions influencing GPP and RH (Appendix B) when using a different median age were compensated by shifts in the base rates. For example, with a lower median age, which produces higher age scaling factors for GPP, the GPP base rates tended to decrease, and vice versa. This implies that bias in assigned stand ages can be corrected for by offsets in the base rates without loss of accuracy, and because individual base rates are assigned per surface type, the only information lost when assigning uniform stand ages is due to the smoothing out variability within surface types. Concerning the comparison of the base case results versus the oneDis and medAge scenarios, variations in the assigned base rates were only moderate, and none of the model runs tended toward unrealistic surface fluxes.

[59] Reducing the original 10 ecoregions to only 2 climate zones (2zones scenario) reduced both RMSE and  $R^2$  only slightly compared with the OneDis scenario, whereas results were better than those found for the noAge simulations. However, a significant loss in output accuracy was found for the noEco scenario that completely neglects ecoregions. Assigning just two climate zones, humid and semiarid, still roughly captures the major differences in plant functional types that reflect the vastly different water availability between Western and Eastern Oregon (section 2.1). Because the transition between both regions follows a gradient extending East from the Cascade crest, even the simplification into just two zones has an effect on the flux simulations

in the Cascades area, as reflected by the moderate decline in the quality of the optimized results for the MP site. Neglecting ecoregions altogether forces the model to find, e.g., a set of base rates averaged for Western fir forests and Eastern pine stands, which cannot be justified from the ecological standpoint.

[60] The results presented in Figure 9 emphasize the major role of ecoregions when defining the model domain surface through surface types. Simplifying the domain setup from the base case down to scenario 3 (2zones), which implies combining more and more biomes with potentially different flux characteristics, only had a moderate effect on the model's performance. At the same time, this simplification reduced the number of surface types by an order of magnitude, from 120 in the base case to just 12 for the 2zones scenario, which limits the number of parameters to be optimized,  $m$ , from 360 to 36 (three base rates per surface type). Further simplifying the setup from scenario 3 (2zones) to scenario 4 (noEco) significantly reduces the correlation between simulated and observed CO<sub>2</sub> concentrations compared with the base case. This simplification has a serious ecological effect, i.e., the elimination of the climate gradient between humid Western Oregon and semiarid Eastern Oregon in the model, while the reduction of degrees of freedom is only small (from 36 to 18) compared with the changes that came with the previous simplification steps. We conclude that the reduction in model performance that comes with scenario 4 can be attributed to the oversimplification in the description of the model domain, or, in other words, the loss of ecoregions, and is not a consequence of statistical properties such as the lower degrees of freedom in the optimization process.

[61] For all model runs, the optimized flux base rates were assumed to be constant over time (section 2.6.2). We acknowledge that better agreement between the observed and optimized CO<sub>2</sub> concentrations can potentially be obtained through the assignment of parameter sets in monthly or seasonal time steps. The complex mechanisms driving biosphere CO<sub>2</sub> fluxes can only be approximated by simple diagnostic models such as BioFlux, so it cannot be ruled out that some processes or feedbacks that vary on seasonal time scales are not accurately simulated. Using base rates that are stable over time, the optimization is forced to find a parameter set that represents the mean of those seasonal trends. Such systematic bias would be improved through the assignment of monthly base rates because it can be assumed that processes not captured by the BioFlux model are more stable on short time frames compared with longer ones. However, variations in monthly base rates may also be caused by model-data mismatch and not be related to a poorly constrained process, so any accuracy gain due to shorter optimization time steps could as well be a statistical artifact based on the extended degrees of freedom.

#### 4. Conclusion

[62] We presented a regional-scale atmospheric modeling approach to constrain terrestrial biosphere CO<sub>2</sub> budgets. The modeling framework is built on a prior domain setup that ingests several remote sensing data sets to differentiate the domain into so-called "surface types." Optimizing flux base

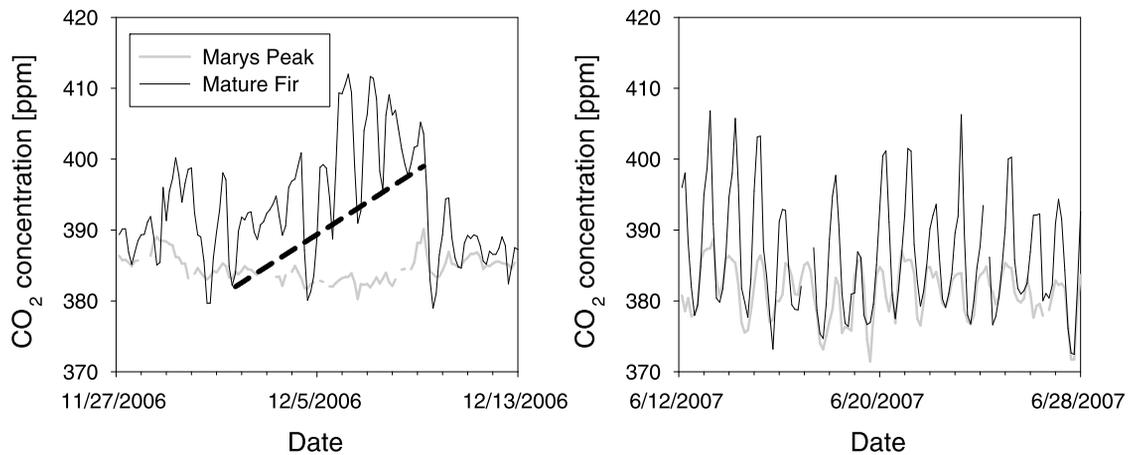
rates for each of these surface types effectively decouples the number of degrees of freedom in the optimization process from the horizontal resolution of the regular biosphere model surface grid and, therefore, permits describing the model surface in high level of detail as required for regional analyses.

