Tellus (2006), 58B, 344–358
Printed in Singapore. All rights reserved

TELLUS

Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data

By DANIEL M. MATROSS[1]*, ARLYN ANDREWS[2], MAHADEVAN PATHMATHEVAN[1], CHRISTOPH GERBIG[3], JOHN C. LIN[4], STEVEN C. WOFSY[1], BRUCE C. DAUBE[1], ELAINE W. GOTTLIEB[1], VICTORIA Y. CHOW[1], JOHN T. LEE[5], CONGLONG ZHAO[6], PETER S. BAKWIN[2], J. WILLIAM MUNGER[1] and DAVID Y. HOLLINGER[7], [1]*Department of Earth and Planetary Sciences Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138, USA;* [2]*Global Monitoring Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, 80305, USA;* [3]*Max-Planck-Institut für Biogeochemie, D-07745 Jena, Germany;* [4]*Department of Earth Sciences, University of Waterloo, Waterloo, ON, Canada, N2L 3GL;* [5]*Environmental Physics Group, University of Maine, Orono, Maine, 04469, USA;* [6]*Cooperative Institute for Research in Environmental Science, University of Colorado, Boulder, CO, 80309, USA;* [7]*United States Department of Agriculture Forest Service, Durham, NH, 03824, USA*

(Manuscript received 16 January 2006; in final form 12 June 2006)

ABSTRACT

We derive regional-scale ($\sim 10^4$ km^2) CO_2 flux estimates for summer 2004 in the northeast United States and southern Quebec by assimilating extensive data into a receptor-oriented model-data fusion framework. Surface fluxes are specified using the Vegetation Photosynthesis and Respiration Model (VPRM), a simple, readily optimized biosphere model driven by satellite data, AmeriFlux eddy covariance measurements and meteorological fields. The surface flux model is coupled to a Lagrangian atmospheric adjoint model, the Stochastic Time-Inverted Lagrangian Transport Model (STILT) that links point observations to upwind sources with high spatiotemporal resolution. Analysis of CO_2 concentration data from the NOAA-ESRL tall tower at Argyle, ME and from extensive aircraft surveys, shows that the STILT–VPRM framework successfully links model flux fields to regionally representative atmospheric CO_2 data, providing a bridge between 'bottom-up' and 'top-down' methods for estimating regional CO_2 budgets on timescales from hourly to monthly. The surface flux model, with initial calibration to eddy covariance data, produces an excellent a priori condition for inversion studies constrained by atmospheric concentration data. Exploratory optimization studies show that data from several sites in a region are needed to constrain model parameters for all major vegetation types, because the atmosphere commingles the influence of regional vegetation types, and even high-resolution meteorological analysis cannot disentangle the associated contributions. Airborne data are critical to help define uncertainty within the optimization framework, showing for example, that in summertime CO_2 concentration at Argyle (107 m) is $\sim$0.6 ppm lower than the mean in the planetary boundary layer.

1. Introduction

Terrestrial carbon budgets at regional and continental scales ($\sim 10^4$ to 10^6 km^2) are key to assessment of human impact on ecosystems and atmospheric composition. However, obtaining and validating regional-scale fluxes of CO_2 and other important trace gases has proven to be a scientific challenge. Efforts to quantify distributions of sources and sinks of CO_2 have focused on global inverse modelling of CO_2 concentration data from the global monitoring network. Most recently, the globe has been split into 10–20 large regions (Gurney et al., 2002, 2004). Aggregation errors and errors in atmospheric transport, both within the boundary layer and between the boundary layer and free troposphere, can be formidable obstacles to using this approach to obtain reliable quantitative estimates of carbon fluxes at regional and continental scales (Gloor et al., 1999). Global-scale inversions cannot account for important planetary boundary layer processes which affect the concentrations on which they are based,

*Corresponding author.
e-mail: matross@fas.harvard.edu
DOI: 10.1111/j.1600-0889.2006.00206.x

subjecting them to an additional representation error (Kaminski and Heimann, 2001).

Observations of CO_2 over the continent can potentially provide the information needed to determine regional fluxes. The analysis must quantitatively account for the large variability introduced by sources and sinks in the near-field of the measurement location (Gerbig et al., 2003a), requiring detailed understanding of terrestrial CO_2 sources and sinks at spatial and temporal scales much finer than those used in global inversions. In particular, climate variations and human impacts are often most readily evident at the regional and continental scales. Methods to quantify CO_2 sources and sinks at this scale, intermediate between global and very localized, are notably lacking. Moreover, different regions can vary markedly in response of the carbon cycle to a changing climate (Friedlingstein et al., 2003; Fung et al., 2005). From a policy perspective, the inability to reliably quantify carbon exchange at the regional scale presents a potential stumbling block to future regulatory goals, for example, developing markets for verifiable trading of CO_2 emissions.

Eddy covariance measurements are a rich source of information on temporal variability and environmental controls of CO_2 exchange between the atmosphere and terrestrial ecosystems (Law et al., 2002); they have semi-continuous temporal coverage at an increasing number of sites across the continent. Unfortunately, it is difficult to reliably scale up eddy-flux measurements to regional scales from the localized footprint ($\sim 10^1$ to 10^2 km^2) of the measurements or to test regional flux models developed using micrometeorological measurements.

Concentration data for CO_2 from tall towers ($>\sim$100 m) provide a potentially powerful constraint on 'bottom-up' flux models, because the relatively large footprint provides an integrated signal of CO_2 exchange at the regional scale (Gloor et al., 2001). Previous efforts to interpret the signal of regional CO_2 exchange using tall tower concentration data have focused on simple one-dimensional atmospheric boundary layer budgets that rely on gradients in CO_2 concentrations between the boundary layer and the free troposphere (Bakwin et al., 2004; Helliker et al., 2004). These methods are limited to monthly resolution by the need to smooth and average over several synoptic events. They also use a marine boundary layer surrogate for the free tropospheric CO_2 concentration over the continent, because of limited observations.

Gerbig et al. (2003a) and Lin et al. (2006) showed, however, that free tropospheric concentrations depart significantly from the MBL reference over the continent, leading to biases in calculations of regional CO_2 flux. These differences result largely from the time lag for vertical propagation of marine boundary layer concentration changes upwards from the surface into the free troposphere (Gerbig et al., 2003a), as well as from meridional transport via meandering of the polar jet, and deep convective events (Gerbig et al., 2003a).

This paper uses CO_2 concentrations from a tall tower at Argyle, Maine to critically test a bottom-up CO_2 surface flux model based on assimilation of eddy covariance fluxes. The surface flux model is an application of the Vegetation Photosynthesis and Respiration Model (VPRM, http://people.deas.harvard.edu/wofsy/VPRM.submitted.zip, hereafter referred to as Pathmathevan et al., 2006), a diagnostic CO_2 flux model with a minimum number of parameters (3 per vegetation class) driven by remote sensing and weather data. We focus on CO_2 sources and sinks for northern New England and southern Quebec in summer 2004, when we made extensive aircraft flights in the region. The Stochastic Time-Inverted Lagrangian Transport (STILT) Model (Lin et al., 2003), an adjoint transport model with high spatiotemporal resolution, provides the quantitative link between surface fluxes calculated with the VPRM and the time-series of concentration observations at the Argyle, Maine. We then examine the number of degrees of freedom in the STILT + VPRM inverse problem, applied at this single tall tower, to lay the groundwork for future inverse studies seeking to obtain regional or continental fluxes using a network of tall tower measurements.

2. Methodology

Gerbig et al. (2003b) developed the receptor-oriented modelling framework that provides the basis for this study, consisting of four major components: (1) influence functions from the STILT model that quantitatively link upstream spatially/temporally resolved surface sources/sinks to concentration measurements at a receptor point (i.e. a measurement location) (2) a lateral continental CO_2 boundary condition for North America from Pacific observations; (3) Fossil fuel CO_2 inventories and (4) a model for surface CO_2 fluxes; in this case, the VPRM of Pathmathevan et al. (2006).

2.1. STILT adjoint atmospheric model

STILT is analogous to the adjoint of an Eulerian transport model, with footprint elements representing the sensitivity of the mixing ratio at receptor location to any given surface flux (Errico, 1997). Information about a footprint, the upstream source region of surface influences on air at a particular measurement point, comes from computing transport of an ensemble of particles—representing air parcels—backwards in time using winds and turbulence statistics from a high-resolution meteorological assimilation. STILT links the local concentration $C(\mathbf{x}_r, t_r)$ of a conserved tracer, measured at a receptor location $\mathbf{x}_r$ at time t_r to the surface sources S for the tracer emitted at upstream locations $\mathbf{x}$ at prior time t, by computing the influence function $I(\mathbf{x}_r, t_r|\mathbf{x}, t)$ (Lin et al., 2003) through eq. (1a).

$$C(\mathbf{x}_r, t_r) = \int_{t_0}^{t_r} dt \int_V d^3x\, I(\mathbf{x}_r, t_r|\mathbf{x}, t) S(x, t) + \int_V d^3x\, I(\mathbf{x}_r, t_r|\mathbf{x}, t_0) C(\mathbf{x}, t_0). \quad (1a)$$

The first term in eq. (1a) represents ΔCO_2 at the receptor due to surface sources in domain V between time t_0 and t_r. The second term in (1a) is the advected contribution from the initial tracer field, which in this case is represented by a boundary condition based on observations in the mid-Pacific (see below). STILT represents surface fluxes as volume sources distributed from the surface through a mixing height h. Lin et al. (2003) recast the first term in eq. (1a) using eq. (1b), which incorporates a surface flux $F(\mathbf{x}, t)$ and a footprint element f, which is mathematically defined in eq. (2).

$$\Delta C(\mathbf{x}_r, t_r) = \int_{t_0}^{t_r} dt \int_A d^2x\, f(\mathbf{x}_r, t_r|\mathbf{x}, t) F(\mathbf{x}, t) \tag{1b}$$

$$f = \int_0^h dz I(\mathbf{x}_r, t_r|\mathbf{x}, t) \times m_{air}/(h\rho_{air}) \tag{2}$$

The value of the footprint calculation depends on the initial column height, h, below which turbulent mixing is strong enough to mix the surface signal. STILT calculates the height of the planetary boundary layer based on a modified Richardson number method (Lin et al., 2003) and h is determined as a fraction of that value. Gerbig et al. (2003b) found that simulated footprints were insensitive to values of h between 10% and 100% of the planetary boundary layer height.