[63] Modeled CO<sub>2</sub> concentration time series showed good agreement with observational data at the two examined monitoring sites in the Oregon domain, indicating that both transport processes and spatial and temporal variability in surface CO<sub>2</sub> fluxes are well captured by the modeling framework. The simulated a posteriori surface flux maps showed plausible absolute values and spatial patterns. Sensitivity tests on the temporal resolution revealed that best results were obtained with an aggregation of atmospheric data to time intervals of 4 h. The inverse modeling framework proved to be rather insensitive to the spatial resolution settings as long as horizontal aggregation is 16 km or smaller, although the sensitivity to small-scale variability in the near field could potentially be increased through denser observation networks. For larger grid sizes, horizontal shifts in ecoregion and land cover assignment in the near field could lead to significant additional scatter. For the Oregon domain, loss of information due to averaging of fine-scale heterogeneity was insignificant for grid sizes of 6 km or smaller. Sensitivity tests on the definition of surface types indicated that dividing disturbance into fire versus harvest regimes played a minor role in this model setup as long as the regional-scale characteristics are captured through the ecoregions. In contrast, the definition of ecoregions to describe major ecophysiological differences along climatic gradients is shown to be of paramount importance for the performance of the model. Model performance was improved by the use of highly resolved age information, which resulted in a better fit of simulated CO<sub>2</sub> concentrations to their observed references. The correlation between the observed and modeled time series may have been further improved by allowing temporal variability in flux base rates, but this strategy carries the risk of contaminating flux results with potential model-data mismatch effects.

[64] Our future studies will aim to further reduce parameter uncertainties by including more observation sites and longer data time series with multiple years of observations. An extended observational database will ultimately increase the resolution in the information gain across the state and result in a more effective use of the rich prior information assimilated through the proposed inversion framework. The addition of new sites, ideally located in remote parts of the state to fill in areas currently rarely covered by the footprints of the existing towers, will reduce flux uncertainties linked to the reliance on a few observation sites and their corresponding transport or measurement biases.

#### Appendix A: Disjunct Tall Tower Concept

[65] Vertical gradients of atmospheric CO<sub>2</sub> concentrations within the planetary boundary layer (PBL) can be used to characterize the PBL vertical mixing processes. Small differences between the surface and upper layers indicate well-mixed conditions, with surface layer measurements being representative for the boundary layer. Given the limitations of upper air observations in our modeling domain, we used



**Figure A1.** Comparison of 3 h averaged CO<sub>2</sub> concentrations at the (black) MF and (gray) MPk sites for two 16 day periods in (left) fall 2006 and (right) summer 2007. The black dashed line on the left was included to highlight the ramplike buildup of surface layer CO<sub>2</sub> concentrations during a capping inversion.

CO<sub>2</sub> concentration time series from a mountaintop site on Mary’s Peak in the Oregon Coast Range (MPk, 44.40°N, 123.55°W, 1248 m asl) to serve as the upper measurement height of a “disjunct” tall tower, with the lower levels composed of the observations at the mature fir (MF) and mature pine (MP) sites. The MPk site is instrumented with a monitoring system for CO<sub>2</sub> concentrations nearly identical to the ones installed at the MF and MP sites and has been operating since October 2006. Because of its isolated location as the highest point in the Coast Range, the MPk site can be assumed to represent the upper part of the PBL well. The horizontal distance to the MF site is about 16 km to the north, whereas the MP site is situated about 160 km to the east; accordingly, a bias due to the horizontal advection can be neglected when using the MPk time series as an upper level for the MF disjunct tall tower, whereas it needs to be taken into account for the MP site.

[66] To evaluate the vertical gradient of CO<sub>2</sub> concentrations for the MF and MP disjunct tall towers, we used 3 h averaged data from the period October 2006 to November 2007. For the MF site, afternoon minimum CO<sub>2</sub> concentrations correlated well between the lower levels (MF observations) and the upper level (MPk observation) for most of the days in the observation period, indicating small vertical gradients and thus a well-mixed boundary layer. However, results differed for two general seasons, namely, a “warm” season (March–September) and a “cold” season (October–February). In the entire warm season, afternoon minimum concentrations for both surface layer sites were closely correlated with the upper air observations, with an example shown for a 16 day period in Figure A1 (right). During the cold season, periods could be identified when afternoon minimum concentrations at the lower levels were decoupled from the upper air observations at the MPk site, leading to a ramp-like slow buildup of atmospheric CO<sub>2</sub> over time [example shown in Figure A1 (left), highlighted by black dashed line]. Under these conditions, the observations at the MF site are not representative for the PBL anymore, and simulation with the inverse modeling framework fails.

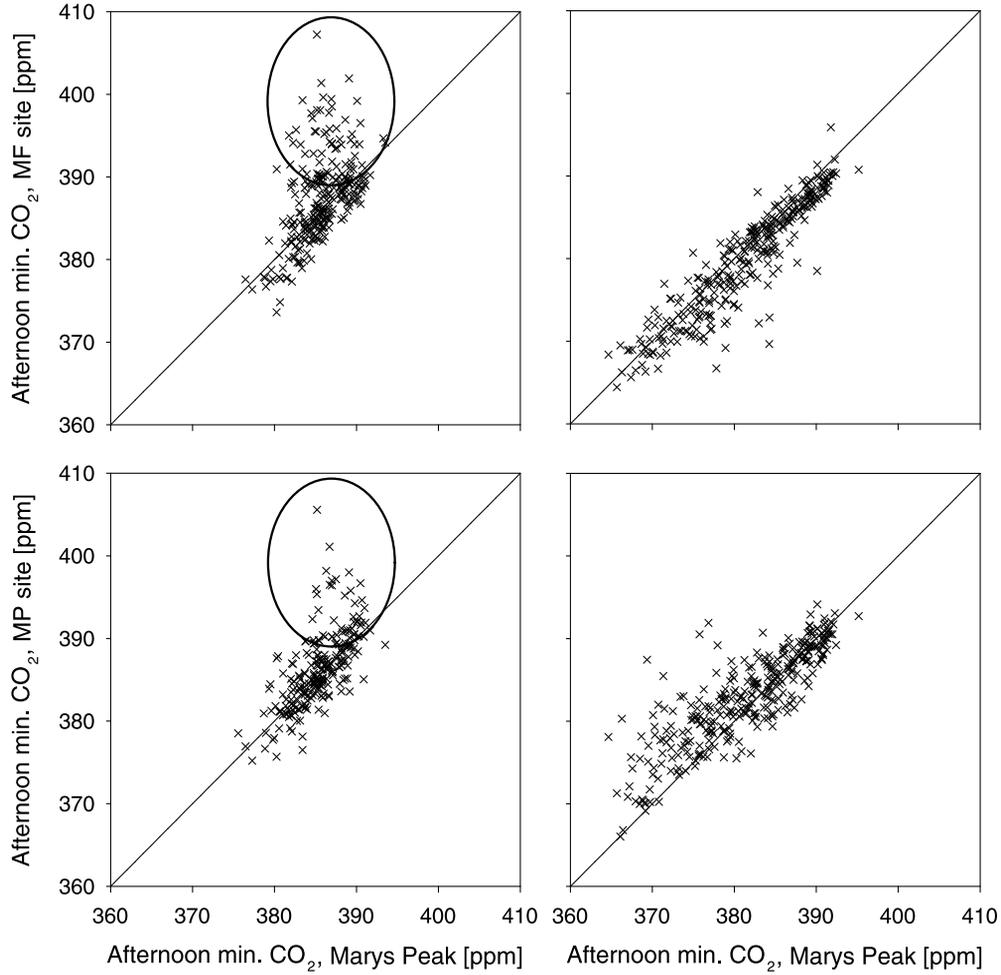
[67] Scatter plots between afternoon minimum CO<sub>2</sub> concentration measured near the surface (MF and MP data) and in the upper boundary layer (MPk data) are shown in Figure A2. This analysis confirms that vertical gradients of CO<sub>2</sub> concentration are small at both sites during the warm season, with results clustering around the 1:1 lines (Figures A2, right). During the cold season, however, part of the data set systematically breaks away from the 1:1 relationship, with clearly higher concentrations in the surface layer (Figures A2, left, encircled areas). Compared with the MF disjunct tall tower, results at the MP site are slightly biased by the horizontal distance between MP and MPk site locations, so the scatter around the 1:1 line in Figure A2 (bottom) is increased. Still, the ramplike patterns of increasing CO<sub>2</sub> concentrations while at the same time increasing the offset from the upper air observations at the MPk site were clearly distinguishable also at this site.

[68] A correlation analysis with additional meteorological observations revealed that the capping inversion events that decoupled the surface layer observations from the upper air observations were characterized by cold temperatures, weak turbulence (as indicated by a low friction velocity), and calm winds from easterly directions. These ancillary information can be used to filter data sets not representative of the PBL in case no upper air observations from MPk are available, e.g., because of a data gap.

## Appendix B: Biosphere Carbon Flux Model Algorithms

[69] The sigmoid ramplike functions to calculate the influence of minimum temperature and daytime VPD are given in equations (B1) and (B2).  $T_{\min}$  and  $VPD_d$  are taken from and derived from the SOGS interpolated surface meteorology, respectively. Inflection points and influence widths are optimized for each surface type (see section 2.3.2). See Table B1 for parameter definitions.

$$T_{sc} = \frac{1}{1 + \exp\left[\frac{-(T_{\min} - T_{\text{inf}})}{T_{\text{wid}}}\right]}} \quad \text{and} \quad (\text{B1})$$



**Figure A2.** Comparison of afternoon minimum CO<sub>2</sub> concentrations between surface layer (MF and MP data) and upper boundary layer (Mary's Peak data). (left) Results [(top) MF and (bottom) MP sites] for the “cold” season (October–February). (right) Results [(top) MF and (bottom) MP sites] for the “warm” season (March–September). The thin black line indicates the 1:1 line, and the black circles mark the data points influenced by capping inversions.

$$\text{VPD}_{\text{sc}} = 1 - \frac{1}{1 + \exp\left[\frac{-(\text{VPD}_{\text{d}} - \text{VPD}_{\text{inf}})}{\text{VPD}_{\text{wid}}}\right]}. \quad (\text{B2})$$

For all forest land cover types, functions were parameterized on the basis of the model output of Biome-BGC [Thornton *et al.*, 2002] to simulate the effect of stand age on GPP (B3) and RH (B4).

$$\text{Age}_{\text{GPP}} = \min[1, \text{gp}_1 + \text{gp}_2 \cdot \exp(\text{gp}_3 \cdot \text{age})] \quad \text{and} \quad (\text{B3})$$

$$\text{Age}_{\text{RH}} = \text{rp}_1 [0.5 + \text{rp}_2 \cdot \exp(\text{rp}_3 \cdot \text{age}) + 0.5(1 - \text{rp}_4^{\text{age}})], \quad (\text{B4})$$

where

age median stand age (years);  
 $\text{gp}_x$ ,  $\text{rp}_x$  scaling parameters (see Table B2).

The exponential functions to assess the influence of soil temperature ( $\text{TS}_{\text{sc}}$ ) and soil water content ( $\text{SW}_{\text{sc}}$ ) on RH

include the weighting factors  $\text{TS}_{\text{wgt}}$  and  $\text{SW}_{\text{wgt}}$ , which could be optimized by surface type. For this study, we used prescribed values (see Table B1). Again, the parameterization of these functions is based on output of the Biome-BGC model (see Turner *et al.* [2007] for details).

$$\text{TS}_{\text{sc}} = \exp(\text{TS}_{\text{wgt}} \cdot \text{TS}) \quad \text{and} \quad (\text{B5})$$

$$\text{SW}_{\text{sc}} = \frac{1 - 0.86 \cdot \exp(\text{SW}_{\text{wgt}} \cdot \text{SWC})}{0.7486}, \quad (\text{B6})$$

where

$\text{TS}_{\text{wgt}}$  weighting factor for soil temperature influence on RH;  
 $\text{SW}_{\text{wgt}}$  weighting factor for soil water content influence on RH.

For input data preparation, daytime and daily average temperatures are approximated as weighted means of daily minimum and maximum temperature [e.g., Thornton *et al.*,