As currently implemented, the transport fields to drive STILT can come from operational global forecast or reanalysis products (e.g. Eta Data Assimilation System, Nested Grid Model, European Center for Medium range Weather Forecast model) or from mesoscale models run specifically for periods and domains of interest [e.g. Regional Atmospheric Modelling System (RAMS), Weather Research and Forecasting model (WRF)]. A stochastic Markov chain is used to represent subgridscale turbulence. STILT can explicitly model convection, for example, in convective storms, if the driver provides convective mass fluxes. The footprint f is derived from the local density of particles by counting the number in surface-influenced boxes and determining the amount of time each particle spends in each surface volume element during each time step; the results can be visualized by plotting the time- and area-integrated footprint of f [<<f>> units: ppm/(μmole m^{-2} s^{-1})]. The principal advantages of STILT include the great care taken to conserve mass and to maintain well-mixed conditions (viz. to obey the second law of thermodynamics), the considerable computational advantage provided by the backward-time, receptor-oriented formulation: that is, running a minimum number of representative particles backwards in time, each of which influences the receptor point, and the high spatial and temporal resolution achieved at modest computational cost (Gerbig et al., 2003b; Lin et al., 2003).

2.2. Empirical lateral tracer boundary condition

A lateral tracer boundary condition is required to connect regional simulations to the global background distribution. We use a statistical approach based on available observations in the Pacific to characterize spatial and temporal dependence of tracer variations. The boundary condition for North America is imposed at 145° W, representing tracer concentrations over the mid-Pacific ocean before air parcels enter the dominant westerly flow over North America. Most particles cross 145° W after being transported back for roughly 6 d from receptor points in the United States (Gerbig et al., 2003b).

Gerbig et al. (2003b) describe the full development of the lateral boundary condition, including a complete categorization of the extensive airborne and station data used in its development. The statistical analysis consists of a Fourier decomposition of observed CO_2 time-series from marine surface stations to yield an analytical representation, followed by fitting a Green's function to aircraft CO_2 measurements, to represent the vertical propagation of the seasonal signal at the into the middle and upper troposphere. For this study, we updated the station data from Pacific ground stations in the NOAA GMD network (Cape Kumakahi, HI; Cold Bay, AK; and Barrow, AK) to include the period from 1 January 1980 to 31 December, 2004. In addition, we updated upper air profiles from regular NOAA GMD aircraft flights over Carr, CO, Poker Flats, AK and Park Falls, WI during 2003 and 2004, along with measurements from the Niwot Ridge surface station. The result is a time–latitude–height boundary condition, based on data from 1980 to 2004, consisting of meridional (145° W) cross-sections for CO_2 with spatial resolution of 0.5 km altitude by 2.5° latitude and daily time resolution.

2.3. Fossil fuel inventory

The fossil fuel inventory is unchanged from Gerbig et al. (2003b). Emissions of CO_2 from fossil fuel, cement production and gas flaring come from the $1^\circ \times 1^\circ$ database compiled by Marland et al. (1997), with methodology described by Andres et al. (1996). A linear extrapolation is applied, propagating the trend between 1992 and 1996, and resulting in a 10% total increase in emissions since 1996. Time-of-day and day-of-week scaling factors are applied to account for time dependence of emission fluxes following Ebel et al. (1997). As described in Gerbig et al. (2003b), the time factors for carbon monoxide are applied to CO_2, with the amplitude reduced by a factor of 2.5. The day-of-week factors are 0.95 on weekends and unity on weekdays. Time-of-day factors average to unity, and range from 0.69 in the middle of the night to 1.3 for the rush-hour peak.

2.4. Vegetation photosynthesis and respiration model

The VPRM is a powerful new data-driven diagnostic biosphere flux model fully described in Pathmathevan et al. (2006),

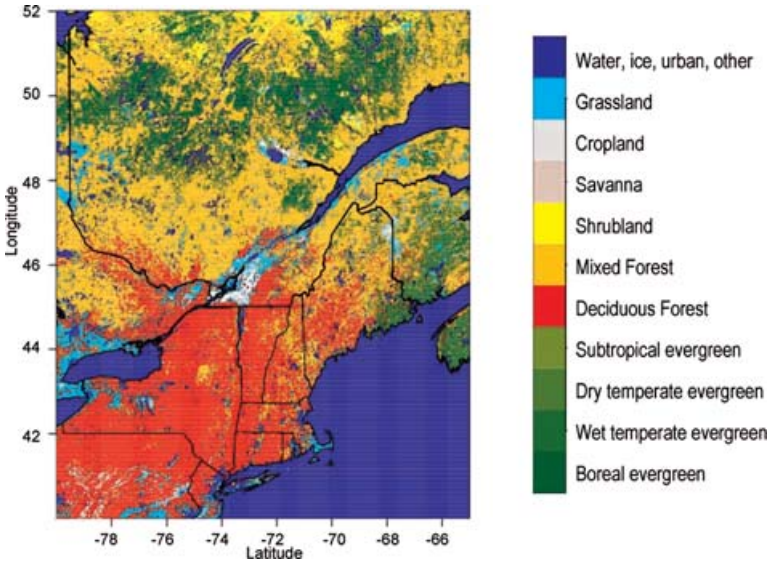

Fig. 1. VPRM vegetation classification (1 km resolution) for the northeast United States and southern Quebec, the area with significant influence on Argyle observations. This is modified from the GLCC 2.0 database (Loveland et al., 2000; Pathmathevan et al., 2006).

extending the Vegetation Photosynthesis Model developed by Xiao et al. (2004) to include respiration and saturation of photosynthesis at high light levels. The VPRM conceptually partitions sunlight between photosynthetically active vegetation and non-photosynthetic components within the leaf and canopy. Satellite data provide independent information on the spatial and phenological variations of gross primary production using the Enhanced Vegetation Index (EVI) and Land Surface Water Index (LSWI), both from MODIS-Terra. Model parameters are initially determined through fitting to eddy covariance data from AmeriFlux sites. The model uses temperature from the same meteorological files used by STILT, and incident solar radiation from those same fields or from retrievals based on data from the North American Land Data Assimilation System (NLDAS; http://ldas.gsfc.nasa.gov/, Mitchell et al., 2003). Net flux is computed every hour on a grid of 1/4° longitude by 1/6° latitude.

The VPRM bins the GLCC 2.0 1-km resolution vegetation inventory (Loveland et al., 2000; http://edcsns17.cr.usgs.gov/glcc/) into 11 classes (10 vegetation + 1 flux neutral class; Fig. 1) and calculates net ecosystem exchange (NEE) of CO_2 for each vegetation class in each grid square separately, scaled by vegetation fraction. NEE is the sum of two model terms: a light-dependent term, identified with canopy photosynthesis (GEE), and a temperature-dependent term identified with ecosystem respiration (R). GEE, as calculated in eq. (3), is assumed to be a function of shortwave incident flux (SW) and to the observed EVI (Huete et al., 1997).

$$\mathrm{GEE} = \lambda \times T_{\mathrm{scale}} \times W_{\mathrm{scale}} \times P_{\mathrm{scale}} \times \frac{1}{(1 + \frac{\mathrm{SW}}{\mathrm{SW}_0})} \times \mathrm{SW} \times \mathrm{EVI}, \tag{3}$$

where SW_0 is the half-saturation value for photosynthesis derived from eddy covariance data for each vegetation type.

The calculation of GEE in eq. (3) includes several scalar functions. These light-, temperature- and water-dependant scalars are defined as needed for individual vegetation classes in order to accommodate the large range of light- and water-utilization strategies observed in nature. The first scalar in eq. (3) is a temperature function (T_{scale}), as calculated in eq. (4), capturing temperature sensitivity of photosynthesis for each vegetation type.

$$T_{\mathrm{scale}} = \frac{(T - T_{\mathrm{min}}) \times (T - T_{\mathrm{max}})}{[(T - T_{\mathrm{min}}) \times (T - T_{\mathrm{max}}) - (T - T_{\mathrm{opt}})^2]} \tag{4}$$

The second scalar in eq. (3) adjusts for water stress (W_{scale} based on the LSWI), as calculated in eqs. (5a) and (5b).

$$W_{\mathrm{scale}}(\mathrm{grassland/savanna}) = \frac{(\mathrm{LSWI} - \mathrm{LSWI}_{\mathrm{min}})}{(\mathrm{LSWI}_{\mathrm{max}} - \mathrm{LSWI}_{\mathrm{min}})} \tag{5a}$$

$$W_{\mathrm{scale}}(\mathrm{other\ vegetation\ types}) = \frac{(1 + \mathrm{LSWI})}{(1 + \mathrm{LSWI}_{\mathrm{max}})} \tag{5b}$$

The third scalar in eq. (3) is a phenology-tracking function based on LSWI (P_{scale}),

$$P_{\mathrm{scale}}(\mathrm{evergreen}) = 1 \quad \{\mathrm{all\ times}\} \tag{6a}$$

$$P_{\mathrm{scale}}(\mathrm{grassland/savanna}) = \frac{(1 + \mathrm{LSWI})}{2} \quad \{\mathrm{all\ times}\} \tag{6b}$$

$$P_{\mathrm{scale}}(\mathrm{other\ vegetation\ types}) = 1 \quad \{\mathrm{full\ canopy\ period}\} \tag{6c}$$

$$= \frac{(1 + \mathrm{LSWI})}{2} \quad \{\mathrm{bud - burst\ to\ full\ canopy}\}. \tag{6d}$$

The bud-burst and full canopy periods of eqs. (6c) and (6d) are based on EVI, as outlined in Pathmathevan et al. (2006).