**Table B1.** Overview on Parameters Used in the Terrestrial Biosphere Carbon Flux Model

Parameter	Description	Unit	Source
Age <sub>GPP</sub>	Age scaling function on GPP	-	Turner <i>et al.</i> [2006]
Age <sub>RH</sub>	Age scaling function on RH	-	Turner <i>et al.</i> [2006]
APAR	Absorbed photosynthetically active radiation	MJ m <sup>-2</sup> time step <sup>-1</sup>	Product of fPAR and PAR
CL <sub>sc</sub>	Cloudiness scaling factor	-	Ratio of SW <sub>pot</sub> and SW <sub>act</sub>
CL <sub>wgt</sub>	Cloudiness influence weight	-	Optimized per biome type
ET	Evapotranspiration	mm d <sup>-1</sup>	MODIS product [Mu <i>et al.</i> , 2007]
fPAR	PAR fraction absorbed by the canopy	-	MODIS product
GPP	Gross primary productivity	g C m <sup>-2</sup> time step <sup>-1</sup>	Model output
GPP <sub>base</sub>	Base rate for gross primary productivity	g C MJ <sup>-1</sup>	Optimized per biome type
P	Precipitation	mm d <sup>-1</sup>	Gridded surface meteorology
PAR	Photosynthetically active radiation	MJ m <sup>-2</sup> time step <sup>-1</sup>	Gridded surface meteorology
Q <sub>10</sub>	Base rate for R <sub>m</sub> temperature influence	-	Prescribed value of 2
RA	Autotrophic respiration	g C m <sup>-2</sup> time step <sup>-1</sup>	Model output
R <sub>g</sub>	Growth respiration	g C m <sup>-2</sup> time step <sup>-1</sup>	Model output
R <sub>g, frac</sub>	Fraction of carbon available for growth	-	Prescribed value of 0.25
RH	Heterotrophic respiration	g C m <sup>-2</sup> time step <sup>-1</sup>	Model output
RH <sub>base</sub>	Base rate for heterotrophic respiration	g C m <sup>-2</sup> time step <sup>-1</sup>	Optimized per biome type
R <sub>m</sub>	Maintenance respiration	g C m <sup>-2</sup> time step <sup>-1</sup>	Model output
R <sub>m, base</sub>	Base rate for maintenance respiration	g C m <sup>-2</sup> time step <sup>-1</sup>	Optimized per biome type
Srad <sub>pot</sub>	Potential incoming shortwave radiation	W m <sup>-2</sup>	Fu and Rich [1999]
Srad <sub>act</sub>	Actual incoming shortwave radiation	W m <sup>-2</sup>	Gridded surface meteorology
SWC	Soil water content	-	Simple water balance model
SWHC	Soil water holding capacity	mm	Prescribed
SW <sub>sc</sub>	Soil water content scaling factor	-	Exponential function
T <sub>air</sub>	Mean air temperature	°C	Gridded surface meteorology
T <sub>inf</sub>	T <sub>min</sub> inflection point	°C	Optimized per biome type
T <sub>min</sub>	Daily minimum temperature	°C	Gridded surface meteorology
T <sub>sc</sub>	Minimum temperature scaling factor	-	Sigmoid function
TS	Soil temperature	°C	Derived from gridded meteorology
TS <sub>sc</sub>	Soil temperature scaling factor	-	Exponential function
T <sub>wid</sub>	T <sub>min</sub> influence width	°C	Optimized per biome type
VPD <sub>d</sub>	Daytime VPD	Pa	Gridded surface meteorology
VPD <sub>inf</sub>	VPD inflection point	Pa	Optimized per biome type
VPD <sub>sc</sub>	VPD scaling factor	-	Sigmoid function
VPD <sub>wid</sub>	VPD influence width	Pa	Optimized per biome type
WUE	Water use efficiency	mm g C <sup>-1</sup>	Prescribed [Law <i>et al.</i> , 2002]

1997]. Both temperatures serve as input to convert the 24 h average VPD provided by SOGS to a daytime averaged value that is required for the BioFlux model. The incoming shortwave radiation ( $S_{rad}$ ) is converted to PAR using a fixed PAR/ $S_{rad}$  ratio of 0.45 and the daytime length provided by DayMet. The cloudiness scalar CL<sub>sc</sub> is the ratio of actual incoming PAR derived from SOGS data and potential PAR (PAR<sub>pot</sub>), which is computed using the algorithm of Fu and Rich [1999].

$$CL_{sc} = 1 - \left( \frac{PAR}{PAR_{pot}} \right). \quad (B7)$$

For the computation of RH, soil temperature is derived as the running mean of daily air temperature during the previous 25 days. Soil water is simulated by a simple water mass balance approach at a daily time step, with water loss taken from gridded MODIS evapotranspiration [Mu *et al.*, 2007] and SOGS precipitation data refilling the soil water storage. The soil water content is then derived by normalizing the actual soil water with the soil water holding capacity, here assumed to be 200 mm for the entire domain.

### Appendix C: Computation of Subdaily Radiation and Temperature

[70] Daily PAR values are split into subdaily time steps by assigning weighting factors to each half-hour time step. These weighting factors follow a four-parameter modified

Gaussian distribution to simulate the daily course of PAR and sum up to 1 to give the relative contribution of each half-hour to the total incoming radiation:

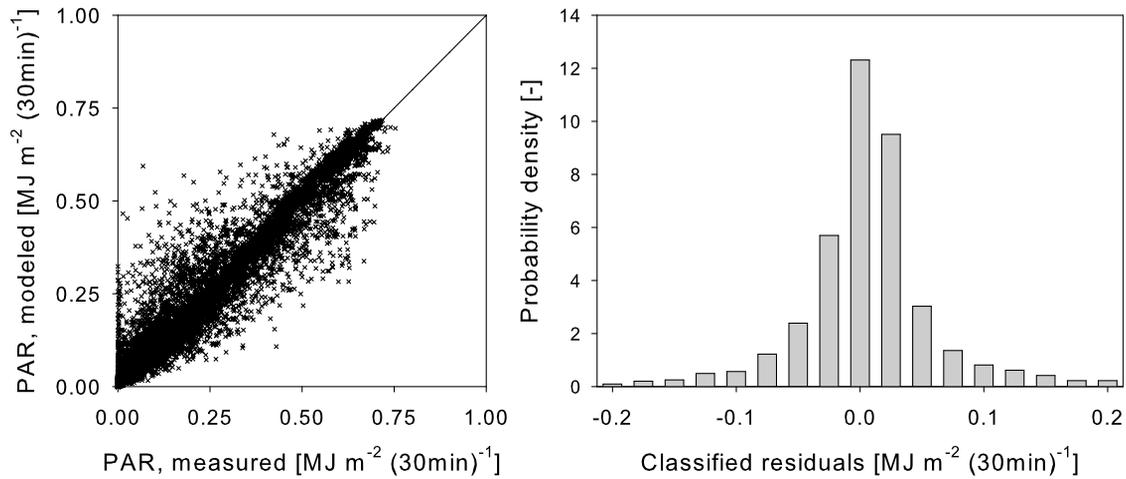
$$PAR_{hr} = PAR \cdot PAR_{wgt} \quad \text{and} \quad (C1)$$

$$PAR_{wgt} = a \cdot \exp \left[ -0.5 \cdot \left( \frac{time - time_{max}}{b} \right)^2 \right], \quad (C2)$$

**Table B2.** Factors for Age Influence Functions on GPP and RH

	Biome Number				
	1	2	3	4	5
gp <sub>1</sub>	0.823	0.852	0.714	0.745	0.650
gp <sub>2</sub>	0.430	0.530	0.360	0.520	0.990
gp <sub>3</sub>	-0.011	-0.013	-0.008	-0.009	-0.440
rp <sub>1</sub>	0.372	0.386	0.318	0.264	0.74
rp <sub>2</sub>	2.190	2.090	2.640	3.300	0.840
rp <sub>3</sub>	-0.154	-0.073	-0.122	-0.064	-0.057
rp <sub>4</sub>	0.982	0.983	0.978	0.990	0.980

Biome settings: 1, cut conifer forest or shrub, dry climate; 2, burned conifer forest or shrub, dry climate; 3, cut conifer forest or shrub, humid climate; 4, burned conifer forest or shrub, humid climate; and 5, deciduous broadleaf forest.