In the calibration of the VPRM to eddy covariance data, GEE is multiplied by an adjustable parameter (λ) for each vegetation type, representing the overall light use efficiency of the ecosystem. Adjustments which scale λ in our inverse modelling studies

represent adjustments of the photosynthesis efficiency for a particular vegetation class over the landscape, as constrained by the tall tower data.

The VPRM utilizes a linear formulation for respiration (R), as calculated in eq. (7). R is a function of a second calibration factor (α) that captures the dependence of respiration on air temperature, when temperatures are above a minimum temperature T_{rmin}, and an additional calibration parameter (β) represents the basal respiration rate (e.g. during winter when $T < T_{rmin}$) for a given vegetation type.

$$R = \alpha \times T + \beta \{T > T_{r\,min}\} \quad R = \alpha \times T_{r\,min} + \beta \{T \leq T_{r\,min}\} \tag{7}$$

The sets of three scaling parameters $\{\lambda, \alpha, \beta\}$ of the VPRM were calibrated for each of ten vegetation classes with eddy covariance data, using all valid hourly NEE measurements for 1–4 yr, at ten AmeriFlux sites and validated using data from ten separate sites. The parameters were *not* initially adjusted to fit atmospheric concentration constraints in the present work. Model results with these 30 calibration parameters (dependent on vegetation class, invariant with time, determined using eddy covariance data) accounted for 60–80% of the variance of hourly flux data at the calibration sites, and 50–75% at validation sites; in most cases predictions of seasonal and annual sums at validation sites were quite close to observed values (Pathmathevan et al., 2006).

2.5. Argyle tall tower measurements

The receptor-oriented modelling framework (STILT, VPRM, fossil fuel inventories, lateral boundary condition) allows calculation of CO_2 concentration at a given receptor point in space for any hour given sunlight and meteorological data. In order to leverage simultaneous measurements being taken during the COBRA-Maine airborne campaign (see below, and Lin et al., 2006), the primary receptor point for this study is the NOAA Argyle tall tower, run by the Global Monitoring Division (GMD) of NOAA's Earth System Research Laboratory (ESRL). STILT is used to determine the influence functions and the upstream footprint, which are multiplied by the spatially resolved, hourly biosphere fluxes from the VPRM and fossil fuel fluxes from the inventory. Contribution from CO_2 advected from the lateral boundary is added and the result compared to the observations at the receptor.

The Argyle site (45.03° N, 68.68° W) is a cell phone tower located in a northern mixed deciduous, evergreen forest in rural central Maine north of Bangor, elevation 50 m above sea level. The experimental set-up is similar to the WLEF tower (Bakwin et al., 1998), with CO_2 and CO concentration measurements taken every 4 minutes at 12 m and 107 m above ground level. CO_2 is measured with a Licor Li-7000 CO_2/H_2O analyser calibrated with five standard gases. CO is measured with a Thermo Electron model 48CTL CO analyser calibrated using two standard gases. The zero concentration reference for the CO analyser is checked by catalytically scrubbing CO from ambient air using Sofnocat as a reagent as part of the standard measurement protocol. Supplementary flask samples are collected weekly from the top level and are shipped to the GMD/ESRL laboratory in Boulder, CO for analysis of CO_2, N_2O, CH_4, SF_6, H_2 and CO concentrations, as well as isotope ratios of ^{13}C and ^{18}O in CO_2. NOAA GMD began sampling at Argyle on 17 September 2003.

The focus of this study is the time period of May to August 2004, coinciding with the COBRA-Maine airborne campaign. COBRA-Maine utilized the University of Wyoming King Air for nearly 200 flight hours over 59 flights in Maine, greater New England and southern Quebec to characterize regional carbon exchange. The campaign was based out of the Bangor International Airport, located roughly 30 km south of Argyle. In total, the aircraft collected over 900 vertical profiles of CO_2, CO, water vapour, ozone and meteorological data from the surface to altitudes up to 8 km. Over 120 of those vertical profiles were flown within 50 km of the Argyle tall tower. The flight data provide characterization of CO_2, CO and atmospheric structure in the boundary layer and free troposphere in the vicinity of Argyle.

2.6. Application STILT + VPRM at argyle tall tower

We utilized the STILT + VPRM receptor-oriented modelling framework to compute hourly concentrations during summer 2004 for the receptor at 107 m above the ground at Argyle. For each hour, STILT determined influence functions using 100 particles, transported back 6 d in time by Eta Data Assimilation System 40 km (EDAS-40) reanalysis fields obtained from the NOAA Air Resources Lab server (Rolph, 1997; ftp://www.arl.noaa.gov/pub/archives/edas40/). For consistency with Gerbig et al. (2003b), we took an initial column height h for the input of surface emissions (see eq. 2) to be 50% of the planetary boundary layer height, as calculated from the reanalysis fields. Because we did not have convective mass fluxes available from EDAS-40, we did not explicitly model convection.

Eight-day EVI and LSWI were computed directly from MODIS radiances at 1 km resolution for each of the eleven vegetation classes (ten vegetated, plus non-vegetated such as ocean), using quality control as outlined in Pathmathevan et al. (2006), and aggregated onto a rectangular surface grid spanning 30° to 65° N and 51° to 140° W with resolution 1/6° latitude by 1/4° longitude, taking trimmed mean values from 1 km data for each parameter, for each vegetation type within a grid square. Temperatures to drive the VPRM were taken from the EDAS-40 analysis fields and shortwave radiation was taken from incoming downward solar radiation tabulated by the North American Land Data Assimilation System (NLDAS; Mitchell et al., 2003).

The convolution of STILT influence functions and VPRM fluxes yields a value for CO_2 concentration at Argyle for any hour. We calculated STILT + VPRM concentrations at Argyle for all hours in two representative 15 d periods, one from 1 June to 15 June and one from 1 August to 15 August. We also determined

concentrations for midday hours (1400 GMT to 2100 GMT) for the entire summer, 15 May to 15 September 2004. EDAS fields do not accurately reflect the near-surface atmospheric environment at night, when the tower is influenced mainly by nearby sources. Hence we focused our model-data comparisons on daytime values.

2.7. Optimization approach

In order to place large-scale constraints on the VPRM parameters using the link to atmospheric concentration data at Argyle, we performed a simplified Bayesian optimization approach (Rodgers, 2000) with two adjustable coefficients for each vegetation type, γ and ρ, one respectively for GEE and R. (Gerbig et al., 2003):

$$\mathrm{NEE} = \gamma \times \mathrm{GEE} + \rho \times R. \tag{8}$$

Here GEE and R are computed from eqs. (3) and (7), respectively. The a priori model ($\gamma = \rho = 1$ for all vegetation classes) has assimilated a large volume of eddy covariance information through the calibration factors (λ, α, β), but no constraint from atmospheric concentration data had been applied. Thus deviations of Bayesian-optimized γ and ρ from 1 encapsulate the deviation between local-scale carbon fluxes calibrated against eddy covariance measurements and regional-scale carbon fluxes constrained against atmospheric CO_2 observations.

As outlined in Rodgers (2000) and Gerbig et al. (2003b), the measurements at Argyle can be related to vegetation signal (ΔCO_2) from each vegetation class through

$$\mathbf{y} = \mathbf{K}\,¡ + ", \tag{9}$$

where $\mathbf{y}$ is a vector of measurements, $\mathbf{K}$ is the Jacobian matrix relating the measurement vector to the state vector, ¡ is the state vector of scaling factors (γ's and ρ's), and " is an error vector accounting for uncertainties in measurements and modelling framework. As applied here, the measurement vector $\mathbf{y}$ has a single element for each hourly value at Argyle, calculated by subtracting from the Argyle observation the computed contributions advected from the boundary and associated with fossil fuel combustion. The Jacobian matrix $\mathbf{K}$ computes the CO_2 signals for each vegetation class as determined by the STILT + VPRM from the surface flux model.

The optimum posterior estimates of the scaling factors γ and ρ for each vegetation class, based on atmospheric concentration information from Argyle, are obtained by minimizing the cost function $\mathbf{J}$ using a standard least squares formulation

$$\mathbf{J}(\Gamma) = (y - K\Gamma)^{\mathrm{T}} S_{\varepsilon}^{-1}(y - K\Gamma) + (\Gamma - \Gamma_{\mathrm{prior}})^{\mathrm{T}} S_{\mathrm{prior}}^{-1}(\Gamma - \Gamma_{\mathrm{prior}}). \tag{10}$$

In eq. (10), there are two error covariance matrices, one for the vegetation signals, $\mathbf{S}_{\varepsilon}$, and one for the prior scaling factors, $\mathbf{S}_{\mathrm{prior}}$. $\mathbf{S}_{\mathrm{prior}}$ comes from work extensively detailed in Pathmathevan et al. (2006) comparing observed eddy covariance fluxes to the initial calibration of the VPRM, resulting in standard deviations on the initial calibration factors that can be normalized and propagated for use as prior uncertainties for scaling factors γ and ρ. The off-diagonal elements of $\mathbf{S}_{\mathrm{prior}}$ are assumed to be 0, since the calibration of the VPRM does not introduce covariance in errors between vegetation classes. The values of the diagonal elements in $\mathbf{S}_{\mathrm{prior}}$ for the different vegetation classes are shown in Table 3. The formulation $\mathbf{S}_{\varepsilon}$ of is detailed below.

Posterior estimates of ¡, optimally consistent with both Argyle measurements and prior estimates of fluxes from the initial calibration of the VPRM, are obtained by minimizing $\mathbf{J}$. ¡ post can be calculated using eq. (11) (Rodgers, 2000).

$$\Gamma_{\mathrm{post}} = (K^{\mathrm{T}} S_{\varepsilon}^{-1} K + S_{\mathrm{prior}}^{-1})^{-1}(K^{T} S_{\varepsilon}^{-1} y + S_{\mathrm{prior}}^{-1}\Gamma_{\mathrm{prior}}) \tag{11}$$

The uncertainty of ¡ post is expressed through a posterior error covariance matrix, $\mathbf{S}_{\mathrm{post}}$, as shown in eq. (12).

$$S_{\mathrm{post}} = (K^{\mathrm{T}}\mathbf{S}_{\varepsilon}^{-1}\mathbf{K} + \mathbf{S}_{\mathrm{prior}}^{-1})^{-1} \tag{12}$$

Following Gerbig et al. (2003b), the error covariance matrix $\mathbf{S}_{\varepsilon}$ was assumed to be additive:

$$\mathbf{S}_{\mathrm{e}} = \mathbf{S}_{\mathrm{veg}} + \mathbf{S}_{\mathrm{part}} + \mathbf{S}_{\mathrm{eddy}} + \mathbf{S}_{\mathrm{transp}} + \mathbf{S}_{\mathrm{aggr}} + \mathbf{S}_{\mathrm{ocean}}. \tag{13}$$

$\mathbf{S}_{\mathrm{veg}}$ is the uncertainty in the vegetation signal (total ΔCO_2), as given by $\mathbf{y}$ in eq. (9). The associated error covariance matrix $\mathbf{S}_{\mathrm{veg}}$ is derived by propagating measurement uncertainty from the Argyle observations (variable between 0.1 and 0.3 ppm) with estimated uncertainties in the advected background and in the contributions from fossil fuels. We used a value of 1.15 ppm as the standard deviation for normally distributed errors in the advected background, as calculated by Gerbig et al. (2003b). This value is supported by measured differences between calculated advected background values and corresponding free troposphere observations in the vicinity of Argyle during COBRA-Maine (Section 3.3) for which the standard deviation was 1.20 ppm. Uncertainty in the contribution due to fossil fuel combustion was assumed to be 30% of the signal (Gerbig et al., 2003b). We did not account for any influence biomass burning may have had on the Argyle measurements; 2004 was generally wet in the region and there were no large fires near Argyle.

$\mathbf{S}_{\mathrm{part}}$ is the random error due to particle statistics, and is taken to be a diagonal matrix with value 13% of the Argyle signal, in accordance with the relative uncertainties for mixed layer receptor points derived in Gerbig et al. (2003a) for simulations with 100 particles. $\mathbf{S}_{\mathrm{eddy}}$ is the error due to unresolved eddies, and is calculated from a statistical comparison of Argyle concentrations to mean concentrations within the mixed layer nearby as measured by the aircraft during COBRA-Maine (Section 3.3).

$\mathbf{S}_{\mathrm{transp}}$ is the sum of two terms. (1) The influence of error in modelled height of the mixed layer is calculated from aircraft profiles done near Argyle during COBRA-Maine (Section 3.3). For profiles within 50 km of Argyle, the mean difference between EDAS-40 mixed layer height and manually observed height was

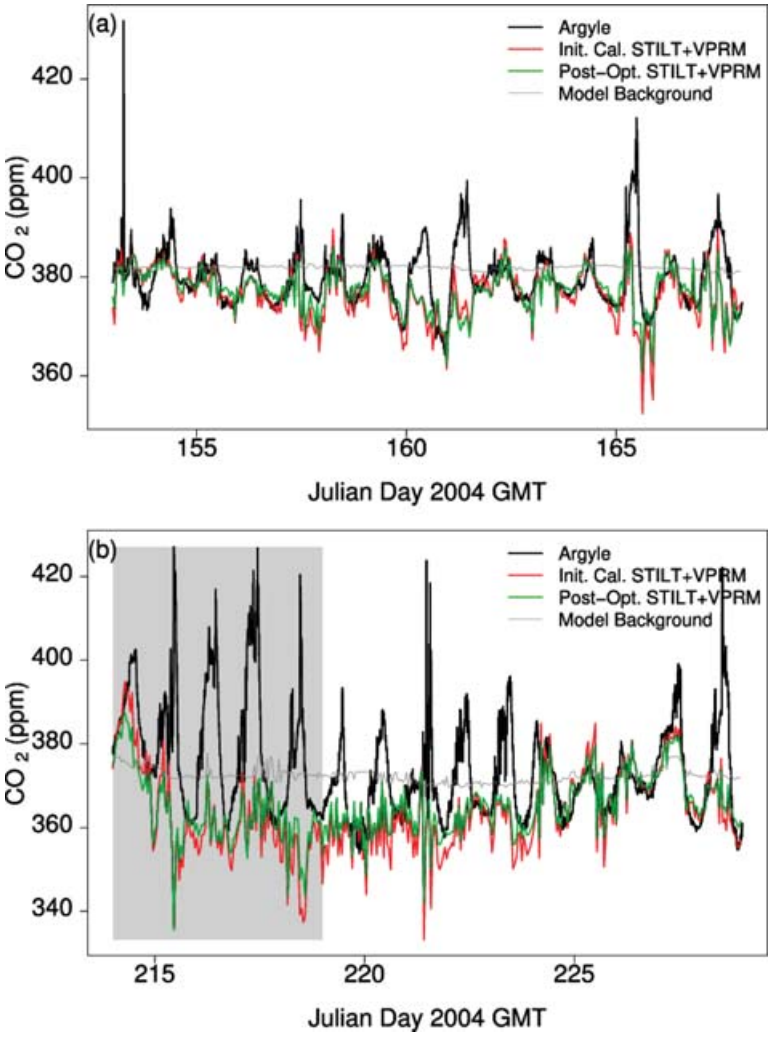

Fig. 2. Time-series of observations from the NOAA CMDL tall tower site (black line) and STILT + VPRM derived hourly CO_2 concentration at the Argyle receptor (red and green lines) point, with advected model background (light grey line) for 1–15 June 2004 (a) and 1–15 August 2004 (b). STILT + VPRM results based on the initial calibration are in red; results with the a posteriori scaling factors from the Bayesian inversion (Section 3.4) are in green. Differences in modelled concentration from the background line are the product of VPRM fluxes and STILT-derived influence functions. Hurricane Alex influenced transport during 1–6 August 2004. This time period is shaded and not included in the Bayesian Inversion.

200 m, with standard deviation of 500 m. The first part of $\mathbf{S}_{\text{transp}}$ was calculated by taking the variance in relative differences between EDAS-40 and manual mixed layer height (mean = 0, standard deviation = 0.42) and multiplying by the vegetation signal for that each time step. (2) Tracer transport error is computed using the method of Lin and Gerbig (2005), in which direct comparison is made between EDAS assimilated winds and data and radiosonde observations. Statistics of transport errors from this comparison are propagated stochastically through STILT as an added error component reflecting uncertainty in the winds. Tracer transport uncertainty is defined for each hour as the associated variance in concentrations, as determined by eq. (1).

$\mathbf{S}_{\text{aggr}}$ represents the error due to aggregation of fluxes into large finite regions, with diagonal elements estimated as in Gerbig et al. (2003b) to be $(3\ \text{ppm})^2$. Similarly, the error due to neglecting oceanic fluxes, $\mathbf{S}_{\text{ocean}}$, is taken directly from Gerbig et al. (2003b) as the upper limit of what those fluxes might have a contributed, 0.1 μmol/m^2/s.

Errors in transport ($\mathbf{S}_{\text{transp}}$) and aggregation ($\mathbf{S}_{\text{aggr}}$) were assumed to have a temporal correlation that decreased exponentially with a decorrelation timescale of 12 hr. Consequently, there were off-diagonal elements in $\mathbf{S}_{\text{transp}}$ and $\mathbf{S}_{\text{aggr}}$ that were calculated by multiplying the diagonal elements with a matrix made from a simple exponential decay function. The 12 hr timescale is meant to capture the synoptic scale variation expected to affect $\mathbf{S}_{\text{transp}}$ and $\mathbf{S}_{\text{aggr}}$. The other terms in eq. (13) were assumed to have negligible temporal correlation.