**Figure C1.** (left) Comparison of measured versus modeled PAR using 30 min averaged values for the year 2004 at the Metolius mature pine site as reference data. (black line) The linear trend of the data (slope = 0.92, intercept = 0.01,  $R^2 = 0.94$ ) closely follows the 1:1 line. (right) Empirical distribution of the residuals between measured and predicted PARs, binned into classes of 0.025. Approximately 75% of the residuals have absolute values lower than  $0.05 \text{ MJ m}^{-2}$ .

where

- PAR<sub>hr</sub> hourly PAR ( $\text{MJ m}^{-2} 30 \text{ min}^{-1}$ );
- PAR daily PAR derived from SOGS ( $\text{MJ m}^{-2} \text{ d}^{-1}$ );
- PAR<sub>wgt</sub> relative contribution of half-hour time step to daily PAR (-);
- time time of day (h);
- time<sub>max</sub> time of maximum PAR (h);
- $a$  maximum PAR scaling factor;
- $b$  curve width parameter;
- $c$  curve shape parameter.

One year of half-hourly PAR measurements from MP (see also section 2.3.2) were used to train and test the algorithm. Parameters were calibrated using 70 selected cloudless days, whereas the performance test is based on the full year of data, independent of cloud cover.

[71] Time<sub>max</sub>, the time of day when incoming PAR peaks, is taken as the average of sunrise and sunset times. The maximum PAR scaling factor ( $a$ ) can be approximated as a function of day length ( $R^2 = 0.99$ ):

$$\text{time}_{\max} = \frac{\text{sunrise} + \text{sunset}}{2} \quad \text{and} \quad (\text{C3})$$

$$a = -0.0001052 \cdot \text{DL}^3 + 0.00442 \cdot \text{DL}^2 - 0.06573 \cdot \text{DL} + 0.4038, \quad (\text{C4})$$

where

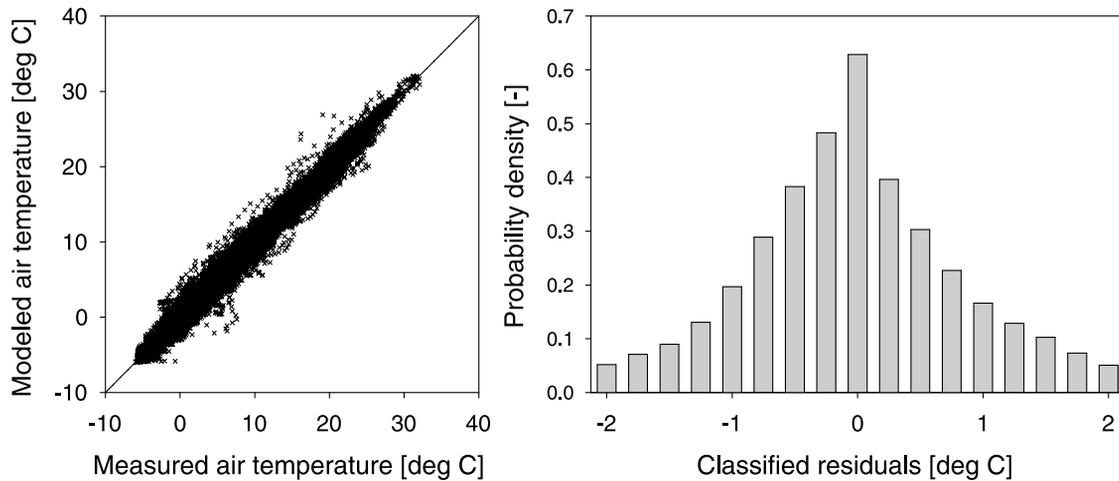
- sunrise time of sunrise (h);
- sunset time of sunset (h);
- DL day length (sunset - sunrise) (h).

Sunrise and sunset times required in both (C3) and (C4) are computed using textbook equations for solar declination angles on the basis of geographical position and date.

Values for the curve shape parameter ( $c$ ) varied in the range of 2.67–3.1, with an average value of 2.97, and had no detectable trend during the year. It is therefore set to a fixed value of 3. The curve width parameter ( $b$ ) was found to be linearly dependent on day length ( $b = 0.273 \times \text{DL}$ ,  $R^2 = 0.97$ ). In a final step, the sum of the half-hourly weighting factors needs to be normalized by the daily sum of weighting factors to ensure their total sum adds up to exactly 1 (usually the sum of half-hourly weights overestimates daily PAR by 2%–5%).

[72] For the set of selected days used to train the model, the correlation between measured and parameterized PAR was very high (average daily  $R^2 = 0.997$ ). Application of the algorithm on the full year of data increased the scatter but still yielded very good agreement between model results and reference (see Figure C1). The increased scatter can mainly be attributed to the fact that the developed algorithms assume a uniform cloud cover during each day; that is, it distributes the incoming PAR taken from the SOGS data set in an ideal modified Gauss curve during the day. This approach will provide poor results when cloud cover changes, e.g., from overcast in the morning to clear skies in the afternoon, when PAR would be overestimated during the first half of the day and underestimated in the second. Although this may increase model uncertainty at certain times, errors are not systematic but cancel over time.