3. Results and Discussion

3.1. Comparison of observations and initially calibrated STILT + VPRM results

STILT + VPRM concentrations at Argyle are consistently very close to daytime observations during most periods using the coefficients derived from eddy flux data without adjustment ($\gamma = \rho = 1$ for all vegetation classes). Figure 2 shows modelled CO_2

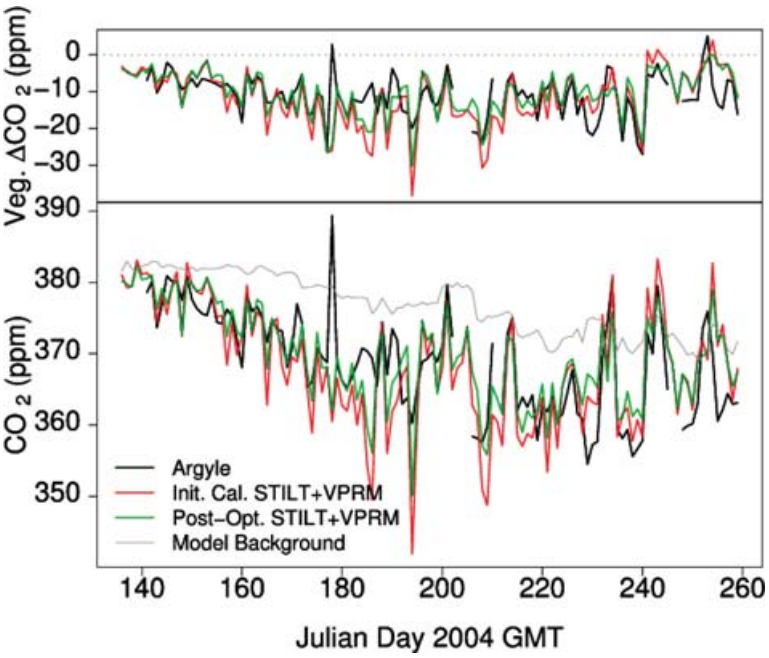

Fig. 3. (Top) Time-series of daily afternoon mean $CO_2{}^{veg}$ (i.e. ΔCO_2 sum from all vegetation classes) for 15 May–15 June 2004. There is one point per day along each line. STILT + VPRM results based on the initial calibration are in red; results with the a posteriori scaling factors from the Bayesian inversion (Section 3.4) are in green. The Argyle signal (black) is determined by subtracting from the observed concentrations the advected background (Section 2.2) and advected fossil fuel contribution (Section 2.3) determined by STILT + VPRM. (Bottom) Time-series of mean afternoon observations from Argyle (black line) and corresponding mean STILT + VPRM-derived afternoon CO_2 concentration (red and green lines) point, along with advected model background (grey line) for 15 May–15 June 2004. The fossil fuel contribution is not explicitly shown, but included in the STILT + VPRM calculations of CO_2 concentration.

concentrations and model background values calculated from the advected continental boundary condition along with those measured at Argyle tower, for two 15 d periods, in June and in August 2004. STILT + VPRM also captures summertime seasonal trends of ambient concentrations observed at Argyle. Figure 3 (top) shows the STILT + VPRM estimates of $CO_2{}^{veg}$ compared directly to with the Argyle data, from which we have subtracted the advected background and fossil fuel contributions. Note the drawdown that increases in early summer, then decreases later in the season, reflecting activity of forests in the region. Most synoptic variations are well simulated (e.g. Fig. 3, bottom; days 230–233 and days 240–245). These reflect large-scale changes in environmental drivers (sunlight, temperature), changes in advected background and fossil fuel influence, and changes in the particular regions influencing measurements at Argyle.

Due to limitations of the meteorological driver (EDAS), STILT has difficulty accurately reproducing transport on stable nights. This is manifest in the time-series by the variable fidelity of STILT + VPRM predictions of the large positive CO_2 concentration excursions in the observations, associated with build-up of CO_2 in the stable nocturnal boundary layer. Significant systematic errors also accrue during periods of convection, when STILT may calculate erroneous transport and footprint fields because EDAS reanalysis fields do not represent convective mass transport. We see this effect as errors in CO_2 concentration at the receptor point at times when frontal systems are near the receptor (Gerbig et al., 2003b). For example, during the period around 0000 GMT on 10 June 2004 (Fig. 2a, Julian day 161), NCEP surface charts and GOES satellite data show a distinct cold frontal passage at Argyle.

On seasonal timescales, a period of frequent stormy weather, such as 19 June to 9 July (Fig. 3, Julian days 170 to 190), shows clear-cut deviations from observations. We will in the future use a mass flux scheme (e.g. Grell and Devenyi, 2002) with transport fields from mesoscale models that can output convective mass fluxes (e.g. WRF, RAMS), in the hope of enhancing the power of STILT + VPRM to consistently link regional flux fields to observed concentration values including convective influence (Gerbig et al., 2005).

In Fig. 2b, the STILT + VPRM does not match the observations during the period from 1 August to 6 August 2004 (Julian days 214–219, shown in grey). We attribute this to the inability of the transport model to capture atmospheric transport associated with Hurricane Alex. Alex tracked to the northeast staying a few hundred kilometres offshore, partly out of the EDAS domain. It strengthened to Category 3 as it moved through latitude 40° N on 4 August. This period has been removed from our Bayesian inversion (Section 3.4).

3.2. *Footprint and vegetation influence*

Two vegetation types dominate summertime influence at Argyle: mixed forest and deciduous forest. They contribute on average at least an order of magnitude greater ΔCO_2 at Argyle during summer 2004 than the other vegetation types (Table 1). Wet temperate evergreen forest (Eastern white pine) provides the greatest influence of the remaining classes.

Footprint calculations display the principal areas that influence Argyle CO_2 data, mostly located in Maine and Quebec. Figure 4 shows the 5 d time-integrated STILT footprint for Argyle averaged over each daytime hour during the period from 15

Table 1. Mean ΔCO_2 (ppm) for different vegetation classes at Argyle

	Boreal evergreen	Wet temp. evergreen	Dry temp. evergreen	Subtropical evergreen	Deciduous forest	Mixed Forest	Shrub	Savannah	Crop	Grass
June	−0435	−0.512	−0.017	−0.065	**−4.99**	**−5.95**	+0.215	−0.046	−0.260	+0.357
July	−0.790	−1.06	−0.012	−0.034	**−5.48**	**−9.31**	+0.132	−0.029	−0.655	−0.171
August	−0.662	−0.475	−0.011	−0.010	**−3.46**	**−7.29**	+0.191	−0.049	−0.407	−0.253

Mean ΔCO_2 contribution of each VPRM vegetation class to the modelled concentration at Argyle for daytime hours in June, July and August 2004. The two primary contributors for each month are in boldface type.

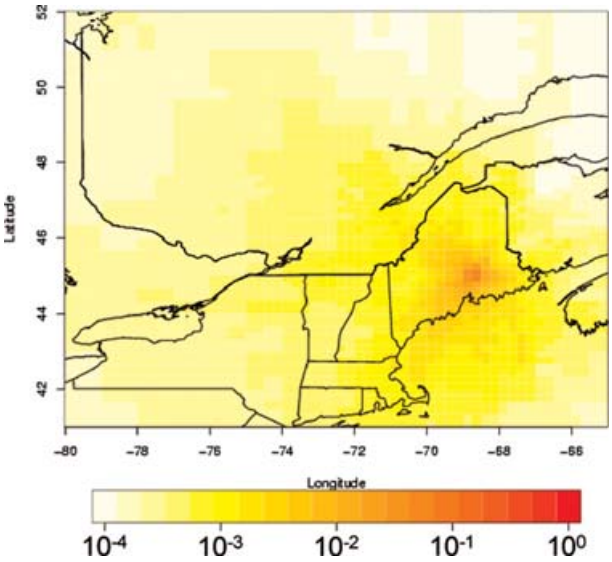

Fig. 4. Visualization of the average 5 d footprint (<<f>>, see eq. 2) for Argyle tower for daytime hours during the period 15 May to 15 September 2004. For each hour, <<f>> is calculated by taking the footprint function and integrating over all areas and all times for 5 d upstream, and the mean result is shown. Values less than 10^{-5} are not coloured.

May to 15 September 2004. The area coinciding with the average footprint is dominated by deciduous and mixed forest (Fig. 1) as expected. These are among the ecosystem types for which the VPRM was most effective in capturing seasonal GEE and respiration variations, and where the model performed best in cross-validations (Pathmathevan et al., 2006). The signal from wet temperate evergreens comes from east of Argyle in Nova Scotia, New Brunswick, and easternmost Maine. Unfortunately there is not really a suitable eddy flux site for this biome, and the VPRM had to use Niwot Ridge, a montane forest, as the calibration site. Influence from this biome also coincides with unsettled weather in many cases.

Fossil fuel influence on Argyle was consistent across the summer. The monthly means for summer are shown in Table 2. Fossil fuel influence at Argyle on a any given day can be vary, depending on whether air arrived at Argyle from relatively rural southern Quebec or from the major cities of eastern United States. Aside from synoptic timescale variability, fossil fuel influence is relatively constant reflecting very large-scale influences far from the site.

Table 2. Fossil fuel ΔCO_2 (ppm) at Argyle

	Mean	Standard deviation
June	2.5	3.1
July	3.5	2.9
August	3.5	2.7

Mean and standard deviation of monthly ΔCO_2 contribution from fossil fuels at Argyle, as calculated by STILT + VPRM.

Monthly average maps of net ecosystem exchange are shown in Fig. 5, calculated using the VPRM to determine the mean of all the hourly values at each grid square for each month in summer, 2004. Within the region that influences concentration measurements at Argyle tower (see Fig. 4), there is a large increase in uptake from May to June, then a gradual decrease through July and August, corresponding to the seasonal pattern expected for a mid-latitude area dominated by deciduous and mixed forest. Overall, VPRM-derived fluxes appear to capture both the spatial heterogeneity and temporal variations needed to meaningfully assess patterns in regional terrestrial carbon exchange.

Using the daytime results calculated for the entire summer, we computed average morning concentration values (10 a.m.–1 p.m. eastern daylight time; 1400–1700 GMT) and afternoon concentration values (2 p.m.–5 p.m. eastern daylight time; 1800–2100 GMT) for each day. We removed from the Argyle data periods with low turbulent mixing (mean friction velocity, $u^* < 0.2$ m/s), to avoid times where the tower is not well coupled to the overlying atmosphere (Lin et al., 2003; Gerbig et al., 2003b). Figure 6 compares these averages to Argyle concentration data averaged over the same periods of the day. Averaging over several hours allows reduction of hour-to-hour noise in both measurements and model results, while preserving the signal contained in diurnal patterns that reflect the daily cycle of photosynthetic uptake.

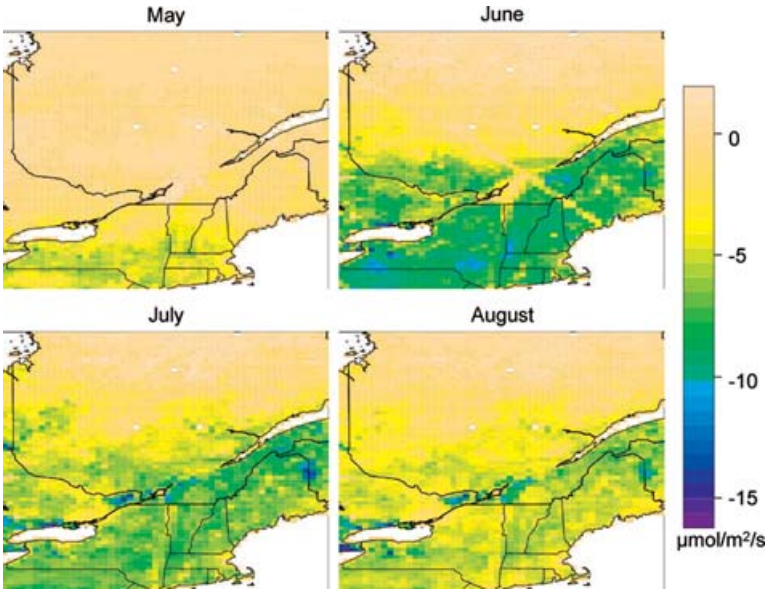

Fig. 5. Initially calibrated VPRM calculated mean monthly (all hours) fluxes for May (a), June (b), July (c) and August (d) 2004 in μmol/m^2/s, based on hourly calculations.

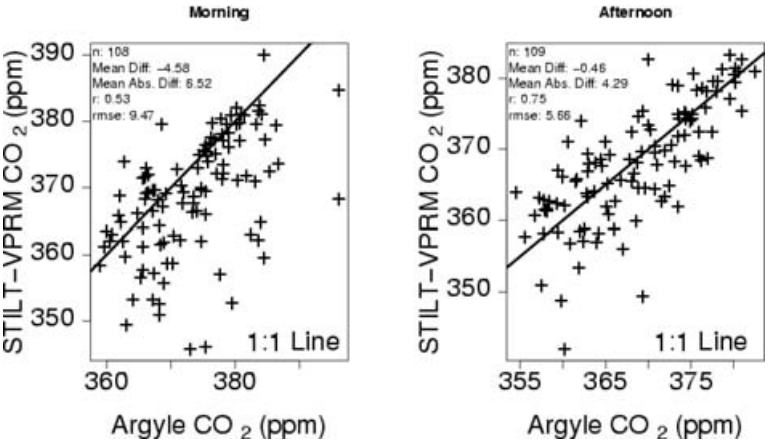

Fig. 6. Direct comparison of STILT + VPRM derived concentrations prior to Bayesian optimization versus Argyle observations for averaged over morning (1400 GMT–1700 GMT) (a) and afternoon (1800 GMT–2100 GMT) (b). The performance of the model is best in afternoon, when conditions are most likely to be well mixed. The 1:1 line is also shown. Periods with consistently low turbulent mixing, indicated by mean friction velocity, u^*, less than 0.2 m/s (seven points in each case) have been eliminated.

STILT + VPRM captures day-to-day variations in CO_2 concentration observed at Argyle somewhat better in the afternoon than in the morning, with root mean square error of 5.66 ppm versus 9.53 ppm, respectively. The mean difference and mean absolute difference in the morning are about the same as the root-mean-square, indicating systematic underestimation of CO_2. The distribution of residuals for the afternoon is more Gaussian with significantly smaller bias. Morning mean observations at Argyle are likely influenced in some cases by residual build-up of CO_2 from the previous night, a locally influenced effect that is insignificant most afternoons. The difference between morning and afternoon values might also indicate errors in the unadjusted VPRM respiration or uptake parameters, particularly in the mixed forest and deciduous forest vegetation surrounding Argyle. More likely, the morning bias reflects the inability of EDAS-40 to adequately capture growth of the planetary boundary layer.

3.3. *Comparisons with COBRA-Maine airborne data*

Aircraft data collected during the COBRA-Maine campaign give us valuable information through comparisons with Argyle observations and STILT + VPRM results. Comparisons with tower data allow direct quantification of error distributions (e.g. representation error) needed to perform a Bayesian optimization. Figure 7 shows the characteristic pattern during an afternoon cross-section observed travelling eastwards across Maine to the coastal region. STILT + VPRM effectively captures the patterns along the cross-section, but the west-to-east gradient is overestimated within the boundary layer and there is a noticeable deviation of the mid-tropospheric values. These differences are a convolution of transport errors, including boundary layer height (especially in the coastal zone), and errors in the VPRM parameters, with a potential minor contribution from misclassification of vegetation. Differences are also caused by processes not

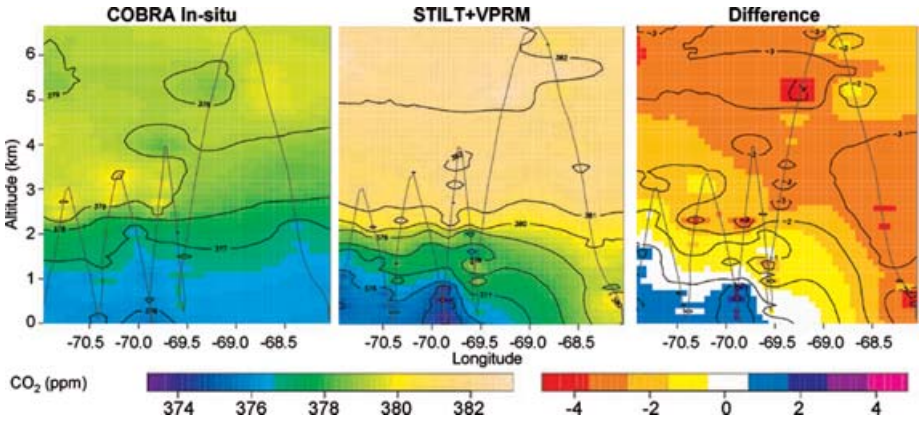

Fig. 7. Comparison of COBRA-Maine in-situ CO_2 (ppm) with STILT + VPRM derived CO_2 (ppm) along a flight track that traversed the state of Maine eastwards on the afternoon of 11 June 2004. Longitude-altitude cross-sections using a distance-weighted interpolation, with the flight track line in light grey are shown for observations (left), the STILT + VPRM results (centre), and the difference between the two (right). The points used in the interpolation were a subset of flight data, with a point for taken when positioned changed by 10 km along the horizontal flight track or 20 mb in the vertical.

adequately captured within STILT + VPRM framework: global scale and far-field effects on the mid-troposphere, convective redistribution, and vegetation processes (e.g. nutrient cycling, disturbance, hydrological processes) not included in the VPRM. It is clear, however, that even perfect optimization of the surface flux model could not perfectly match observations due to the strong influence of transport error, for example, due to errors in the boundary layer height.

Integration of concentrations to a fixed height may provide a better measure of regional fluxes in the model, since these are sensitive to regional surface fluxes but relatively insensitive to details of regional atmospheric dynamics (see, e.g. Chou et al., 2002). In COBRA-Maine, the observed and STILT + VPRM values were typically very close. For example, in Fig. 7, mean CO_2 observed above 2 km was 378.94 versus 381.0 from STILT + VPRM, and below 2 km, 376.4 ppm versus 376.3, respectively. The difference between CO_2 below and above 2 km, a key measure of model performance, was 2.5 ppm observed versus 3.7 ppm modelled, typical of the level of agreement we observed.

Aircraft data allow us to assess directly the representation error incurred by model treatment of measurements at Argyle as the mean in the regional planetary boundary layer, an important prior uncertainty for inclusion within a Bayesian optimization. We expect a small daytime gradient to develop in CO_2 concentration within the boundary layer due to biospheric uptake, potentially leading to a biased underestimation of mean PBL concentrations at the relatively low tower measurement height. Of the over 120 vertical profiles flown during COBRA-Maine within 50 km of Argyle, manual examination of individual profiles of water vapour, potential temperature, CO_2, ozone, and turbulence showed a clearly discernable inversion marking the top of the afternoon mixed layer in 4a cases. For each of those cases, we found the average CO_2 concentration measured in situ by the aircraft as it vertically traversed the boundary layer and compared it to the value at Argyle tower. The Argyle observations at 107 m underestimated the mean boundary layer CO_2 by 0.89 ppm on average (Fig. 8). Observations at 25 m underestimated the mean PBL concentration by much more, 1.49 ppm on average. Overall, this represents a relatively small

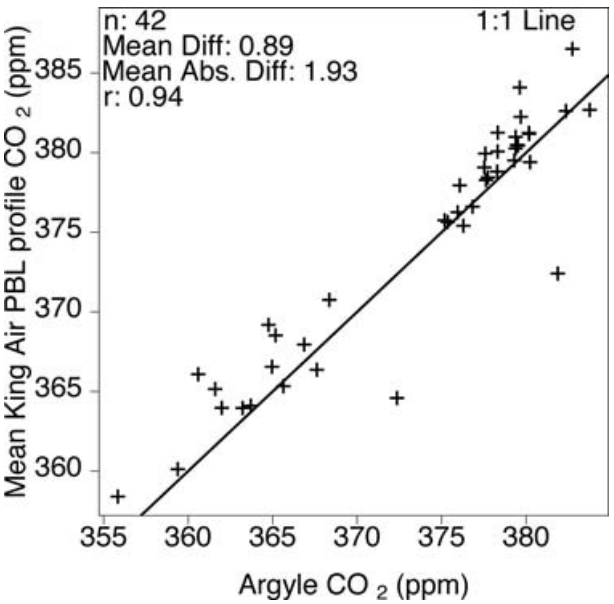

Fig. 8. Comparison of CO_2 concentration observed at 107 m on the NOAA CMDL Argyle tall tower and mean boundary layer CO_2 concentration, determined from COBRA-Maine airborne data. Each point represents the mean concentration below the top of the mixed layer measured by the aircraft as it ascended or descended through the boundary layer within 50 km of Argyle versus the mean tower concentration for the same time period. The top of the mixed layer was determined manually from vertical profiles of water vapour, potential temperature, ozone, and turbulence.

potential source of uncertainty compared to others discussed in Gerbig et al. (2003b). However, to our knowledge, this is the most extensive systematic comparison of tower measurements to mean boundary layer values obtained by aircraft to date, with many more data points than previous studies (Bakwin et al., 2003). It reinforces that tall towers are less prone to influence from local concentration gradients and more representative of a larger surrounding area than short ones, but systematic differences of order 1 ppm are likely even at tall towers.