[73] Temperature interpolation to subdaily temporal resolution assumes a “classic” temperature course during each day, with the daytime minimum in the morning and the maximum in the afternoon. Using the same reference data as for the PAR interpolation described above, the specific times for these temperature extremes were approximated on the basis of the parameterized sunrise and sunset times:  $t_{\min} = \text{sunrise}$  ( $R^2 = 0.98$ );  $t_{\max} = \text{sunrise} + (\text{sunset} - \text{sunrise}) \times 0.716$  ( $R^2 = 0.94$ ). Using these times, each day is split into three sections: before the morning minimum, between morning minimum and afternoon maximum, and after the afternoon maximum. Temperatures to interpolate between



**Figure C2.** (left) Comparison of measured versus modeled air temperatures using 30 min averaged values for the year 2004 at the Metolius mature pine site as reference data. (black line) The regression line (slope = 0.99, intercept = 0.02,  $R^2 = 0.98$ ) closely follows the 1:1 line. (right) Empirical distribution of residuals, binned into classes of 0.25, indicates no systematic bias.

are SOGS minimum and maximum temperatures for the given day, the maximum temperature of the previous day, as well as the minimum temperature of the following day. A regular sine function was used for increasing temperatures in equation (C6), whereas a squared sine function was found to describe the decreasing temperatures in equations (C5) and (C7) best:

$$t < t_{\min}(\text{DOY}) : T = T_{\min}(\text{DOY}) + \left\{ 0.5 \cdot \left[ \sin \left[ \frac{1}{2} \pi + \pi \cdot \left( \frac{t - t_{\max}(\text{DOY} - 1) + 24}{t_{\min}(\text{DOY}) - t_{\max}(\text{DOY} - 1) + 24} \right) \right] + 1 \right] \right\}^2 \cdot [T_{\max}(\text{DOY} - 1) - T_{\min}(\text{DOY})], \quad (\text{C5})$$

$$t_{\min}(\text{DOY}) < t < t_{\max}(\text{DOY}) : T = T_{\min}(\text{DOY}) + 0.5 \cdot \left\{ \sin \left[ \frac{3}{2} \pi + \pi \cdot \left( \frac{t - t_{\min}(\text{DOY})}{t_{\max}(\text{DOY}) - t_{\min}(\text{DOY})} \right) \right] + 1 \right\} \cdot [T_{\max}(d) - T_{\min}(d)], \quad \text{and} \quad (\text{C6})$$

$$t > t_{\max}(\text{DOY}) : T = T_{\min}(\text{DOY} + 1) + \left\{ 0.5 \cdot \left[ \sin \left[ \frac{1}{2} \pi + \pi \cdot \left( \frac{t - t_{\max}(\text{DOY})}{t_{\min}(\text{DOY} + 1) - t_{\max}(\text{DOY}) + 24} \right) \right] + 1 \right] \right\}^2 \cdot [T_{\max}(\text{DOY}) - T_{\min}(\text{DOY} + 1)], \quad (\text{C7})$$

where

- $t$  actual time of day (h);
- $t_{\min}$  time of daily minimum temperature (h);
- $t_{\max}$  time of daily maximum temperature (h);
- DOY actual day of year (-);
- $T_{\min}$  minimum temperature (°C);
- $T_{\max}$  maximum temperature (°C).

Applied during the full year of data, the plot comparing measured and parameterized 30 min averaged temperatures demonstrates a close 1:1 relationship (see Figure C2). Deviations between individual values are in the range  $[-8, 8]$ .

Absolute deviations fall below 1°C for approximately 73% of the data, and for only 8% of the data, the deviations exceed 2°C. With an average deviation of  $-0.04^\circ\text{C}$ , no systematic offset is introduced into the data by applying the described algorithm.

[74] **Acknowledgments.** This research was supported by the Office of Science (BER), U.S. Department of Energy (grant DE-FG02-06ER63917), 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 Univ.), John Lin (Univ. of Waterloo, Canada), and Christoph Gerbig (MPI for Biogeochemistry in Jena, Germany) for developing and providing the STILT software package; Marcos Longo (Harvard Univ.), Janusz Eluszkiewicz, and Thomas Nehrkorn (both AER Inc) for mesoscale modeling support; Wouter Peters and Andrew Jacobson (both NOAA ESRL) for support with the CarbonTracker product; Andrew Michaelis and Ramakrishna Nemani (both NASA Ames) for providing customized SOGS data sets; Peter Thornton (ORNL) for customized DayMet data; Qiaozhen Mu (NTSG group, Univ. of Montana) for MODIS ET data sets; Marc Fisher (LBNL) for customized versions of the fossil fuel data sets; Matthias Falk (Univ. of California, Davis) for the provision of the Wind River eddy-covariance data sets; Juriaan Spaaks (Univ. of Amsterdam, Netherlands) for providing the SCEN software package; Dave Ritts (Oregon State Univ.) for remote sensing product support; Christoph Thomas (Oregon State Univ.) for the provision of the Metolius eddy-covariance data set; Kent Davis and Jon Boro (both Oregon State Univ.) for setting up and maintaining the CO<sub>2</sub> concentration measurement program; Britton Stevens (NCAR) and Steven Wofsy for advice in the development of our continuous CO<sub>2</sub> concentration monitoring systems; and Julie Styles for developing the initial version of the BioFlux model.

## References

- Baker, D. F., et al. (2006), Variational data assimilation for atmospheric CO<sub>2</sub>, *Tellus, Ser. B*, 58(5), 359–365.
- Baldocchi, D. (2008), Breathing of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems, *Aust. J. Bot.*, 56(1), 1–26.
- Baldocchi, D., et al. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bull. Am. Meteorol. Soc.*, 82(11), 2415–2434.
- Bousquet, P., et al. (1999), Inverse modeling of annual atmospheric CO<sub>2</sub> sources and sinks. 1. Method and control inversion, *J. Geophys. Res.*, 104(D21), 26161–26178.
- Bousquet, P., et al. (2000), Regional changes in carbon dioxide fluxes of land and oceans since 1980, *Science*, 290(5495), 1342–1346.

- Campbell, J. L., and B. E. Law (2005), Forest soil respiration across three climatically distinct chronosequences in Oregon, *Biogeochemistry*, 73(1), 109–125.
- Campbell, J. L., et al. (2009), Carbon dynamics of a ponderosa pine plantation following a thinning treatment in the northern Sierra Nevada, *For. Ecol. Manage.*, 257(2), 453–463.
- Cohen, W. B., et al. (2002), Characterizing 23 years (1972–95) of stand replacement disturbance in western Oregon forests with Landsat imagery, *Ecosystems*, 5(2), 122–137.
- Cohen, W. B., et al. (2003), An improved strategy for regression of biophysical variables and Landsat ETM+ data, *Remote Sens. Environ.*, 84(4), PII S0034-4257(0002)00173-00176.
- 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, doi:10.1002/joc.1688.
- Deng, A., and D. R. Stauffer (2006), On improving 4-km mesoscale model simulations, *J. Clim. Appl. Meteorol.*, 45(3), 361–381.
- Donato, D. C., et al. (2006), Post-wildfire logging hinders regeneration and increases fire risk, *Science*, 311(5759), 352–352.
- Donato, D. C., et al. (2009), Vegetation response to a short interval between high-severity wildfires in a mixed-evergreen forest, *J. Ecol.*, 97(1), 142–154.
- Duan, Q. Y., et al. (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, 28(4), 1015–1031.
- Duan, Q. Y., et al. (1993), A shuffled complex evolution approach for effective and efficient global minimization, *J. Optimiz. Theory App.*, 76(3), 501–521.
- Enting, I. G. (2005), *Inverse Problems in Atmospheric Constituent Transport*, 412 pp., Cambridge Univ. Press, New York.
- Environmental Protection Agency (EPA) (2008), *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2006*, 473 pp, U.S. Environmental Protection Agency, Washington, D. C.
- Falk, M., et al. (2008), Flux partitioning in an old-growth forest: Seasonal and interannual dynamics, *Tree Physiol.*, 28(4), 509–520.
- Fu, P., and P. M. Rich (1999), *Design and Implementation of the Solar Analyst: An ArcView Extension for Modeling Solar Radiation at Landscape Scales*, edited by Paul M. Rich, ESRI, San Diego, Calif.
- Gash, J. H. C. (1986), A note on estimating the effect of a limited fetch on micrometeorological evaporation measurements, *Boundary Layer Meteorol.*, 35, 409–414.
- Gerbig, C., et al. (2003), Toward constraining regional-scale fluxes of CO<sub>2</sub> 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.
- Gerbig, C., et al. (2006), What can tracer observations in the continental boundary layer tell us about surface-atmosphere fluxes?, *Atmos. Chem. Phys.*, 6, 539–554.
- Gerbig, C., et al. (2008), Vertical mixing in atmospheric tracer transport models: Error characterization and propagation, *Atmos. Chem. Phys.*, 8(3), 591–602.
- Gloor, M., et al. (2001), What is the concentration footprint of a tall tower?, *J. Geophys. Res.*, 106(D16), 17831–17840.
- Gourdji, S. M., et al. (2008), Global monthly averaged CO<sub>2</sub> fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, *J. Geophys. Res.*, 113, D21115, doi:10.1029/2007JD009733.
- Gourdji, S. M., et al. (2009), Regional-scale geostatistical inverse modeling of North American CO<sub>2</sub> fluxes: A synthetic data study, *Atmos. Chem. Phys. Discuss.*, 9, 22407–22458.
- Gurney, K. R. (2008), *The Vulcan Project*, release 1.0, edited, Purdue University.
- Gurney, K. R., et al. (2002), Towards robust regional estimates of CO<sub>2</sub> sources and sinks using atmospheric transport models, *Nature*, 415(6872), 626–630.
- Gurney, K. R., et al. (2008), *The Vulcan Inventory*, Version 1.0, edited, Purdue University.
- Hudiburg, T., et al. (2009), Carbon dynamics of Oregon and Northern California forests and potential land-based carbon storage, *Ecol. Appl.*, 19(1), 163–180.
- Intergovernmental Panel on Climate Change (2007), *Climate Change 2007: Synthesis Report*, 52 pp.
- Irvine, J., et al. (2002), Water limitations to carbon exchange in old-growth and young ponderosa pine stands, *Tree Physiol.*, 22(2–3), 189–196.
- Irvine, J., et al. (2007), Postfire carbon pools and fluxes in semiarid ponderosa pine in Central Oregon, *Global Change Biol.*, 13(8), 1748–1760.
- Jacobson, A. R., et al. (2007), A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: 2. Regional results, *Global Biogeochem. Cycles*, 21(1), GB1020, doi:10.1029/2006GB002703.
- Jolly, W. M., et al. (2005), A flexible, integrated system for generating meteorological surfaces derived from point sources across multiple geographic scales, *Environ. Modell. Softw.*, 20(7), 873–882.
- Kaminski, T., et al. (2001), On aggregation errors in atmospheric transport inversions, *J. Geophys. Res.*, 106(D5), 4703–4715.
- Krinner, G., et al. (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cycles*, 19(1), GB1015, doi:10.1029/2003GB002199.
- 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., et al. (2001), Spatial and temporal variation in respiration in a young ponderosa pine forest during a summer drought, *Agric. For. Meteorol.*, 110(1), 27–43.
- Law, B. E., et al. (2002), Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation, *Agric. For. Meteorol.*, 113(1–4), 97–120.
- 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.
- Law, B. E., et al. (2006), Carbon fluxes across regions: Observational constraints at multiple scales, in *Scaling and Uncertainty Analysis in Ecology: Methods and Applications*, edited by J. Wu et al., pp. 167–190, Springer, New York.
- Lin, J. C., and C. Gerbig (2005), Accounting for the effect of transport errors on tracer inversions, *Geophys. Res. Lett.*, 32(1), L01802, doi:10.1029/2004GL021127.
- 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.
- Lokupitiya, R. S., et al. (2008), Estimation of global CO<sub>2</sub> fluxes at regional scale using the maximum likelihood ensemble filter, *J. Geophys. Res.*, 113(D20), D20110, doi:10.1029/2007JD009679.
- Luyssaert, S., et al. (2008), Old-growth forests as global carbon sinks, *Nature*, 455(11), 213–215.
- 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.
- Metropolis, N., et al. (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, 21(6), 1087–1092.
- Michalak, A. M., et al. (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, *J. Geophys. Res.*, 109(D14), D14109, doi:10.1029/2003JD004422.
- Mu, Q., et al. (2007), Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, *Remote Sens. Environ.*, 111, 519–536.
- Mueller, K. L., et al. (2008), Global monthly averaged CO<sub>2</sub> fluxes recovered using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements, *J. Geophys. Res.*, 113(D21), D21114, doi:10.1029/2007JD009734.
- Myneni, R. B., et al. (2001), A large carbon sink in the woody biomass of northern forests, *Proc. Natl. Acad. Sci. U. S. A.*, 98(26), 14,784–14,789.
- Pasquill, F. (1972), Some aspects of boundary layer description, *Q. J. R. Meteorol. Soc.*, 98, 469–494.
- Peters, W., et al. (2005), An ensemble data assimilation system to estimate CO<sub>2</sub> surface fluxes from atmospheric trace gas observations, *J. Geophys. Res.*, 110(D24), 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.
- Peylin, P., et al. (2005), Daily CO<sub>2</sub> flux estimates over Europe from continuous atmospheric measurements: 1. Inverse methodology, *Atmos. Chem. Phys.*, 5, 3173–3186.
- Potter, C. S. (1999), Terrestrial biomass and the effects of deforestation on the global carbon cycle – Results from a model of primary production using satellite observations, *Bioscience*, 49(10), 769–778.
- Potter, C., et al. (2003), Global teleconnections of climate to terrestrial carbon flux, *J. Geophys. Res.*, 108(D17), 4556, doi:10.1029/2002JD002979.
- Rannik, Ü., et al. (2000), Footprint analysis for measurements over a heterogeneous forest, *Boundary Layer Meteorol.*, 97(1), 137–166.
- Rayner, P. J., et al. (1999), Reconstructing the recent carbon cycle from atmospheric CO<sub>2</sub>, δ<sup>13</sup>C and O<sub>2</sub>/N<sub>2</sub> observations, *Tellus, Ser. B*, 51(2), 213–232.