Airborne in-situ concentration measurements collected in the free troposphere above the top of the mixed layer are useful in evaluating the advected continental boundary condition value that provides the background upon which the concentration changes due to surface fluxes are imposed. In order to assess potential error in upstream boundary condition, Gerbig et al. (2003b) evaluated the mean difference between the empirical boundary condition developed from vertical profiles in the Pacific and the accumulated library of corresponding Pacific airborne measurements, obtaining 0.22 ppm, with a standard deviation of 1.15 ppm. Latitudinal gradients in the Pacific boundary condition were not prominent with values around 3 to 4 ppm between 10° N and 72° N at most altitudes throughout 2004. In August and September, the potential bias in our treatment of the boundary condition was as large as 1 ppm. Figure 9 shows a comparison of the mean advected boundary condition at Argyle used in STILT + VPRM calculations to in-situ mean free tropospheric (>3 km) CO_2 measured by the King Air for 59 profiles

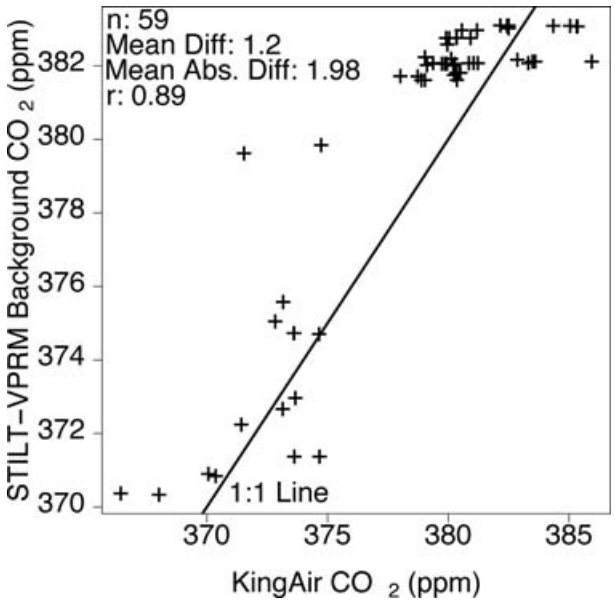

Fig. 9. Comparison of mean free troposphere CO_2 concentration determined from COBRA-Maine airborne observations and corresponding background value at Argyle tower, determined from the advected boundary condition in the receptor-oriented model (Section 2.2). Each point represents one aircraft profile within 50 km of Argyle.

from COBRA-Maine. The mean difference is 1.2 ppm, which is of the same order as would be expected from the analysis in Gerbig et al. (2003b), and the same order as the tower-PBL gradient. These results are applied to $\mathbf{S}_{veg}$ (see eq. (13); Section 2.7) in Bayesian optimization below.

Two types of errors due to transport contribute most strongly to potential errors in the advected boundary condition. First, because we assume a boundary condition in the central Pacific, the small fraction of particles leaving the STILT domain to north (or south) are given values fit to north (or south) Pacific concentrations, where values in the Arctic or subtropics are needed. These are not available due to the scarcity of observations. Second, because EDAS-40 does not include convective transport, there is the potential for divergence between modelled background and observed free troposphere due to misrepresentation of convective influence.

3.4. Bayesian optimization incorporating tower concentration observations

The Argyle data set provides a constraint for powerful inversion techniques, such as Bayesian inversion or Ensemble Kalman Filtering to further optimize the VPRM model. Since the VPRM already incorporates knowledge about the functional dependence of carbon exchange through the AmeriFlux eddy-covariance data, STILT + VPRM already goes beyond a straightforward top-down estimation of fluxes. The Bayesian optimization can extend bottom-up extrapolation of eddy correlation fluxes by imposing large-scale constraints from tall tower data. We, therefore, undertook an inverse analysis below, exploratory in nature, using the single tower data set from Argyle within the STILT + VPRM framework to assess the information available from eddy covariance and tall tower concentration data.

We performed a Bayesian optimization for scaling factors γ and ρ for each vegetation class (Table 3). This analysis used 863 daytime hours (excluding hours with missing Argyle data) during the growing season, 15 May to 15 September 2004. The atmospheric data provide the most information for the mixed and deciduous forest vegetation classes as expected from the STILT footprint in Fig. 4. In all cases, a posteriori scaling factors are relatively close to unity, consistent with the relatively good performance of the input VPRM. Prior to optimization, the root mean square error (RMSE) of STILT + VPRM CO_2 concentrations from the observed concentrations for all daytime hours was 7.31 ppm. Optimization improved the RMSE to 5.79 ppm, a 20% reduction.

Errors in computed concentrations reflect errors in the a priori values of γ and ρ for each class, which can in turn be interpreted as errors in the initial calibration factors used to fit the VPRM. These are mixed with errors in the footprint calculation and other model errors. Applying the γ and ρ factors to the STILT + VPRM results shown in Figs. 2 and 3 increases correlation of model results to observations, but changes the

Table 3. STILT + VPRM Bayesian optimization based on Argyle tower concentration data

	A priori scaling factor	A posteriori scaling factor	A priori uncertainty	A posteriori uncertainty	% reduction in uncertainty
$\gamma_{\text{borealevergreen}}$	1	0.96	0.018	0.015	16.7
$\rho_{\text{borealevergreen}}$	1	0.96	0.112	0.073	34.8
$\gamma_{\text{wettemperateevergreen}}$	1	1.01	0.013	0.011	15.4
$\rho_{\text{wettemperateevergreen}}$	1	1.24	0.060	0.037	38.3
$\gamma_{\text{drytemperateevergreen}}$	1	1.02	0.022	0.022	0.0
$\rho_{\text{drytemperateevergreen}}$	1	0.94	0.072	0.069	4.2
$\gamma_{\text{subtropicalevergreen}}$	1	0.96	0.014	0.012	14.3
$\rho_{\text{subtropicalevergreen}}$	1	1.21	0.045	0.036	20.0
$\gamma_{\text{deciduous}}$	1	0.71	0.023	0.009	60.9
$\rho_{\text{deciduous}}$	1	0.42	0.344	0.017	95.0
$\gamma_{\text{mixedforest}}$	1	0.67	0.016	0.007	56.3
$\rho_{\text{mixedforest}}$	1	0.79	0.148	0.020	86.5
$\gamma_{\text{shrubland}}$	1	1.00	0.015	0.015	0.0
$\rho_{\text{shrubland}}$	1	0.89	0.068	0.063	7.4
γ_{savanna}	1	0.99	0.015	0.015	0.0
ρ_{bsavanna}	1	1.02	0.076	0.076	0.0
γ_{cropland}	1	0.86	0.030	0.020	33.3
ρ_{cropland}	1	0.50	0.170	0.052	69.4
$\gamma_{\text{grassland}}$	1	0.92	0.027	0.025	7.4
$\rho_{\text{grassland}}$	1	0.08	0.212	0.148	30.2

Bayesian optimization of γ and ρ scalar multipliers for VPRM gross ecosystem exchange and respiration, respectively, for 10 vegetation classes using Argyle tower concentration measurements. Constraints were hourly daytime (1400–2100 GMT) CO_2 concentration observations, over 114 d (N = 863), computed from STILT–VRPM, May 15 to September 15, 2004. In the a priori case $\gamma = 1$ and $\rho = 1$.

overall patterns only slightly. This indicates the prior values (i.e. the initial calibrations to eddy covariance data) carry significant weight in the optimization. Further, deviations are mostly caused by processes not adequately captured within STILT + VPRM framework, especially transport error.

Atmospheric concentration data do not provide significant constraints for scaling factors from either dry temperate evergreen or wet temperate evergreen vegetation classes, those that Table 1 shows have the next strongest influence on Argyle, after deciduous forest and mixed forest. Within the mixed forest and deciduous forest, atmospheric concentration data provides strong constraints on two scaling factors and only moderate constraints on the other two. Hence, Argyle data alone are insufficient to fully constrain the VPRM model, even where only two vegetation classes have major influence. In order to fully optimize the VPRM for New England and southern Quebec, additional tower sites and/or inclusion of aircraft data will be needed.

4. Conclusions

We have used STILT + VPRM, a model-data assimilation framework with a minimum number of parameters, to estimate terrestrial carbon flux with high temporal resolution on regional to continental scales. Results from the VPRM, using parameters derived from AmeriFlux eddy covariance data, showed good agreement with CO_2 concentrations at Argyle tall tower and in COBRA-Maine aircraft surveys, for simulations in which there was no adjustment of parameters to fit atmospheric concentration data. This result shows that STILT + VPRM is fully representative of the regional scale, rather than just the local scale. The framework successfully takes eddy flux data from local scales, and derives the emergent functional dependence of flux on environmental conditions to capture surface heterogeneity and to interpret data for carbon fluxes at scales previously inaccessible from global inversions or local-scale eddy covariance measurements. Success of STILT + VPRM hinges on the capability for STILT to link regionally representative point concentration data to upwind sources with high spatiotemporal resolution, plus the capability for the VPRM to capture the rich spatial and temporal complexity of CO_2 fluxes using remote sensing and eddy flux data.

The extremely simple mathematical structure of the VPRM enables efficient optimization of scaling parameters using tall tower or aircraft concentration data. An exploratory Bayesian optimization demonstrated the large amount of information necessary to reasonably constrain even a relatively simple biosphere model such as the VPRM, as well as the critical importance of transport errors.

Our results imply a conservative approach to estimating the amount of information any single tower within a network can provide for carbon accounting. A network may need a denser

distribution of towers than footprint calculations might suggest in order to properly attribute CO_2 changes to different vegetation, because the influences of various vegetation types are commingled by atmospheric transport and thus are not readily separated in an inverse analysis. The Bayesian analysis gives a complete set of constraints, but there is significant weight for the a priori fluxes. We do not believe that the need for a dense tower network is an artifact of the STILT analysis, because our Lagrangian framework has very high-resolution and low artificial numerical dispersion (Lin et al., 2003), and we can reproduce reasonably well the available data set even from aircraft (see above); it appears to be a basic limitation of inversions based on data from tall towers, even parameter inversions. Aircraft data and/or towers with overlapping footprints in space and vegetation class will be needed to provide reliable regional CO_2 budgets.

Proper quantification of the uncertainty within the model-data fusion framework is as important as the flux estimation itself (Raupach et al., 2005). There are limits on how many parameters any single set of atmospheric data might constrain and there are unavoidable covariances between CO_2 fluxes, weather, and tower footprints. Hence, effective inversion requires an accurate and reliable a priori flux models that capture the functional dependence of CO_2 exchange on environmental conditions, such as the calibrated VPRM model presented here.

5. Acknowledgments

We thank the crew from the University of Wyoming King Air, with particular appreciation for the professionalism of pilots Tom Drew and Don Cooksey. We gratefully acknowledge NOAA GMD scientists Tom Conway, Paul Novelli, Colm Sweeney and Patricia Lang for providing the remote station and aircraft profile data we used in updating the continental boundary condition to 2004. COBRA-Maine was made possible by a grant from the National Science Foundation Biocomplexity in the Environment Program (ATM-0221850). D. Matross is supported by a NASA Earth System Science Fellowship (NGT5–30495).

References

Andres, R. J., Marland, G., Fung, I. and Matthews, E. 1996. A $1^\circ \times 1^\circ$ distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990. *Glob. Biogeochemical Cycles* **10**, 419–429.

Bakwin, P. S., Tans, P. P., Hurst, D. F. and Zhao, C. 1998. Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program. *Tellus* **50B**, 410–415.

Bakwin, P. S. Tans, P. P., Stephens, B. B., Wofsy, S. C., Gerbig, C. and co-authors. 2003. Strategies for measurement of atmospheric column means of carbon dioxide from aircraft using discrete sampling. *J. Geophys. Res.* **108**, D4514, doi:10.1029/2002JD003306.

Bakwin, P. S., Davis, K. J., Yi, C., Wofsy, S. C., Munger, J. W. and co-authors. 2004. Regional carbon dioxide fluxes from mixing ratio data. *Tellus* **56B**, 301–311.

Chou, W. W., Wofsy, S. C., Harriss, R. C., Lin, J. C., Gerbig, C. and co-authors. 2002. Net fluxes of CO_2 in Amazônia derived from aircraft observations. *J. Geophys Res.* **107**(D22), 4614, 10.1029/2001JD001295.

Ebel, A., Friedrich, R. and Rodhe, H. (eds.) 1997. *Transport and Chemical Transformation of Pollutants in the Troposphere*, **vol. 7**, *Tropospheric Modelling and Emission Estimation*. Springer Verlag, New York.

Errico, R. M. 1997. What is an adjoint model? *Bull. Am. Meteorological Soc.* **78**, 2577–2591.

Friedlingstein, P., Dufresne, J. L., Cox, P. M. and Rayner, P. 2003. How positive is the feedback between climate change and the carbon cycle? *Tellus* **55B**, 692–700.

Fung, I. Y., Doney, S. C., Lindsay, K. and John, J. 2005. Evolution of carbon sinks in a changing climate. *Proceedings Nat. Acad. Sci.* **102**, 11 201–11 206.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E. and co-authors. 2003a. Towards constraining regional scale fluxes of CO_2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, *J. Geophys. Res.* **108**, D4756, 10.1029/2002JD003018.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E. and co-authors. 2003b. Towards constraining regional scale fluxes of CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor oriented framework. *J. Geophys. Res.* **108**, D4757, doi:10.1029/2003JD003770.

Gerbig, C., Freitas, S. R., Lin, J. C., Longo, M., Matross, D. M. and co-authors. 2005. Consistent formulation of convective transport in a Lagrangian particle model. *European Geosciences Union General Assembly 2005*. 24–29 April, 2005. Vienna, Austria.

Gloor, M., Fan, S. M., Pacala, S., Sarmiento, J. and Ramonet, M. 1999. A model-based evaluation of inversions of atmospheric transport, using annual mean mixing ratios, as a tool to monitor fluxes of nonreactive trace substances like CO_2 on a continental scale. *J. Geophys. Res.* **104**, 14 245–14 260.

Gloor, M., Bakwin, P., Hurst, D., Lock, L., Draxler, R. and co-authors. 2001. What is the concentration footprint of a tall tower? *J. Geophys. Res.* **106**, 17 831–17 840, doi:10.1029/2001JD900021.

Grell, G. A. and Devenyi, D. 2002. A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Letters* **29**, doi:10.1029/2002GL015311.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D. and co-authors. 2002. Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models. *Nature* **415**, 626–630.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Pak, B. C. and co-authors. 2004. Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks. *Global Biogeochemical Cycles* **18**, GB1010, doi:10.1029/2003GB002111.

Helliker, B. R., Berry, J. A., Betts, A. K., Bakwin, P. S., Davis, K. J. and co-authors. 2004. Estimates of net CO_2 flux by application of equilibrium boundary layer concepts to CO_2 and water vapor measurements from a tall tower. *J. Geophys. Res.* **109**, D20106, doi:10.1029/2004JD004532.

Huete, A. R., Liu, H. O., Batchily, K. and Leeuwen, W. V. 1997. A comparison of vegetation indices global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **59**, 440–451.

Kaminski, T. and Heimann, M. 2001. Inverse modeling of atmospheric carbon dioxide fluxes. *Science* **294**, 5541.

Law, B. E., Falge, E., Gu, L., Baldocchi, D. D., Bakwin, P. and co-authors. 2002. Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation. *Ag. and Forest Met.* **113**, 97–120, doi:10.1016/S0168-1923(02)00104-1.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C. and co-authors. 2003. The Stochastic Time-Inverted Lagrangian Transport Model (STILT): quantitative analysis of surface sources from atmospheric concentration data using particle ensembles in a turbulent atmosphere. *J. Geophys. Res.* **108**, D4493, doi:10.1029/2002JD003161.

Lin, J. C. and Gerbig, C. 2005. Accounting for the effect of transport errors on tracer inversions. *Geophys. Res. Lett.* **32**, L01802, doi:10.1029/2004GL021127.

Lin, J. C., Gerbig, C., Wofsy, S. C., Daube, B. C., Matross, D. M. and co-authors. 2006. What have we learned from intensive atmospheric sampling field programs of CO_2? *Tellus* **58B**, this issue.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z. and co-authors. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Intl. Journal Remote Sensing* **21**, 1303–1330.

Marland, G., Boden, T. A., Brenkert, A. L., Andres, R. J. and Olivier, J. G. 1997. CO_2 from fossil fuel burning: Updates on the magnitude, distribution, and uncertainty of emission estimates. In *Fifth International Carbon Dioxide Conference*, Cairns, Queensland, Australia, 4.

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C. and co-authors. 2004. The multi-institution North American Land Data Assimilation System (NLDAS): utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.* **109**, D07S90, doi:10.1029/2003JD003823.

Pathmathevan, M., Wofsy, S. C., Matross, D. M., Xiao, X., Lin, J. C. and co-authors. 2006. A satellite-based biosphere parameterization for net ecosystem CO_2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM). http://people.deas.harvard.edu/šwofsy/VPRM˙submitted.zip

Raupach, M. R., Rayner, P. J., Barrett, D. J., Defriess, R. S., Heimann, M. and co-authors. 2005. Model-data synthesis in terrestrial carbon observation: method, data requirements, and data uncertainty specifications. *Global Change Biology* **11**, 378–397, doi:10.1111/j.136502486.2005.00917.x.

Rodgers, C. D., 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Sci., River Edge, New Jersey, 238 p.

Rolph, G. D. 1997. NOAA NCEP Eta Data Assimilation System (EDAS) archived by NOAA Air Resources Laboratory (The EDAS Archive). U.S. Department of Commerce, NOAA Air Resources Laboratory (ARL).

Xiao, X. M., Zhang, Q. Y., Braswell, B., Urbanski, S., Boles, S. and co-authors. 2004. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sensing of the Environment* **91**, 256–270.