- Reichstein, M., et al. (2007), Determinants of terrestrial ecosystem carbon balance inferred from European eddy covariance flux sites, *Geophys. Res. Lett.*, *34*(1), L01402, doi:10.1029/2006GL027880.
- Running, S., et al. (2000), Global terrestrial gross and net primary production from the Earth Observing System, in *Methods in Ecosystem Science*, edited by O. E. Sala et al., Springer Verlag, New York.
- Schmid, H. P. (1997), Experimental design for flux measurements: Matching scales of observations and fluxes, *Agric. For. Meteorol.*, *87*, 179–200.
- Schmid, H. P. (2002), Footprint modeling for vegetation atmosphere exchange studies: A review and perspective, *Agric. For. Meteorol.*, *113*(1–4), 159–183.
- Schmid, H. P., and T. R. Oke (1990), A model to estimate the source area contributing to turbulent exchange in the surface layer over patchy terrain, *Q. J. R. Meteorol. Soc.*, *116*, 965–988.
- Schuepp, P. H., et al. (1990), Footprint prediction of scalar fluxes from analytical solutions of the diffusion equation, *Boundary Layer Meteorol.*, *50*, 355–373.
- Schuh, A. E., et al. (2009), Seeing the forest through the trees: Recovering large-scale carbon flux biases in the midst of small-scale variability, *J. Geophys. Res.*, *114*(G3), G03007, doi:10.1029/2008JG000842.
- Schwarz, P. A., et al. (2004), Climatic versus biotic constraints on carbon and water fluxes in seasonally drought-affected ponderosa pine ecosystems, *Global Biogeochem. Cycles*, *18*(4), GB4007, doi:10.1029/2004GB002234.
- Shaw, D. C., et al. (2004), Ecological setting of the Wind River old-growth forest, *Ecosystems*, *7*(5), 427–439.
- Skamarock, W. C., et al. (2005), *A Description of the Advanced Research WRF Version 2*, 100 pp., National Center of Atmospheric Research, Boulder, Colo.
- Stauffer, D. R., and N. L. Seaman (1994), Multiscale four-dimensional data assimilation, *J. Appl. Meteorol.*, *33*(3), 416–434.
- Stauffer, D. R., et al. (1991), Use of four-dimensional data assimilation in a limited-area mesoscale model. Part II: Effects of data assimilation within the planetary boundary layer, *Mon. Weather Rev.*, *119*(3), 734–754.
- Stephens, B. B., et al. (2006), An autonomous inexpensive robust CO<sub>2</sub> analyzer (AIRCOA), paper presented at 13th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques, WMO/IAEA, Boulder, Colo.
- Stephens, B. B., et al. (2007), Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO<sub>2</sub>, *Science*, *316*(5832), 1732–1735.
- Tarantola, A. (1987), *Inverse Problem Theory. Methods for Data Fitting and Model Parameter Estimation*, 613 pp., Elsevier Sci., New York.
- Thornton, P. E., and S. W. Running (1999), An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation, *Agric. For. Meteorol.*, *93*(4), 211–228.
- Thornton, P. E., et al. (1997), Generating surfaces of daily meteorological variables over large regions of complex terrain, *J. Hydrol.*, *190*(3–4), 214–251.
- Thornton, P. E., et al. (2002), Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests, *Agric. For. Meteorol.*, *113*(1–4), 185–222.
- Turner, D. P., et al. (2003), Effects of land use and fine-scale environmental heterogeneity on net ecosystem production over a temperate coniferous forest landscape, *Tellus, Ser. B*, *55*(2), 657–668.
- Turner, D. P., et al. (2004), Monitoring forest carbon sequestration with remote sensing and carbon cycle modeling, *Environ. Manage.*, *33*(4), 457–466.
- Turner, D. P., et al. (2006), A diagnostic carbon flux model to monitor the effects of disturbance and interannual variation in climate on regional NEP, *Tellus, Ser. B*, *58*(5), 476–490.
- Turner, D. P., et al. (2007), Scaling net ecosystem production and net biome production over a heterogeneous region in the western United States, *Biogeosciences*, *4*(4), 597–612.
- Valentini, R., et al. (2000), Respiration as the main determinant of carbon balance in European forests, *Nature*, *404*(6780), 861–865.
- Vrugt, J. A., et al. (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.
- Wang, W., et al. (2005), *User's Guide for Advanced Research WRF (ARW) Modeling System Version 2.1*, edited, NCAR.
- Zhao, M., et al. (2005), Improvements of the MODIS terrestrial gross and net primary production global data set, *Remote Sens. Environ.*, *95*(2), 164–176.
- Zupanski, D., et al. (2007), Carbon flux bias estimation employing maximum likelihood ensemble filter (MLEF), *J. Geophys. Res.*, *112*(D17), D17107, doi:10.1029/2006JD008371.

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.gockede@oregonstate.edu)

A. M. Michalak, Department of Civil and Environmental Engineering, University of Michigan, 183 EWRE, 1351 Beal Ave., Ann Arbor, MI 48109-2125, USA.

D. Vickers, College of Oceanic and Atmospheric Sciences, Oregon State University, Ocean Admin. Bldg. 104, Corvallis, OR 97331-5503, USA.